Validity and the Research Process

David Brinberg
Joseph E. McGrath

SAGE PUBLICATIONS Beverly Hills London New Delhi

For information address:

SAGE Publications, Inc.
275 South Beverly Drive
Beverly Hills, California 90212

SAGE Publications India Pvt. Ltd.
M-32 Market
Greater Kailash I
New Delhi 110 048 India

SAGE Publications Ltd
28 Banner Street
London EC1Y 8QE
England

Printed in the United States of America

Library of Congress Cataloging in Publication Data

Brinberg, David.
Validity and the research process.

Bibliography: p.
Includes index.
1. Social sciences—Research—Methodology.
I. McGrath, Joseph Edward, 1927- II. Title.
H62.B662 1985 300'.72 85-1903
ISBN 0-8039-2302-1

FIRST PRINTING

Validity and the Research Process

Contents

Preface

Validity is a much used term in the behavioral and social sciences. In recent years, however, we have been impressed by how heterogeneously that term is used. Moreover, each of those many uses of validity seems to be free-floating—not related to one another and not tied to the research process, either. The discussion about validity issues has focused on identifying strategies for increasing the validity in empirical findings. Such strategies have been discussed as if they amounted to a set of procedures that, if followed, would increase validity. In other words, those strategies are regarded as algorithms for "getting" validity into your study.

This book is partly a reaction to such diverse and separate uses of the concept of validity, and partly a reaction to the view of validity as something to be acquired by diligent application of certain techniques. We constructed a schema that lays out the research process in some detail and complexity, and presents the logical relations of different aspects of validity to different stages within that research process. This book is about that schema, which we call the Validity Network Schema, or VNS.

The VNS is very complex and abstract. In developing it, we were forced to address issues in epistemology and other areas of the philosophy of science. It proved to be all too easy to get very far afield from our initial goal. So we have specified some limits for what we will and will not try to deal with in this book.

First, we will not examine epistemological issues in depth, but only discuss those issues briefly, insofar as needed to describe our schema. Therefore, we want to note here our position on several philosophical issues. On the issue of reality, we both lean heavily toward a position of "hypothetical realism" (see Campbell, 1981; Brewer and Collins, 1981). That position, in essence, says that we do not know whether there is really a "real world." We are confident that, if there is a "real world," we can know it only "through a glass, darkly." But at the same time, we believe that the *underjustified presumption* of the existence of a real world makes sense to use for those intellectual endeavors we call science, just as it does for our everyday experience. We also follow the view of Campbell and colleagues (see Campbell and Fiske, 1959; Campbell and Stanley, 1966; Cook and Campbell, 1979) regarding the value of searching for convergence among multiple operations, as the main means by which we can increase confidence in our findings.

We think validity is a matter of logical possibilities as well as a matter of empirical outcomes. We will not try to develop or assess quantitative indices of various validities, nor statistical or mathematical models for them in this book. For some aspects of validity, development of such quantitative indices is an important task; but that is not what this book is about. For other aspects of validity, as we view the matter, the key issues are logical, intuitive, and deductive, and are not amenable to resolution on grounds of quantitative evidence. For them, seeking quantification is inappropriate.

The VNS is a relatively general system that could be used to provide a conceptualization of the research process for a wide variety of fields. In this book, however, we will limit our direct consideration to the behavioral and social sciences. Furthermore, we will give most attention to, and draw examples from, those substantive areas about which we have some expertise: social psychology, consumer behavior, and organizational research.

Our intention is to lay out and clarify what amounts to a meta-theory of the research process. We will present material at several levels of detail in some places, but we will not try to be completely detailed in all parts of the schema. We will suggest many possible issues that can be examined using the VNS, but we

will not try to be comprehensive in those suggestions. Throughout the presentation we will use examples that illustrate our points, but we will not carry specific examples of research projects or findings through the whole VNS. For instance, we will suggest in Chapter 6 that the VNS may provide a useful tool to help conduct meta-analyses of bodies of literature, but we will not carry through such a meta-analysis for any specific findings.

Finally, the VNS contains many distinctions and relations between terms. In developing such a schema (as the reader doubtless knows from personal experience), there is a strong Procrustean force toward "neatness" and fit. We have tried to resist and guard against that force, by a variety of techniques for providing intellectual "checks and balances." But we are confident that we have succeeded only in part. The reader is warned—and encouraged—to transform or extrapolate parts of the schema as is useful.

A number of colleagues helped us in our thinking and writing about the VNS. We would especially like to thank Irwin Altman, Marta Axelson, Donald Fiske, J. Richard Hackman, Elizabeth Hirschman, James Jaccard, and John Lynch for their useful insights and encouraging feedback. We also want to express our appreciation for the feedback we received on earlier versions of these materials from several classes of students at the University of Illinois, Urbana, University of Maryland, and Baruch College. Finally, we want to express our gratitude for the opportunity to spend several summers as fellows at the Baldwin Research Institute, where most of the ideas for this book were generated.

David Brinberg
Joseph E. McGrath

Chapter 1

INTRODUCTION

Validity is not a commodity that can be purchased with techniques. Validity, as we will treat it, is a concept designating an ideal state—to be pursued, but not to be attained. As the roots of the word imply, validity has to do with truth, strength, and value. The discourse of our field has often been in tones that seem to imply that validity is a tangible "resource," and that if one can acquire a sufficient amount of it, by applying appropriate techniques, one has somehow "won" at the game called research. We reject this view. In our view, validity is not like money—to gain and lose, to count and display. Rather, validity is like integrity, character, or quality, to be assessed *relative to purposes and circumstances.*

In modern social science, validity has been given a variety of meanings: convergence, correspondence, differentiation, equivalence, generality, repeatability, and some others. Many of these meanings apply, we think, within different stages of the research process. But too often the various aspects of validity have been treated as if they were separate and free-floating—unconnected to one another or to the research process.

In this book we try to integrate these diverse and separate uses of the concept of validity. We have constructed a schema that lays out the research process and considers the logical relations of different aspects of validity to different phases or stages within that process. Recently, we published a skeletal version of it

(Brinberg and McGrath, 1982; McGrath and Brinberg, 1983; McGrath and Brinberg, 1984), calling it a Validity Network Schema (or VNS). In this book we present that Validity Network Schema in full detail, elaborate many of its parts, and try to show how use of that schema can help the behavioral and social science researcher to understand and carry out research. We have provided a glossary of terms for the VNS, which contains terms first introduced in the text in boldface.

In our view, many kinds of research activities play a part in the knowledge generating enterprise that we call behavioral and social science research. In the VNS, we lay out the research process in terms of several domains, levels, and stages, and describe a number of alternative paths through that process. Those paths reflect different styles of, orientations toward, and purposes for doing research. We do not argue that one of those paths is the preferred one or is better than the others. Quite to the contrary! Any one path is limited in what we can gain from it. Every path is flawed in what we can achieve with it, but each is flawed in different ways. We will argue that the full research endeavor requires pursuit of multiple paths.

SOME BASIC ASSUMPTIONS ABOUT RESEARCH

Research involves (a) some content that is of interest, (b) some ideas that give meaning to that content, and (c) some techniques or procedures by means of which those ideas and content can be studied. We will later call these three aspects the **substantive, conceptual**, and **methodological domains**, respectively.

Research is the study of **relations**. It is always done in terms of relations between some units (later we will call the units **elements**) that are proper parts of some surrounding context (later we will call these contexts **embedding systems**) within which both the units and the relations are embedded.

Doing research does not involve just one set of activities. For example, it is not simply obtaining some measurements (and only doing that). We argue here that the total research process is extensive and complex, and can be divided into three major stages. The **central stage** is the part of the process that is usually meant when we talk about "doing a study": combining some content, some ideas, and some techniques to arrive at some "findings." But to do this part, the researcher must somehow come by those ideas, contents, and techniques. We argue that this central stage, "doing a study," is stage two of the research process and must be preceded by much prior work. That work involves generation, identification, development, and clarification of concepts, methods, and substantive phenomena. It is usually done by people other than those who will later be "doing a study." In the VNS, we call that work stage one, or the **prestudy stage.**

After generating a set of findings in the central stage, there is still a need to explore the scope and limits of those findings before they can have meaning as part of a body of knowledge. These activities—exploring the scope and limits of a set of findings—are stage three, or the **follow-up stage** of research.

In the central stage there are several different styles or approaches by which one can do a study. We will deal with three study paths: the **experimental path**, the **theoretical path**, and the **empirical path**. We also will deal with the following three orientations to research: **basic research** focuses on conceptual issues; **applied research** focuses on some substantive system; and **technological research** focuses on methodological matters. All these paths and orientations are legitimate and valuable parts of the overall research process.

AN OVERVIEW OF THE VALIDITY NETWORK SCHEMA

The Validity Network Schema (VNS) starts with certain assumptions and primitive terms:

(1) ***Domains.*** Research involves three interrelated but analytically distinct domains: the conceptual, the methodological, and the substantive.

(2) ***Levels.*** With respect to each of these domains, research involves relations between elements within an embedding system. The nature of the elements, relations, and embedding system differs among the domains.
(3) ***Stages.*** The complete research process involves three major stages, each with several paths.
(4) ***Validity.*** The concept of validity takes on fundamentally different meanings within each of the three stages.
(5) ***Paths.*** There are three alternative paths for carrying out the central stage of the research process. These paths reflect different styles of doing research and encounter different validity issues.

In our Validity Network Schema, all research involves the combination of some set of concepts, some set of methods for making observations and comparing sets of observations, and some set of substantive events that are to be the focus of study. *The research process is the identification, selection, combination, and use of elements and relations from the conceptual, methodological, and substantive domains.* Different research areas deal with different portions of the substantive domain and different research approaches use elements and relations from different portions of the conceptual and methodological domains. But any given research study makes use of some set of elements and relations from each of the three domains.

Domains and Levels

The substantive domain refers to the real-world systems and phenomena that are the focus of research. The conceptual domain refers to ideas that are abstract representations of aspects of such substantive phenomena. The methodological domain refers to techniques that are the means by which we gather, process, and interpret information about substantive phenomena. Within each of the domains, we are concerned with relations among elements within an embedding system.

The substantive domain. Elements are **phenomena**, which are states and actions of entities (agents and objects, human and

nonhuman). Relations are **patterns of phenomena** (relations between two or more states or actions of entities). The embedding system refers to the **substantive systems**, at higher levels of organization, that provide the temporal, locational, and situational context within which the entities and their states and actions are embedded.

The conceptual domain. Elements are attributes or **properties of phenomena** (properties of states and actions of entities, events, and contexts from the substantive domain). Relations refer to the set of **conceptual relations** that can specify the form of the logical-causal-temporal pattern among two or more properties. The embedding system refers to what might be called the set of paradigmatic assumptions, or the **conceptual paradigm**, within which those properties and relations are specified and studied.

The methodological domain. Elements are **methods or modes of treatment** (techniques for gathering information about phenomena). Relations are **comparison techniques** (procedures for comparing or assessing the patterns of relations among two or more phenomena). The embedding system refers to **research strategies** within which methods and comparison techniques are executed.

Stages

In the VNS, the complete research process is divided into three stages. Stage one is a preparatory stage. It involves generation, development, clarification, and evaluation of elements and relations within each of the three domains. It is necessary groundwork that must be done before stages two and three can be carried out. Stage one is the part of the research process that most nearly reflects the generative or constructionist paradigm for research.

Stage one researchers are system experts in various systems within each of the three domains. Some are experts in various kinds of methodology, some in various conceptual systems, and some in various substantive systems. Usually, different sets of people do stage one research in each of the three domains. Furthermore, these are often different people than those who do stage two research. Sometimes, though, stage one work is actually the result of stage two research of the past, that was thought to be especially definitive with respect to one of the domains. Stage one is the topic of Chapter 2.

Stage two in the VNS is the central stage of the research process. It is the part we usually mean when we talk about "doing a study." It involves developing sets of empirical findings by combining some subset of elements and relations from each of the three domains. It is stage two that best reflects the logical-empiricist or hypothetico-deductive paradigm for research.

Stage two involves three steps. The first step involves choosing the elements and relations from one of the domains, the one of central interest to the researcher. The second step has to do with combining elements and relations from that domain with those from a second domain, to form an intermediate or instrumental structure. The third step has to do with bringing elements and relations from the third domain into that structure, thereby generating a **set of empirical findings**. With three domains, there are three places to start, reflecting three different orientations to research. We call these orientations *basic research, applied research*, and *technological research* for efforts that start in the conceptual, substantive, and methodological domains, respectively.

There also are three different combinations of two domains that can be combined into the instrumental structures that are the result of the first two steps. Those three ways constitute three different paths or styles for conducting stage two research. We call these the *experimental path*, the *theoretical path*, and the *empirical path*. They are discussed later in this chapter. The orientations and paths differ in terms of which domain is emphasized, which is given lowest priority, and what validity issues get addressed. All three paths result in a *set of empirical findings*. Stage two is the topic of Chapter 3 and Chapter 4: Chapter 3 deals with paths and orientations, and Chapter 4 deals with validity as fit.

Stage three involves following up the findings of stage two, by replication and by a systematic search for both the range and the boundaries of those findings. That search, we argue, needs to be done in relation to all three domains. The purpose of stage three activity is to *verify, extend, and delimit* some particular set of stage two findings. As stage one is "generative" or "constructive," and stage two is "logical-empiricist" or "hypothetico-deductive," stage three most nearly reflects a "generalizability" or "credibility" paradigm for research.

The three stages are related to one another in complex ways. Stage two is what we most often think of when we talk about a research study. It has reference to research on a specific problem, which we will here call the *focal problem*. The focal problem is defined by the substantive phenomena, concepts, and methods that are used in stage two activities. Work in stage one is preparatory for the study of many problems, of which any given stage two focal problem is but one. Work in stage three explores the range and limits of some set of findings from stage two work on some focal problem, both by using the same sets of concepts, phenomena, and methods (i.e., by replication) and by systematically varying certain facets of those concepts, methods, and phenomena (i.e., by searching for the scope and limits of those findings). Stage three is the topic of Chapter 5.

Validities

Validity has different meanings in the three stages of the research process. In stage one, validity means **value** or worth. People within each of the domains working in stage one identify, develop, and clarify elements and relations that they consider to be "of value"; that is, those they consider important, meaningful, or useful. Researchers engaged in stage two activities—very often not the same persons who did the stage one work—usually have special interest in one domain and less interest, and indeed less expertise, in one or both of the other two domains. Often, stage two researchers accept without much thought the elements and relations that are currently available from the stage one work of others, for at least one and often two of the domains. In other words, stage two researchers borrow, often rather casually, from substantive system experts, conceptual system experts, and methodological system experts. Inevitably, there are underlying assumptions and constraints built into any set of elements and relations that might be selected for use from any one of the domains. These assumptions and constraints often are not fully appreciated by the stage two researcher who uses those elements and relations. We will call the set of validity issues encountered in stage one the **valuation validities**.

In stage two, validity means **correspondence** or fit. In the VNS, the fit has to do with fit between relations among the elements from the three domains. The second step in stage two involves the fit between relations from two of the domains. The third step involves the fit between the instrumental structure formed in step two and relations from the third domain. There are three ways to combine things from three domains. Therefore, there are three routes through stage two, the three research paths or styles noted earlier. Questions of fit arise at steps two and three for each of the three paths. They all involve the underlying notion of correspondence or fit. We term them, collectively, the **correspondence validities**.

In stage three, validity means **robustness, generalizability**, or so-called external validity. Stage three activities have to do with increasing our confidence concerning the interpretation of a certain set of stage two findings. Work in stage three has to do with reducing our uncertainty about the range of variation—of substantive, conceptual, and methodological elements and relations—over which the stage two findings will and will not hold. Stage three validity issues are all related to the idea of generalizability or robustness. We call these issues, collectively, the **generalization validities**.

Alternative Study Paths

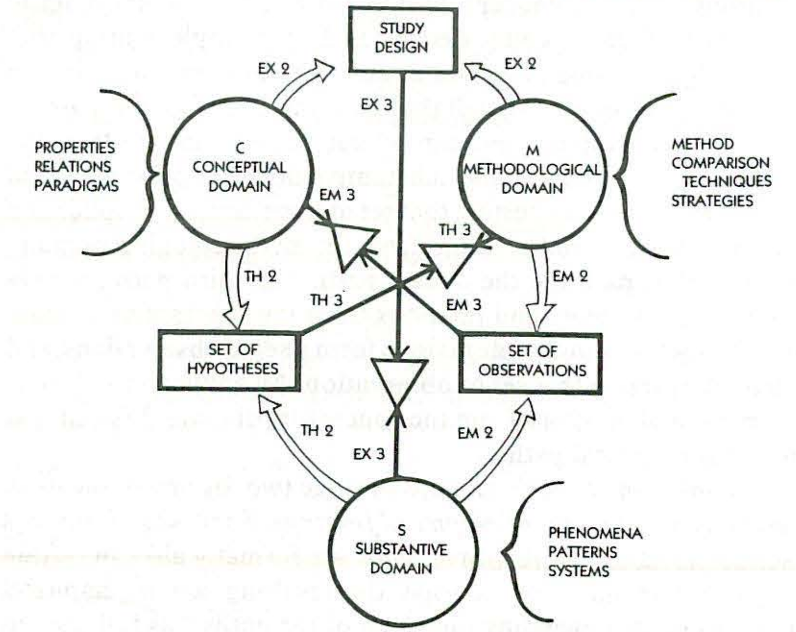
The three alternative paths of stage two constitute distinct research styles. One of them involves combining elements and

relations from the conceptual domain and the methodological domain to form a **study design**, and then implementing that design by applying it to some elements and relations in the substantive domain. We call that path the experimental path. A second path involves combining elements and relations from the conceptual domain and the substantive domain to form a **set of hypotheses**, and then testing that set of hypotheses by application of some elements and relations from the methodological domain. We call that path the theoretical path. The third path involves combining elements and relations from the substantive domain and the methodological domain to form a **set of observations**, and then interpreting that set of observations by application of some elements and relations from the conceptual domain. We call that path the empirical path.

Completion of the three steps of stage two, by any of the three paths, results in a *set of empirical findings*. Such sets of findings derived from any of the three paths are formally alike in certain ways, but in many other ways the resulting sets of empirical findings differ depending on which of the paths was followed in attaining them. Similarities and differences among the results of these three paths, and some further complications that arise by way of different orientations or approaches to research, are discussed in Chapter 3.

Some General Features of the VNS

We see the research process as complex and multifaceted, with various stages, steps, paths, and levels. Alternative paths through the process are equally valuable. There is no one "correct" path. Nor is any single path, used alone, sufficient. Multiple paths are *essential*. Each path has weaknesses. Different paths have different weaknesses. Using multiple paths to study a single problem allows for convergence across paths, each of which is fallible, thus permitting their different potential strengths to offset their differing inherent weaknesses. Furthermore, the various paths through the research process encounter—and allow the researcher to deal with—different validity issues. When work



EXPERIMENTAL PATH: Building a design, and implementing it by using it on a set of substantive events.

THEORETICAL PATH: Building a set of hypotheses, and testing them by evaluating them with an appropriate set of methods.

EMPIRICAL PATH: Building a set of observations, and explaining them by construing them in terms of a set of meaningful concepts.

Figure 1.1: The VNS System: Domains, Levels, and Paths

has been built entirely on one path, certain validity issues will not have been addressed; hence, that work will be more limited than it need be.

In all of these matters, here and throughout the book, it is important to distinguish among several possible levels of discourse. We do not intend to make normative prescriptions about how the scientist "ought" to operate. To the contrary, we do not think there is, or could be, a "best way" to contribute to the

TABLE 1.1

The VNS System: Validities and Stages of the Research Process

Stage One: Validity as Value

(Central tasks of stage one: Identification, development; and clarification of elements, relations, and embedding systems, for each of the three domains.)

Domain	*Criteria for Evaluating Elements, Relations*
Conceptual (C)	Parsimony, internal consistency, subsumptive power, testability, etc.
Methodological (M)	Efficiency, power, unbiasedness, explicitness, reproducibility, etc.
Substantive (S)	System effectiveness, cost/benefit, feasibility, etc.

Stage Two: Validity as Correspondence

(Central tasks of stage two: selection, combination, and use of elements and relations from all three domains to produce a set of empirical findings.)

Paths	*Step 2*	*Step 3*	*Product*
Experimental (Ex)	Study design	Implementation →	
Theoretical (Th)	Set of hypotheses	Test of hypotheses →	A set of empirical findings
Empirical (Em)	Set of observations	Interpretation →	

Stage Three: Validity as Robustness

(Central tasks of stage three: verification, extension; and delineation of particular stage two findings.)

Replication:	Are the (stage two) findings reproduced when all facets of C, M, and S are kept the same?
Convergence Analysis:	Over what range (of values of all facets of C, M, and S) do the (stage two) findings hold?
Boundary Search:	Beyond what range (of values of all facets of C, M, and S) do the (stage two) findings fail to hold?

knowledge accrual process that we call behavioral and social science research.

We also do not intend to be providing descriptive accounts of how scientists do operate, either in the specific form of "life histories" of particular research endeavors or in the more general form of accounts of how experienced researchers tend to carry on their work. Providing such descriptive accounts is certainly a respectable and useful endeavor, but it is not ours.

We do intend to provide an account of the logic of the knowledge accrual process itself. Such an account requires a commitment on our part to some epistemology. Our leaning is toward Campbell's "hypothetical realism" (Brewer and Collins, 1981; Campbell, 1981). Using that viewpoint, one proceeds *as if* there were a real and knowable world beyond the phenomenological evidence of our senses, at the same time recognizing that such a belief is itself an underjustified and perhaps unjustifiable presumption.

Thus, we are not attempting to describe how behavioral and social scientists *do* act in their pursuit of knowledge, and certainly not how they *ought to* act. Rather, we are attempting to describe the requirements inherent in the knowledge accrual process itself. These requirements do not specify a set of activities per se that are in some sense a proper way to do science. Rather, they specify a set of criteria that the results of any such activities must meet in order for them to yield *knowledge in which we can have confidence*—or, to use terms we will build upon later, to reduce our uncertainty on the matters in question.

PLAN OF THE BOOK

The complex set of relations among the three domains, the levels within the domains, the three stages of the research process, and the alternative paths and orientations within stage two are all depicted in Figure 1.1 and Table 1.1. At the end of each chapter, we will include a glossary of technical terms introduced in that chapter. We used the three stages of the research process as the bases for the organization of the book. In Chapter 2, we examine stage one, the prestudy stage within which materials are developed in all three domains for later use. Chapters 3 and 4 deal with stage two. In Chapter 3, we discuss alternative study paths and research orientations. In Chapter 4, we describe an array of validity issues that arise at various sites within stage two. In Chapter 5, we discuss stage three, the follow-up or generalization stage of the research process. In the final chapter, Chapter 6, we suggest some uses, implications, and limitations of the VNS.

GLOSSARY OF TERMS

Substantive Domain: contains the phenomena, processes, or focal problems of interest.

- **Element Level: Phenomena;** that is, states and actions of entities. An entity can be a human or nonhuman agent or object. Sometimes referred to as **actors behaving in context**. For example, a particular behavior (action) of an individual (actor) in a situation (context).
- **Relations Level: Patterns of phenomena;** that is, the relation between two or more states and actions of entities; the process of interest. For example, the relation between a particular behavior (action of an individual) and a particular feature (state) of a task.
- **Embedding Systems: Substantive systems;** that is, social units at a higher level of organization in which the elements and relations are embedded; the temporal, locational, and situational context surrounding the elements and relations. For example, an organization may be the embedding system for work teams that are the focal units of study.

Conceptual Domain: contains ideas, concepts, and their relations as well as the philosophical assumptions underlying them.

- **Element Level: Properties of phenomena;** that is, the ideas and concepts used to describe and explain a phenomenon. For example, personality traits (property of individual) may be used to explain individual behavior.
- **Relations Level: Conceptual relations;** that is, the temporal, logical, and causal pattern specified among two or more concepts. For example, a researcher may specify a curvilinear relation between stress and individual performance.
- **Embedding Systems: Conceptual paradigm;** that is, the sets of philosophical assumptions within which the concepts and their relations are embedded. For example, homeostasis is an under-

lying assumption of much of the work in the behavioral and social sciences.

Methodological Domain: contains the methods, designs, and research strategies used to examine concepts and phenomena.

Element Level: Methods or modes of treatment; that is, methods for measuring, manipulating, and controlling variables (properties of phenomena) that allow the researcher to gather information about phenomena. For example, instruments to measure a person's attitude or methods for manipulating that attitude.

Relations Level: Comparison techniques; that is, techniques that allow the researcher to make comparisons or assess covariation among the variables. Researchers usually refer to this level in the methodological domain as designs (although that term takes on a more specific meaning in the VNS), for example, factorial designs, Solomon 4-group design, and cross-lagged panel design.

Embedding Systems: Research strategies in which the elements and relations are embedded. For example, laboratory experiments and sample surveys are both types of research strategies that are methodological embedding systems.

STAGES, PATHS, AND MEANINGS OF VALIDITY

STAGE ONE: Prestudy or generative stage in which a researcher develops, clarifies, and refines the elements and relations within one of the three domains—conceptual, methodological, and substantive. For example, a philosopher or model builder may be viewed as a stage one researcher in the conceptual domain; a statistician or psychometrician may be viewed as a stage one researcher in the methodological domain, and a system manager or system expert may be viewed as a stage one researcher in the substantive domain.

Stage One Validity: Value or worth; that is, the criteria that researchers use to evaluate the elements and relations within each of the domains is based on their value or worth or usefulness or importance. The set of validity issues in stage one are described as **valuation** validities. The specific criteria are discussed in Chapter 2.

STAGE TWO: Central or execution stage in which a researcher combines the elements and relations selected from each of the three domains. The outcome of stage two is a **set of empirical findings.**

Stage Two Validity: Correspondence or fit; that is, the degree to which the features of the relations a researcher is examining match across domains. The set of validity issues in stage two are called **correspondence validities**. The tasks associated with that match and the specific features of a relation are presented in Chapter 4.

Stage Two Path: Experimental path in which a researcher first combines elements and relations selected from the *conceptual* and *methodological* domain to form a **study design**, and then implements that design by selecting phenomena and patterns among phenomena from the *substantive* domain.

Stage Two Path: Theoretical path in which a researcher first combines elements and relations selected from the *conceptual* and *substantive* domains to form a **set of hypotheses**, and then test those hypotheses by applying measures and comparison techniques selected from the *methodological* domain.

Stage Two: Empirical path in which a researcher first combines elements and relations selected from the *methodological* and *substantive* domain to form a **set of observations,** and then attempts to explain those observations by selecting a set of concepts from the *conceptual domain.*

Stage Two Orientations:

Basic orientation, in which a researcher has primary interest in and concern about the elements and relations of the *conceptual* domain.

Applied orientation, in which a researcher has primary interest in and concern about the elements and relations of the *substantive* domain.

Technological orientation, in which a researcher has primary interest in and concern about the elements and relations of the *methodological* domain.

STAGE THREE: Follow-up or generalization stage in which a researcher attempts to determine (a) whether the outcome of stage two (set of empirical findings) will **replicate**, (b) whether those findings will **converge** across variations in the elements and relations from each of the three domains, and (c) what are the **boundaries** beyond which those findings will not hold.

Stage Three Validity: Robustness or generalizability; that is, the range of variation of elements and relations (from each of the domains) over which a set of findings (developed in stage two work) holds (the scope), and the boundaries (of variations in these elements and relations) beyond which the set of findings does not hold (the limits). The set of validity issues is described as **generalization** validities. These validity issues are discussed in detail in Chapter 5.

Chapter 2

STAGE ONE
Validity as Value

Stage one of the research process is preparatory for the two later stages of the process. It involves exploring and elaborating the contents of the three domains. Typically, any one researcher or practitioner is likely to have an interest in, and a talent for, development of just one of the three domains.

There is a set of people, for example, who are methodological specialists. They spend their time and energies generating, developing, and clarifying methodological tools that might be used in research in a wide variety of content areas. These tools would include methods for measurement of properties of phenomena, techniques for selection and allocation of cases, techniques for manipulation and control of variables, and techniques for aggregating and analyzing sets of empirical observations. Sometimes this methodological work is done in close conjunction with some stage two research—with an effort to develop a set of empirical findings about a particular substantive phenomenon. But often, too, the methodological specialist works solely in the methodological domain. Such efforts, to generate, develop, and clarify methods, comparison techniques, and research strategies are what we would regard as stage one work. We will call the people who do it **methodological system experts.**

Similarly, there is a set of people who spend a major portion of their time and effort attempting to generate, develop, and clarify materials from the conceptual domain: concepts or properties of potential interest, conceptual relations by which those properties are connected, and features of conceptual paradigms within which those concepts and relations are embedded. As with methodological work in stage one, some of this work is done in close conjunction with stage two efforts. But some of it also is done solely as development of the conceptual domain. We will regard generating, developing, and clarifying concepts and relations in the conceptual domain as stage one work. We will describe the people who do it as **conceptual system experts.**

There also is a set of people whose chief interest and efforts involve identifying, developing, and clarifying phenomena and patterns of phenomena within the substantive domain. Typically, such people have a strong interest only in particular substantive systems. Very often, too, the people who do this work do not think of themselves as scientists, researchers, or scholars, but rather as practitioners or system experts. As with the other domains, some of this work is done in conjunction with some stage two research. But some of it is done strictly for the sake of exploring and improving activities within the particular substantive systems. Such efforts are what we will call stage one activities. We will call the people who carry out those activities in the substantive domain **substantive system experts.**

Stage one, therefore, involves three quite disparate kinds of activities that get carried on by three quite different sets of people. The methodological specialist is clearly recognized as a researcher and is often quite highly regarded in many areas of the behavioral and social sciences. The methodologist often has training both in some content area and in a methodological speciality, often in mathematics. The most notable among these methodologists are statisticians, but many other categories would also fit: scaling experts; people who develop verbal instruments such as tests, as well as those who develop physical instruments for measuring important properties of *actors behaving in contexts*; people who develop techniques for recording behavior, hence preserving it for

later "scoring" and analysis; people who specialize in study planning or experimental design; and so forth.

Scholars who are specialists in the conceptual domain can be thought of as philosophers, logicians, or model builders, as much as behavioral or social scientists. They are likely to have had training in some substantive areas, but to put their main efforts into the development and clarification of ideas that may well span a number of specific research areas.

The specialist in any given substantive system is seldom thought of as a researcher at all, but rather as a practitioner who is an expert in the real-world systems in question. Often substantive system specialists are leaders in some business, legal, military, educational, or human service organizations. They may be specialists in some particular subset of system processes (e.g., production, policymaking, sales, marketing) or behavior processes (e.g., training, evaluation, selection, communication) that transcend specific systems and are important in many different substantive systems. To state such a list of possible system experts is to begin hinting at some of the phenomena and patterns that such system specialists might generate, develop, and clarify as they carry out the work that we will call stage one activities in the substantive domain. Such system experts could include physicians, attorneys, ministers, bankers, accountants, teachers, politicians, and industrialists.

The work of stage one, in any of these three domains, is preparatory, exploratory, and above all *generative.* It consists of finding or inventing elements and relations—concepts and conceptual relations, methods and comparison techniques, phenomena and patterns among them—that are or might be of value for stage two and stage three work.

The specification of what is "of value" in this context, is itself an important part of stage one activity. The meaning of value, in this context, differs for the different domains. Furthermore, what is "of value" in each of the domains shifts from time to time. Those shifts, themselves, are part of the process by which a field builds and changes its dominant paradigms. (Paradigms and shifts in them will be discussed in Chapter 6).

These introductory comments suggest two key questions that will provide the structure for this chapter:

(1) What are the contents of each of the three domains? What are the elements, relations, and embedding systems for substance, concepts, and methods?
(2) What is the meaning of "value" in each of the three domains? What criteria do those engaged in stage one work in methodological, conceptual, and substantive systems use to evaluate their efforts and results?

Those two key questions will be addressed, in turn, in the two main sections of this chapter. Within each section, the key question will be asked with respect to each of the three domains.

WHAT IS IN THE DOMAINS?

We are postulating a three-domain, three-level system from which the materials of the research process are drawn. The three domains are the substantive, the conceptual, and the methodological. The three levels are *elements, relations,* and *embedding system* (see Table 2.1).

Research is about relations. Therefore, within each domain, the relations level is the main focus of our interest. Elements are the *parts* of those relations. The embedding system is the *surround*, or higher-order system, within which those elements and relations are embedded (McGrath and Altman, 1966). It is the context within which the relations are interpreted. The specific content of the elements, relations, and embedding system differ for the different domains. The content for each of the three domains is the topic of this section.

Even though the primary focus, or level of reference, or unit of analysis is the relations level, with elements viewed as parts and with the embedding system viewed as surround, it is easier to talk

TABLE 2.1
Domains and Levels in the VNS System

	Substantive	*Conceptual*	*Methodological*
Element Level	Phenomena	Properties of phenomena	Mode of treatment of properties of phenomena
Relation Level	Patterns of relations among phenomena	Logical, causal, chronological relations among phenomena	Comparison techniques for modes of treatment
Embedding System Level	Substantive systems	Conceptual paradigm	Research strategies

first about elements, then about relations between those elements, and then about the embedding system within which elements and relations exist. So, to aid clarity of presentation we will discuss the elements level first, then the relations level, then the system level, as we treat each domain.

The Substantive Domain

Of the three domains, the substantive is the most elusive. In a sense, its contents are ineffable. As soon as one begins to talk about what is "in" the substantive domain, one makes use of concepts or methods or both. Furthermore, the substantive domain is in a sense more fundamental: It is what is "there" prior to and independent of the intellectual enterprise we call research.

Elements of the substantive domain are phenomena (see Table 2.2). We define phenomena as states and actions of entities. Runkel and McGrath (1972) define the basic unit of study for the social and behavioral sciences as "actors behaving toward objects in context." We will consider such units—"actors behaving toward objects in context"—as another way to describe the phenomena of the substantive domain. For the social and behavioral sciences, the entities of most concern (that is, the actors in Runkel and McGrath's terms) are various levels of social unit: individuals, groups, organizations, cultures. We are interested in the states and actions of those

TABLE 2.2
Levels of the Substantive Domain

Element Level

Elements = Phenomena
Phenomena = states and actions of entities
= "actors behaving toward objects in context"

Relation Level

Relations = patterns of occurrence of two or more phenomena

Embedding System Level

Embedding System = Substantive system at higher level of organization within which the elements and relations are embedded

entities both as "agents," and as objects of actions by others and by outside forces.

Relations in the substantive domain are *patterns* of two or more such phenomena (i.e., states or actions of entities). That is, *relations are the patterns of occurrence of two or more instances of actors behaving toward objects in context.* The embedding system, in the substantive domain, refers to those higher-order system levels within which the entities, their states and actions, and the patterns by which they are related to one another are embedded. Note that what are regarded as element level entities from one point of view might well be regarded as embedding systems from another viewpoint, and vice versa.

The relations level in the substantive domain deals with the *patterns* of occurrence of states and actions of entities. *Identifying, analyzing, and understanding such patterns, for the states and actions of the entities chosen for study from the substantive system, is the purpose of the scientific enterprise.* By definition, different fields of science differ from one another in terms of which substantive systems (i.e., which sets of phenomena) they involve. They also differ from one another in terms of which methodological and conceptual tools they use, but these differences often seem to be based more on custom than on inherent differences.

In any given case, what entities and events have been selected for study in part determines what is regarded as the embedding system, or perhaps vice versa. For example, if the focal unit chosen for consideration is "individual," then "organization" might be an important embedding system for that effort. But organization is not the only possible choice for embedding system if individual is the element level unit. Instead, certain classrooms might be appropriate embedding systems for such work, or, in general, a certain surrounding culture or subculture. On the other hand, if one is considering certain kinds of cells as entities of interest, then the appropriate embedding system for consideration might be certain organs, or certains organisms, or perhaps an embedding system that is an artificially concocted "culture."

But we can view the matter from the opposite perspective. If an organism is selected as the embedding system, then the element level entities that might be studied include cells, organs, chemical constituents, personality traits, or aggressive behaviors. In any case, *the embedding system always is some system that is at a higher level of organization than the entities of concern.* Within the embedding system, the entities can be regarded as parts, and the patterns of relations among the states and actions of those entities—which are the focus of study—can be regarded as internal processes of that embedding system.

It is somewhat useless to try to enumerate the materials of the substantive domain because there are so many possible sets of entities and events that might be of interest from one perspective or another—many possible sciences, so to speak. And, as noted before, when one begins to delimit the substantive domain one begins to intrude on material that is properly in either the conceptual domain or the methodological domain. We hope that the meaning and limits of the substantive domain will become clearer as the other domains are discussed.

The Conceptual Domain

The elements of the conceptual domain, the concepts of interest, are attributes or *properties* that refer to states and actions of

TABLE 2.3
Levels of the Conceptual Domain

Element Level
Properties Representing States and Actions of Entities

Examples of properties:
Cognitive, affective, conative processes

Relation Level
Logical, Causal, Chronological Relations

Examples of relational terms:
Causal relations: necessary, sufficient, suppressing, facilitating
Quantitative relations: <, = ,>
Functional forms: linear, monotonic, nonmonotonic, step function
Logical forms: inclusion, union, intersection, exclusion, AND, OR, XOR, implication,
Logical relations: asymmetry, recursivity, reflexivity

Embedding System Level
Conceptual Paradigm

Potential paradigmatic assumptions
Homeostasis (equilibrium)
Entropy
Teleological orientation

Examples of conceptual paradigms:

Behaviorism	Functionalism	Psychoanalytic
Information processing	Phenomenology	Gestalt theory

entities. The relations in the conceptual domain are patterns of *logical, causal, or chronological association between those properties* (states and actions of entities). The embedding system in the conceptual domain is the set of conceptual assumptions—the "conceptual paradigm," so to speak—that underlies the concepts and relations being considered (see Table 2.3).

For the behavioral and social sciences, concepts of interest have to do with properties of actors behaving toward objects in context, and of nonhuman events and conditions occurring in those same contexts. The battery of concepts likely to be involved in the conceptual domain in any given study depends, in part, on

the particular science; hence, on the sets of phenomena to be studied. It also depends on the particular theoretical perspectives of the scholars involved; hence, on the conceptual assumptions within which the concepts are embedded. A list of concepts for the behavioral and social sciences would be a litany of the "areas" currently studied in those fields: attitudes, motives, interaction, cognitive processes, learning, reinforcement, altruism, aggression, influence, attraction, status, conflict, coordination, exchange, and so forth.

At the relations level in the conceptual domain, we are dealing with ways in which two or more concepts can be related to one another. There are several sets of logical, chronological, or causal relations between elements that can be used to illustrate the contents of this level of the conceptual domain. There is a set of potential logical relations between properties, including the ideas of necessary and sufficient causes, and the ideas of facilitating and suppressing relations. There is a set of quantitative relations, of a discrete type (<, =, >, and the like). There is a set of terms referring to the functional form of a relation (e.g., linear, monotonic nonlinear, nonmonotonic, step function). There is a set of relational terms drawn from Boolean logic and other formal logics (e.g., Inclusion, Union, Intersection, Mutual Exclusion, Logical And, Or, Exclusive Or, Negation, Implication). There is a set of terms describing relations—symmetrical/asymmetrical, recursive/nonrecursive, reflexive/nonreflexive, and the like. These are potential sets of relations for the conceptual domain.

The embedding system level, in the conceptual domain, consists of a number of properties or principles or assumptions that together form a "conceptual paradigm." Features of those conceptual paradigms might reflect such ideas as homeostasis (or equilibrium), a teleological orientation, and assumptions about the nature of causal relations. Overall conceptual paradigms might include such "systems" as behaviorism, phenomonology, gestalt theory, information processing, and functionalism. Developmental work in stage one for the conceptual domain consists in part of reaching out for new or different conceptual systems as well as for new elements and relations.

The Methodological Domain

In the methodological domain, elements (methods) concern **modes for treatment** (i.e., techniques for measuring, manipulating, or controlling) of properties of phenomena (states or actions of the entities of interest). The relations level has to do with techniques for making comparisons between such treatments (i.e., between measures and manipulations of properties of the states and actions of entities). The embedding system in the methodological domain refers to **strategies** or settings within which research is carried out, together with the set of assumptions that goes along with each strategy (see Table 2.4).

In the methodological domain, the elements (methods) include procedures for measuring some property of some state or action of some entity. Runkel and McGrath (1972) call these the "Y treatment" of a variable. Methods, or modes for treatment of variables, also include procedures for imposing a specific value of a property on some case. These include both the manipulation of a variable (the X treatment in Runkel and McGrath's terminology) and the experimental control of a variable (the K treatment, in Runkel and McGrath's terminology).

The relations level in the methodological domain refers to techniques for making comparisons between the outcome of two or more treatments. Comparisons involve the following:

(a) A set of one or more elements that have been given "treatment Y" and that are considered "dependent variable(s)" or measures of the phenomena of interest.
(b) A set of one or more elements that have been given either "treatment X" or "treatment Y" and that are considered as "independent variables" or potential antecedent conditions for those phenomena of interest.
(c) A third set of elements that have received any of various control treatments (held constant, treatment K; matched, treatment M; randomized, treatment R; or ignored, treatment Z) and that are to be considered "other properties" that provide the context and limiting conditions for the comparison.

TABLE 2.4
Levels of the Methodological Domain

Element Level: Modes for Treatment

Treatment Y: measures
Treatment X: manipulations
Treatment K: hold constant
Treatment M: matching
Treatment R: randomization
Treatment Z: ignoring

Relation Level: Comparison Techniques

Set a: dependent variables
Is there one or more measures (Y treatment)?
For each: Is it measured once or more than once?

Set b: independent variables
Are any variables manipulated (X treatment)?
For each X treatment:
Does it have two levels or more than two?
Is it presented once or more than once?
Are the X treatments crossed, nested, or confounded?
Are any independent variables measured (Y treatment)?
For each Y treatment:
Is it measured once or more than once?

Set c: control variables
Control (K)
Randomize (R)
Match (M)
Ignore (Z)

Does assignment of cases to conditions involve a procedure that is:
Random; nonrandom but known; nonrandom and unknown

Embedding System Level: Research Strategies

Strategic considerations:
Realism of context
Precision of measurement and control of behavior
Generalizability over population

Examples of research strategies
Field studies
Experimental simulations
Judgment studies
Formal theories
Field experiments
Laboratory experiments
Sample surveys
Computer simulations

Such comparisons either assess the strength, stability, and shape of association or *covariation* between the variables in sets a and b, or they assess *differences* on the dependent variables (set a) associated with deliberately induced differences on the independent variables (set b).

Such comparisons, either differences or associations, may vary on a number of features such as:

- whether the planned comparison involves one or more than one dependent variable;
- whether the dependent variable(s) are measured once or more than once;
- whether there is one or more than one independent variable;
- whether each independent variable is measured or manipulated;
- whether each of the independent variables is measured once or more than once or, for manipulations, is presented one or more than once; and
- whether the third set of treatments (the control procedures) involves a random procedure for assigning cases to conditions or a nonrandom but known procedure or a nonrandom and unknown procedure.

These and other distinctions are features of relations level comparison techniques in the methodological domain.

In the methodological domain, the embedding system level refers to overall *research strategies* within which methods (i.e., modes of treatment) and comparison techniques are embedded. Runkel and McGrath (1972) have identified eight such strategies and argued that they are related to one another in a circumplex pattern. The eight strategies located in that circumplex are: field studies, field experiments, experimental simulations, laboratory experiments, judgment studies, sample surveys, formal theories, and computer simulations. That strategy circumplex, the underlying rationale for it, and the set of research strategies located in it, provide one possible schema for considering the embedding system level of the methodological domain.

CRITERIA OF VALUE IN THE THREE DOMAINS

We have said, both in chapter one and in previous publications on the VNS (Brinberg and McGrath, 1982; McGrath and Brinberg, 1983; McGrath and Brinberg, 1984), that validity in stage one has the meaning of *value*. This use of the term value can lead to confusion because the term has more than one meaning in social and behavioral sciences. To avoid the confusion stemming from our use of the term value in this context, we will distinguish two meanings of the term value. Then we will move on to a discussion of the criteria of value that are used in each of the three domains.

Value as Worth Versus Value as Preferences

When we say that validity in stage one has to do with value, this does not refer to value as in attitudes or preferences of the researcher. Rather, it refers to value as in *worth*. For example, to say that something is valuable or that it has added value or that it has high "production values" is to use that meaning. The second of these two meanings, value preferences, is an idea that certainly has a place in our discussion of the research process. And indeed, the researcher's values, in the sense of attitudes or preferences, do play an important part in influencing the research process. Some of those influences will be described in this and later chapters. But, the underlying meaning of validity in stage one is the other meaning of the term value. One of the definitions of "value" in Webster's Collegiate Dictionary (fifth edition) is "the quality or fact of being excellent, useful or desirable; worth in a thing." It is this sense of the term, value, as worth, that we use to define validity in stage one of the research process.

The idea of value or worth forms a basis for the idea of *criteria* or standards of acceptability for elements and relations within each of the domains. In the methodological domain, the criteria of worth have to do with usefulness of the methods and comparison techniques as tools for gaining and clarifying information within some research strategy. In the conceptual domain, the criteria of worth have to do with meaningfulness of the concepts and relations as the basis for making interpretations within some conceptual paradigm. In the substantive domain, the criteria of worth have to do with importance of the phenomena and their patterns as crucial features in the operation of some substantive system of interest.

The criteria of value in the methodological domain are scientific and practical, such matters as generalizability, precision, and cost. Those of the conceptual domain are intellectual and philosophical, such matters as parsimony, internal consistency, and comprehensiveness. Those of the substantive domain are social, economic, technological, political, and moral, such matters as system effectiveness, unit well-being, capability, and growth. Furthermore, the criteria of value in each of the three domains are locked together in interdependent sets, interrelated as conflicting desiderata. Many of the choices among alternatives within each domain involve trade-offs or dilemmas. Each possible choice involves both gains on some criteria and losses on others.

Criteria of Value in the Methodological Domain

The methodological domain is characterized by a set of mutually conflicting desiderata, all of which need to be, but cannot be, maximized simultaneously. These conflicting desiderata pose a set of dilemmas for the researcher: All choices, designed to increase the yield on one of the desiderata, at the same time reduce the level of one or more other desiderata. So, although we can specify a set of criteria for the methodological domain, all of which are desiderata, or attributes of value or worth, we cannot specify a set of procedures that will yield top values on all of those criteria at the same time. Such conflicting desiderata exist at all levels of the methodological domain, elements, relations, and embedding system. They perhaps can be illustrated most clearly at the embedding system level. (Some of this material has been discussed in other publications; such as McGrath et al., 1982; Runkel and McGrath, 1972).

In any social and behavioral sciences study, we can reference each item of information with regard to actors, behaviors, and contexts. In any study, one always wants to maximize three mutually incompatible desiderata:

(1) **Generalizability** with respect to the populations (of actors, situations, conditions, and so on) to which the information applies (Criterion A).
(2) **Precision** with respect to the measurement and control of the behavior variables that are involved (Criterion B).
(3) **Realism** with respect to the contexts, or concrete behavior systems, to which that information is intended to apply (Criterion C).

You always want to maximize all three of these—generalizability, precision, and realism—but you cannot. Research strategies that provide the opportunity to maximize any one of these at the same time virtually guarantee low levels of both of the other two. Suppose you choose a laboratory experiment, in the interest of maximizing criterion B (precision of measurement and control of behavior variables). You will necessarily have an artificial (hence relatively unrealistic) context (i.e., low on criterion C), and low generalizability with regard to populations outside the artificial laboratory conditions (i.e., low on criterion A). In contrast, suppose you choose a field study, in the interest of maximizing criterion C (realism of context). You will necessarily have relatively little precision with respect to measurement and control of variables because you will want not to be intrusive (hence low on criterion B), and you will have relatively little generalizability to populations beyond that specific setting (hence low on criterion A). Suppose instead, that you choose a sample survey, in the interest of maximizing criterion A (generalizability with respect to population). You will necessarily have made the context irrelevant (hence low on criterion C), and have little control or precision of measurement of behavior variables (that is, low on criterion B).

These are not limitations based on the preferences or predilictions of the researcher, nor are they limitations based on scarcity of resources. These limitations are *inherent in the research process itself*. Hence, they cannot be avoided in any single study, and the limitations that they pose for the researcher need to be overcome by conducting multiple studies, using different strategies on the same problem.

Note that the advantages of each strategy are just potential advantages; they can be lost by other methodological choices. For example, you can lose the advantages arising from the potential unobtrusiveness of a field study if you couple it with obtrusive methods for measuring variables (such as self-reports or observations by a visible observer). But the limitations of various strategies are inherent and therefore inevitable limitations and cannot be conjured away by any strategems. For example, there have been attempts to gain the precision of the laboratory experiment along with the realism of the field study by concocting an artificial laboratory situation but packing it with features from some real-world system in an effort to make it seem realistic to the participants. But what this does, at best, is give that laboratory situation a little less precision (by introducing "noise" in the form of what has been termed "mundane realism"), though not giving it any more realism of context (what we will call "experiential realism"). The participants are still quite aware that the situation is not a part of their own lives, but rather exists because of, and serves the purposes of, the researcher and not themselves.

There are also dilemmas at other levels of the methodological domain. The various modes of treatment—measurement (Y), manipulation (X), experimental control (K), matching (M), randomization (R), or ignoring (Z)—offer the investigator a complex set of choices that also pose a series of dilemmas. One of these dilemmas is the conflict between scope of information on the one hand and precision of information on the other. Modes of treatment that retain broad scope of potential information for the study also bring with them high levels of noise. In contrast, modes

of treatment that yield a clearer potential signal do so by narrowing the study scope. For example, if a study makes heavy use of modes of treatment X and K, along with Y, in the interest of increased precision of the resulting information, it will at the same time reduce the scope of that information with respect to the focal problem it is an attempt to study. In contrast, if a study makes heavy use of modes of treatment Y but not X or K, and of mode Z but not R, it can preserve much of the scope of the information in the focal problem, but does so at the cost of a high level of "noise" in the information thus gained.

There is a related dilemma: standardization versus generalizability. On the one hand, there is a need for standardization of all conditions whose variation is not being studied, to avoid confounding the resulting information with variations in other, unstudied factors. On the other hand, there is a need for variability on unimportant features of the study, in the interest of establishing the robustness or generalizability of results (Lynch, 1982).

For example, suppose you wanted to find out about relations between some aspects of group structure and the quality of group task performance. If you hold task type constant for all groups in all conditions (i.e., standardize task type), your results will have restricted meaning. On the other hand, if you have various groups doing tasks of different types (even if you somehow balance task types between conditions to be compared), the variability from those differences may obscure the variations related to group structure that you are trying to study. Thus, standardization and generalizability are another pair of conflicting desiderata; the need to maximize both of them at once, and the impossibility of doing so, is another dilemma of the research process.

There is another dilemma involving the modes of treatment. Treatment R, or randomization, refers to the allocation of "cases" to conditions within a study. Treatment R is a mixed blessing. It is a necessary condition to handle the potentially biasing effects of "all other variables" not treated by one of the direct treatment modes (Y, K, X, M). But, it does not reduce, and indeed it exacerbates, the potential effects of those other variables on random variability among cases within the same treatment conditions. Thus, treatment mode R can aid generalizability, but at the expense of standardization; whereas treatment mode K can aid standardization, but at the expense of generalizability. Treatment mode Z (ignoring "all other variables") is antithetical to both of those desiderata because it contributes negatively to standardization while not contributing at all to generalizability.

The methodological domain teems with dilemmas that involve the need to maximize, simultaneously, two or more conflicting desiderata or criteria. The stage one researcher in the methodological domain needs to make those conflicting choices clear, to seek out and generate new forms or techniques for use as elements or relations within this domain, and to clarify how those new forms fit within the dilemmatic structure of the domain.

Criteria of Value in the Conceptual Domain

The criteria in the conceptual domain also involve interrelated sets of conflicting desiderata, not all of which can be maximized at the same time. Hence, choices within the conceptual domain also involve trade-offs, dilemmas, or compromises.

One prominent criterion in the conceptual domain is the principle of **parsimony**. That principle states that, all other things being equal, interpretations that use fewer concepts and fewer (and less complex) relations to interpret a given body of evidence are *better* than interpretations that use more, or more complex, concepts and relations to do so. In the case of extreme differences between two formulations, in number and complexity of concepts and relations, there is little cause to question the value of parsimony of concepts and simplicity of relations. Fewer is better than more at this limit. Much of what is considered progress in science comes about by the development of unifying concepts or theories that interpret a body of evidence more parsimoniously than did prior conceptions. But at the other limit, when there are only small differences in number and complexity of concepts and relations between two competing interpretations, it is hard to find compelling grounds on which to support parsimony as a general criterion of conceptual value.

The principle of parsimony is usually treated as a quantitative concept because it is applied by examining the relative efficiency with which two formulations account for variance in a focal problem. There are a variety of techniques for developing empirical evidence about whether alternative conceptual formulations with differing numbers of concepts and relations can account for appreciably different amounts in the phenomena being studied. But competing conceptual formulations seldom are related to each other in straightforward ways. Rather, competing conceptual formulations typically differ in the basic distinctions with which they start, and in the ways in which those distinctions are interrelated. They differ qualitatively. They may *account for different portions* of the variance, or they may *account differently for whatever portions* of the variance they do interpret. The principle of parsimony, used as a quantitative concept, seldom offers a basis for choosing between competing conceptual formulations.

A second desideratum in the conceptual domain is **scope.** Other things being equal, it is better for a conceptual formulation to cover a broader, rather than a narrower, focal problem. A third desideratum is **differentiation** of detail. Other things being equal, it is better for a conceptual system to differentiate features of the focal problem in detail, rather than to treat them in more general or abstract form.

But the three desiderata of parsimony, scope, and differentiation are mutually conflicting. They are all desirable, but they cannot all be maximized at the same time. To increase parsimony—fewer and simpler concepts and relations—is, necessarily, to give up either scope or differentiation or both. To increase scope means, necessarily, either to decrease detail or increase the number and complexity of concepts and relations (i.e., decrease parsimony) or both. To increase differentiation means, necessarily, to reduce scope or increase the number and complexity of concepts (i.e., reduce parsimony) or both. As in the methodological domain, these three mutually conflicting desiderata cannot all be maximized simultaneously, and they pose dilemmatic choices for the stage one researcher in the conceptual domain.

As with the research strategies of the methodological domain, there are some compromise choices. These tend to "optimize" on

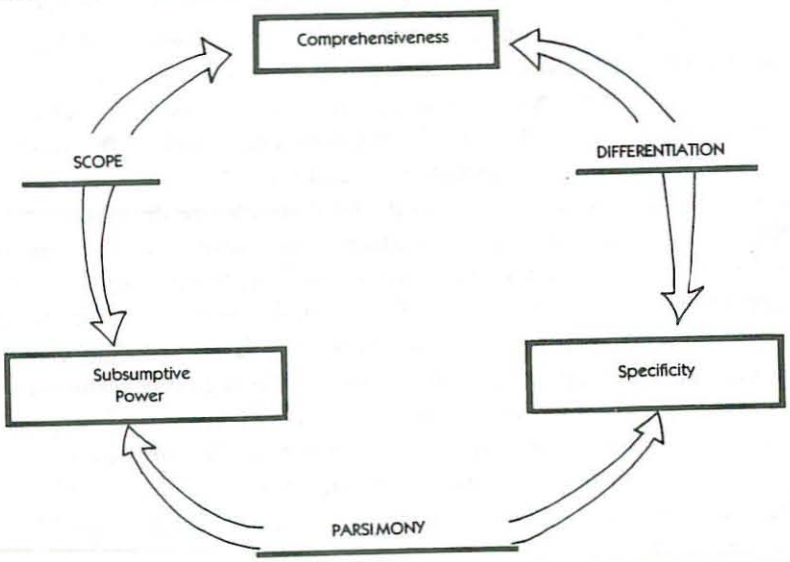


Figure 2.1: Conceptual Criteria as Conflicting Desiderata

two of the three desiderata, but minimize the third. One of these approaches is to seek a high level of **subsumptive power** by using only very abstract terms that keep broad scope and parsimony in number and complexity of concepts. But this minimizes differentiation and detail. Another approach is to seek high **specificity**, by using few but richly detailed concepts to get high differentiation and keep high parsimony. But this can only be done by drastically cutting scope. A third approach is to seek high **comprehensiveness**, by keeping detailed differentiation of concepts but retaining broad scope. This approach, by definition, gives a decrease in parsimony.

Results of these three compromise approaches to solving this dilemma—subsumptive power, specificity, and comprehensiveness—are themselves three mutually conflicting desiderata that cannot all be maximized at once. Such sets of dilemmatic choices face the stage one researcher in the conceptual domain. These desiderata and compromise choices are shown in Figure 2.1.

Another set of criteria in the conceptual domain has to do with the **logical coherence** of the concepts and relations within a given

conceptual formulation. Lack of logical coherence can arise in either of two ways. Some of the concepts and relations within the formulation can be in conflict with one another—that is, the formulation can contain *mutually contradictory* propositions. In addition, the set of propositions—of concepts and relations—that make up a given conceptual formulation can contain one or more *logical gaps,* so that the overall system fails to make specifications for major portions of the substantive phenomena involved. These two breaches of logical coherence are related in a complementary, rather than conflicting, way. A conceptual formulation can avoid both logical gaps and contradictory propositions.

One final pair of desiderata for conceptual formulations, also related to each other in a partially conflicting form, is worth noting here. Conceptual formulations need to deal with *important* aspects of the substantive phenomena and their relations; and they also need for that conceptual formulation to deal with *testable* aspects of those substantive phenomena and relations. The former, importance, is a criterion of conceptual value that ties that domain to the substantive domain—because importance is really a judgment about the contents of the substantive domain. So, in a sense, it is a criterion to be applied not to stage one work in the conceptual domain, but to the stage two work in the structure that we call a *set of hypotheses.* The criterion of testability is a criterion of conceptual value that ties that domain to the methodological domain—because testability is really a judgment about the availability of methods suitable for making such a test. So, in a sense, testability is a criterion to be applied not to stage one work in the conceptual domain but to stage two work in the structure that we call a *study design.* It is nevertheless the case that concepts that are "important" tend to be hard if not impossible to test, and that concepts that are "testable" often tend toward the trivial.

Criteria of Value in the Substantive Domain

It is more difficult to discuss criteria in the substantive domain than in the other two domains, just as it was harder to discuss the contents of that domain. The substantive domain is, to some extent, ineffable. It *is*. But when you talk about it, it tends to slip away, either into methodological or into conceptual matters. When we talk about the substantive domain, we tend either to talk about aspects of it as *conceptualized*, or to talk about aspects of it as *observed or experienced*.

The substantive domain is also difficult to discuss because it is so vast and diverse. Any formulation of elements and relations tends to prejudge which specific portions of the "real world" are to be given emphasis. We want to emphasize those portions of the substantive domain that have to do with the states and actions of entities that are human systems. But it is important to have our treatment include all sorts of states and actions (even including some that have not yet been focused on in research, but could be!). It is also important that our treatment include various system levels: Within individual processes; individuals as systems; interpersonal processes; groups, organizations, communities and other multiperson systems of various kinds; as well as broader social processes, institutions and macro-level social entities (nations, cultures, generations, classes, political systems, economic systems, and the like).

Whereas the conceptual domain involves abstract ideas and its criteria are intellectual; and the methodological domain involves operations and procedures and its criteria are scientific; the substantive domain involves *ongoing, real-world systems*, and its criteria involve behaviors and results that have social, economic, technological, political, and moral content. These criteria of "value" ask whether something is profitable, technologically effective, morally proper, politically permissible, or humane.

Specifically, we can identify three general criteria that seem to be applicable to a wide range of substantive systems, of various kinds and at various system levels. One criterion for any substantive system is **system well-being**—viewed in the broadest sense. All conditions and behaviors that would threaten the health and safety, or constrain the positive development of the system must be viewed as adverse, and those conditions and behaviors that would advance the health, safety, and positive development of the system are viewed as desirable, on that criterion.

A second criterion for any substantive system is **system task performance effectiveness**. All conditions and behaviors that hinder the system in carrying out its tasks (i.e., actions instrumental to its goals) must be regarded as adverse, and those conditions and behaviors that facilitate such task performance must be viewed as desirable, on that criterion.

A third criterion for any substantive system is **system cost** in expenditure of energy and other resources. All conditions and behaviors that *increase* the unit's costs (in energy, money, raw materials, etc.) in its pursuit of task performance effectiveness and system well-being, must be viewed as adverse, and those conditions and behaviors that decrease such costs must be considered desirable, on this criterion.

As in the other domains, these criteria are a set of partially conflicting desiderata, not all of which can be maximized at the same time. They pose a set of dilemmas for anyone making decisions about the system's operation. In the substantive domain, the criteria are partially conflicting at each system level, and there are some conflicts *between* system levels as well.

We are assuming, here, that the portions of the substantive domain that are of importance to the social and behavioral sciences are *social units that constitute a set of partially nested open systems at different system levels* (individuals, groups, organizations, communities, cells, organs, social institutions, neural processes, and so on). Furthermore, we are assuming that, for any given focal system at any given level, that system is *in dynamic interaction with its environment(s)* (that is, with embedding systems at higher levels of organization), and that those environments themselves are changing. Furthermore, we are assuming that any given focal system is composed of *sets of subsystems and processes that are in dynamic interaction with one another and with the focal system itself*. The three criteria—system well-being, system task effectiveness, and system cost—are interdependent and partially conflicting at each system level. And, because there is dynamic interaction between the system, its parts, and its environment, the criteria are interdependent and partially conflicting across system levels as well.

For example, high system task performance effectiveness of a work organization may be purchased at a serious cost in system well-being (and vice versa), just as high task performance effectiveness at the individual level may be purchased at a cost in well-being for the individuals who are the organization's members (and vice versa). Safety, health and environmental risks, or simply task overloads, may be imposed or undertaken in the service of advancing task performance effectiveness. Similarly, increases in expenditures can translate into increases in task performance effectiveness or system well-being or both. And, decreases in expenditures are likely to yield losses in task performance or system well-being or both. Furthermore, these trade-offs can occur between criteria at the same system level (organization or individual, in the example above), or across system levels.

Given such conflicting criteria within and between system levels, when stage one researchers choose a topic for study, they are often choosing for emphasis one of these criteria over the others. Hence, they are "taking sides" in intrasystem disputes that are moral and sociopolitical in flavor. For example, when industrial psychologists identify task performance effectiveness or productivity as an important substantive phenomenon for study in their field, they are, in effect, taking sides in the sense that they are giving preference to one of the conflicting criteria (task performance effectiveness) and to one of the competing levels (individual versus organization). Similarly, when social psychologists set out to study ways to *reduce conflict* in bargaining situations, they are also taking sides—in this case, between a point of view that eschews conflict and a point of view that sees conflict as part of the creative or innovative process. And, depending on what conflict reduction strategies they give their attention to, they are probably taking sides on other kinds of issues as well. Another example of such side-taking would be if one crime prevention program concentrated on what the potential victims ought to do to avoid being victims, whereas another concentrated on what might be done to, or for, or about, the potential criminals to reduce the probability that they will commit the crimes. Still another example of side-taking would be if a program to improve education focused on improving parent-teacher communication, curriculum resources, preschool opportunities, teacher training, or child nutrition.

In all of these examples, the specification of a phenomenon that is to be a potential topic of concern in regard to the system—and, perhaps, a topic of research on the system—is at the same

time the *specification of a desirable state for the system*. It often involves emphasizing some of the set of conflicting desiderata (and some system levels) and ignoring others. It is here, in the substantive domain, that the two meanings of the term *value* get most entwined. For what gets focused upon as "of worth" in these matters—task productivity, conflict reduction, modification of victim behavior, teacher-parent communication—are those features of the systems that are in accord with the investigators' values, attitudes, and preferences.

Researchers' Values and the Research Process

In the context of this discussion on criteria of worth, it is easy to see why there has been such confusion and contention over the questions of when, and how, and if, the researchers' values enter into the process of scientific research. Much of the discussion on one side of the argument urges that research needs to be *value free*. That view holds that research has to do with means not ends, and that its execution should be determined without regard to the worthiness of the ends it will serve. Some of the discussion on the other side of the question argues that research *cannot* be value free, and therefore that it ought to be done to further worthy ends. A third line of argument would hold that although research ought to be value free, human investigators cannot operate in a value-free manner. Therefore, not only do individual researchers need to minimize the impact of their own values on their research, but also the research enterprise, collectively, needs to take investigators' values and preferences into account by somehow counterbalancing them within a substantive area.

The VNS suggests a somewhat different point of view. In the first place, the values (that is, the preferences) of the stage one worker necessarily enter into the choices made as to which elements and relations will be generated, developed, clarified, and made available for stage two work. This intrusion of values into choices is true for each domain, but it is of special force in the substantive domain.

In the methodological domain, the criteria are complex and conflicting, but they are nevertheless often discussed as if they were amenable to explicit evaluation in terms that transcend the attitudes and values of individual researchers. After all, isn't precision a matter of logic, not of taste? Likewise for standardization, variability, and randomness. Similarly, the criteria for the conceptual domain also are often viewed as if they transcended the values of individual researchers.

But are such choices of methods and concepts free of the researchers' values (preferences)? We believe that freedom is more apparent than real. Although choices of elements and relations in the domains seem to be based on "established principles"—derived from the assumptions of the embedding system involved—those established principles actually reflect the social consensus within the scientific and intellectual communities involved. Furthermore, those values/preferences, in regard to scientific and philosophical matters, are already inextricably bound up with the criteria of "value as worth" that are already built into the consensus of researchers in those areas. Thus, we disagree with the generally held view that in method and concepts, at least—if not in choice of substantive problem—social science can be, or become, value free.

In the substantive domain, choices about what is "of worth" are obviously closely tied to the choosers' own personal preferences or, in terms used earlier, the choosers' social, economic, technological, political, and moral values. Certainly, the value entanglement of research choices by stage two researchers is widely recognized (though sometimes deplored, inappropriately we think). But an even more fundamental influence of the values/preferences of the researcher involves the very specification of phenomena, patterns of phenomena, and embedding substantive systems, *that could be researched, that are worthy of attention.* These are the choices made by stage one researchers working in the substantive domain.

Such influence is inevitable—humans being human. To deplore it is pointless because it is inevitable and not to be avoided. Whether or not we wish for scientific research to be value-free, it cannot be.

Note, too, that the stage one workers in the substantive domain are *substantive system experts*, probably not researchers in the usual meanings of that term. They are likely to be attorneys, colonels, mayors, physicians, mental health specialists, school superintendents, coaches, bankers, and the like. When they choose elements, relations, and embedding substantive systems for special attention, those substantive system experts are not making what they see as research choices. Rather, they are making practical choices related to the ongoing operation of systems in which they are likely themselves to be closely involved. So when a manager is formulating concepts vis a vis the system he or she is managing, we can expect that manager to "take sides," and when a union representative is formulating them, we can expect that agent to "take sides," but a different side. When stage two researchers select elements, relations, and embedding systems of interest within the substantive domain, they are very likely to find, and choose for focus, some phenomena, patterns, and systems whose identification, development, and clarification was propelled by *someone else's values*. Such value-driven selectivity is also the case for stage two researchers with respect to the other two domains, for they often select methods propelled by someone else's scientific values and concepts propelled by someone else's philosophical values. The substantive system is no different than the other two in terms of its being value based. It only seems different because we have strong consensual norms that rationalize these choices in the methodological and conceptual domains, whereas value differences in the substantive domain are often more visible and apparently more intense.

Of course, stage two researchers sometimes do stage one work. But the stage one work of those researchers whose main interest is in the development of sets of empirical findings (that is, stage two work), or in verification and extension of such findings (that is, stage three work), is apt to be limited to efforts in but one of the domains—the one of most importance to that researcher. In Chapter 3, we will discuss in some detail the tendency of researchers to begin with a strong preference for one of the three domains and the important consequences that ensue from that tendency.

What can we now say about the intrusion of values—and whose values—into the research process? First, the stage one researcher cannot attain, nor approach, a value-free effort in the methodological or the conceptual domains. Nor can the stage two researcher approach or attain such a value-free situation vis a vis what he or she "borrows" from the methodological and conceptual domains. Almost always, tools from the methodological and conceptual domains contain value-laden assumptions that are not obvious to the stage two researcher who selects them for use.

These hidden assumptions are a kind of value trap for the stage two researcher—and this refers both to *value as worth* and to *value as preference.* That is, they may lead the stage two researcher, unknowingly, to emphasize methodological or conceptual criteria that are really not in keeping with that stage two researcher's current purposes. For example, the underlying assumptions about ratio scales and about equal intervals for some rating scale, or the classical test theory assumptions underlying various scoring procedures may not at all suit the perspective of the investigator who unwittingly incorporates them into his or her stage two work.

But the stage one researcher in the substantive domain is in an even less favorable position to avoid making choices into which his or her values will intrude. Nor can the stage two applied researcher, who does some of his or her own stage one work in the substantive domain, avoid such choices. And the stage two researcher who does not do any stage one developmental work, but rather selects from what is available in the substantive domain, cannot avoid such intrusion by the value choices of someone else.

In the light of these considerations, therefore, *researchers need to take very seriously the matter of what values are associated with the substantive phenomena they study, and with the methods and concepts they use.* They should not try to deny or wish away such influences by inattention to them.

Furthermore, stage two researchers are human beings and citizens. Like everyone else, they are responsible for the consequences of their actions. Therefore, we think researchers should very carefully and deliberately examine the phenomena, patterns, and substantive systems they select for study in the light of their own social, economic, technological, political, and moral values. If they do not, they may realize later (when it is too late) that they

have been contributing to values (criteria) that they do not wish to enhance, or have been doing work that will tend to reduce values (criteria) that they would like to maximize. Whatever their views might be on various issues, within the substantive systems they choose for study, researchers should make sure that their own research activities enhance, or are in accord with, or at least do not violate, their own values and preferences on those issues.

GLOSSARY OF TERMS

Criteria in the Substantive Domain: System well-being, system task performance effectiveness, and system cost. A researcher who focuses his or her efforts in this domain—a **substantive system expert**—will evaluate the particular substantive system under study using these criteria. For example, a researcher examining work groups may be interested in the satisfaction of the members of the work group (system well-being), the performance of the work group (task performance effectiveness), and the amount of time and resources needed to perform a task or maintain system well-being (system cost).

Criteria in the Conceptual Domain: Parsimony, scope, and differentiation. A researcher who focuses his or her efforts in this domain—a **conceptual system expert**—will evaluate the elements and relations selected from the conceptual domain using these criteria. For example, the concepts in Freudian theory can be evaluated in terms of the complexity and number of concepts in the theory (parsimony), the range of phenomena for which the theory can account (scope), and the extent to which the theory describes the phenomena under study in detail (differentiation).

Criteria in the Methodological Domain: Generalizability, realism, and precision. A researcher who focuses his or her efforts in this domain—a **methodological system expert**—will evaluate the elements and relations selected from the methodological domain using these criteria. For example, a laboratory experiment or a field study can be evaluated in terms of the extent to which each of them allows the researcher to make inferences about a population of interest (generalizability), the naturalness of the context or setting (realism), and the amount of precision of measurement and control over extraneous factors that the study permits (precision).

Chapter 3

STAGE TWO Study Paths and Research Orientations

The Validity Network Schema discussed in this book, and especially the alternative paths for conducting stage two research, provides a way to examine the research process in some detail. That will be the goal of the first part of this chapter. In the second part of the chapter, we will present several extended examples of sets of research studies that exemplify the various research paths. The VNS and the study paths also provide a way to examine systematically the familiar distinction between applied and basic research. That distinction will be the topic of the third part of the chapter. Then, in Chapter 4, we will discuss the many validity issues that arise in stage two of the research process.

ALTERNATIVE PATHS FOR CONDUCTING STAGE TWO RESEARCH

Any stage two research effort involves elements and relations from all three of the domains, the conceptual, methodological, and substantive domains. We can regard stage two of the research process as proceeding in three steps. First, the researcher selects elements and relations from one of the domains. We argue that the first domain is the one in which the researcher has some preference or special interest. Second, the researcher brings those elements and relations together with elements and relations from a second domain to form an intermediate or instrumental structure. Third, the researcher brings elements and relations from the third domain into that structure. Because there are three domains, there are three places to start, and there are three instrumental structures that can be built in the first two steps. The three starting places represent three research orientations to be discussed later in this chapter. The three instrumental structures represent three distinct styles or paths for the conduct of stage two research. These three paths pose different opportunities and hazards for the investigator.

An Overview of the Paths

The three alternative paths of stage two constitute distinct research styles. The first path involves combining elements and relations from the conceptual domain and the methodological domain to form a **study design**, and then **implementing** that design by applying it to some elements and relations from the substantive domain. We call that the **experimental path**. The second path involves combining elements and relations from the conceptual domain and the substantive domain to form a **set of hypotheses,** and then **testing** that set of hypotheses by application of some elements and relations from the methodological domain. We call that the **theoretical path**. The third path involves combining elements and relations from the substantive domain and the methodological domain to form a **set of observations**, and then **interpreting** that set of observations by application of some elements and relations from the conceptual domain. We call that the **empirical path**.

Completion of the steps of stage two, by any of the three paths, results in a **set of empirical findings**. In certain ways, all such sets of findings are alike. They all provide empirical evidence on a focal problem. But in many other ways, the resulting sets of empirical findings differ depending on which of the paths was followed in attaining them. For a given focal problem, a set of empirical findings that result from implementing a research design is not the same kind of information as a set of findings that result from testing a set of hypotheses. Neither of those sets of findings are quite the same kind of information as an interpretation of a set of observations.

In the rest of the chapter, we will present a number of examples of research programs or portions of them to illustrate research done by different paths and pathways. We are necessarily making presumptions about what was in the minds of the researchers. In many cases, one can see the overall thrust of the program, and therefore can make some reasonable presumptions about what motivations must have guided the researchers' choices and steps. The written account of the researchers' purposes, plans, and procedures that becomes available in the form of published reports of the research always has been prepared post hoc and usually has been subjected to many modifications that obscure past "facts." To actually know the sequence of steps taken in a given case, it would be necessary (but probably not sufficient) to be present during the events. To actually know the research orientation (that is, what was of most importance to the researcher), it would be necessary (though not sufficient) to be that researcher.

We are interested here not so much in the literal history of research steps as in their logical impact. So we would be more accurate if instead of saying, "So and so followed the experimental path, beginning with a set of methods and . . . ," we said, "So and so's research sounds as if it were done following the experimental path, beginning with a primary interest in certain methods and . . . ." Finally, here and throughout the book, when we choose research to illustrate our points we do not mean to imply that we think it either exemplary work, or an instance of poor research. We choose particular research as illustrations simply because we are familiar with that work and think it exemplifies the points we want to make.

The Experimental Path

The *experimental path* brings together materials from the conceptual and methodological domains in the first two steps of stage two, without bringing in materials from the substantive domain. Such activity involves combining, on the one hand, properties (of phenomena) and conceptual relations (among those properties) that have been drawn from within a *conceptual paradigm*; and, on the other hand, modes of treatment (techniques for measuring, manipulating, and controlling variables) and comparison techniques (techniques for comparing or assessing the relation between variables) that have been drawn from within a methodological paradigm or *research strategy*. The result of such a combination is an instrumental structure that we call a **study design**. When concepts or properties are combined with modes of treatment, the result is a set of (planned) **operations**. When conceptual relations are combined with comparison techniques, the result is a set of **planned comparisons**. Whereas a conceptual paradigm is the embedding system in the conceptual domain, and a research strategy is the embedding system in the methodological domain, the embedding system for the combination of operations and planned comparisons is the instrumental structure we call a study design, which is a plan for a study.

If one begins stage two research along this experimental path, then the third step involves bringing material from the substantive domain into combination with this instrumental structure, the study design. Here, substantive phenomena and patterns of those phenomena that have been drawn from within some

substantive system are brought into combination with the operations and planned comparisons that are part of the study design. When substantive phenomena are combined with operations, the result is a set of **variables**. When patterns of substantive phenomena are combined with planned comparisons, the result is a **set of findings**. The embedding system for the variables and the set of findings is a **body of evidence**.

For example, the development of Fishbein and Ajzen's (1975) theory of reasoned action seems to have followed the experimental path. They began with certain concepts (attitudes, behavioral intentions, beliefs, and the like) and some ideas about relations among them (e.g., that the attitude toward an act represented the sum of the products of each salient belief multiplied by its evaluation). They then selected certain methods of measurement for those concepts (e.g., semantic differential forms of self-report) and certain comparison techniques (e.g., multiple regression analysis). When they had combined those concepts and relations with those methods and comparison techniques, they had a study design. That study design consisted of a set of planned operations and a set of planned comparisons among those variables. They *then* chose some substantive system whose phenomena and patterns they could use as a vehicle for implementing that design. In their developmental work, just what specific set of substantive phenomena would be studied was of less importance than the development and clarification of certain concepts and conceptual relations. Later work using the Fishbein-Ajzen theory, of course, has been strongly concerned with the substantive systems that were being studied (e.g., studies of family planning, voting, smoking behavior, and consumer behavior of various sorts).

The Theoretical Path

The *theoretical path* brings together materials from the conceptual and substantive domains in the first two steps of stage two, without bringing in materials from the methodological domain. Such activity involves combining, on the one hand, properties and relations from within a conceptual paradigm; and, on the other hand, phenomena (states and actions of entities) and patterns of relations among those phenomena, that have been drawn from within an embedding substantive system. The result of such a combination is an instrumental structure that we call a **set of hypotheses** (and that we might have called a theory except that many and varied meanings have become attached to that word in our fields). When concepts or properties are combined with phenomena, the resulting elements are **constructs**. When conceptual relations are combined with substantive patterns the resulting relations are **hypotheses**. Whereas the conceptual paradigm is the embedding system in the conceptual domain, and a particular substantive system is the embedding system in the substantive domain, the embedding system for constructs and hypotheses is the instrumental structure we call the **set of hypotheses** (or theory).

If one begins stage two research along this theoretical path, then the third step involves bringing material from the methodological domain into combination with this instrumental structure, the set of hypotheses. Here, modes of treatment (methods for gathering information) and comparison techniques for using that information are brought into combination with the constructs and hypotheses that are part of the set of hypotheses or theory. When modes of treatment are combined with constructs, the results—as in the experimental path—are measures of variables. When comparison techniques are combined with hypotheses, the results—as in the experimental path—are a set of findings. And, as in the experimental path, the embedding system for variables and findings is a body of evidence.

For example, a set of studies by Lewin and colleagues on changing housewives' food habits during World War II (Lewin, 1953) seems to have followed the theoretical path. Lewin began with a set of substantive problems (how to get housewives to use certain cheap, plentiful, but unpopular cuts of meat) and with certain conceptual tools (valence, forces for and against change, and so forth) that he thought applicable to those problems. He built a set of hypotheses and then developed some experimental methods to manipulate and measure his key concepts so as to test those hypotheses. It is ironic that, although Lewin probably cared more about both the concepts and the substantive systems that he studied than about the methods he used, it is the invention of the latter that had the greatest lasting impact on the field. Various parts of his work were seminal contributions to the development of "experimental social psychology."

Another example of research that seemed to have followed the theoretical path is Newcomb's (1961) work on the acquaintance process. He began with a set of concepts (communicative acts, system balance, mutuality of interpersonal attraction, and the like) and with a particular substantive system (an experimenter-controlled student residence to which the research team had extensive access for an entire semester). He and his colleagues then developed methods for assessing the key constructs repeatedly during the semester, and for manipulating some of them to carry out field experiments, to test the set of hypotheses. As with Lewin, Newcomb undoubtedly had less stake in the methods of measurement used in that study than in the substantive system and the conceptual model being studied.

The Empirical Path

The *empirical path* brings together materials from the methodological and substantive domains in the first two steps of stage two without bringing in materials from the conceptual domain. Such activity involves combining, on the one hand, modes of treatment and comparison techniques drawn from within a research strategy; and, on the other hand, phenomena and patterns from within an embedding substantive system. The result of such a combination is an instrumental structure that we call a **set of observations** (and that we might have called a data set except that the term data seems to have acquired many meanings in our fields). When methods for gathering observations are combined with substantive phenomena the result is a **set of indices**. When methods for analyzing data are combined with patterns of substantive phenomena, the result is a **set of arrays**. The embedding system for those indices and arrays is the instrumental structure, the set of observations.

If one begins stage two research along this empirical path, then the third step involves bringing material from the conceptual domain into combination with this instrumental structure, the set of observations. Here, some concepts or properties of phenomena and some relations among concepts from within a conceptual paradigm are brought into combination with the indices and arrays that constitute the set of observations. When concepts or properties are combined with indices, the result—as for the other two paths—is a set of variables. When conceptual relations are combined with a set of arrays, the result—as for the other two paths—is a set of findings. And, as for the other paths, the embedding system for variables and findings is a body of evidence.

For example, there have been a number of research studies done using data derived from the Consumer Expenditure Survey conducted by the Bureau of Labor Statistics (e.g., Derrick and Lehfield, 1980). Those data have been available and several sets of researchers have conducted extensive analyses on various portions. When any one set of investigators has organized and analyzed some subset of these data, it yields what we term a set of observations. These researchers then attempt to apply sets of concepts and relations (e.g., socioeconomic status, gender, education, and race differences) to transform that set of observations and thereby make it into a "body of evidence."

Another example of the empirical path is work done using the Human Relations Area Files. That set of files is a compilation of ethnographies on many cultures, gathered so that each ethnography contains much information on common topics but also much information that is in some but not all of the files. For many years, social scientists have selected samples of cultures and ethnographies from those files, organized that information into a systematic set of observations (with indices and arrays involving some qualitative and some quantitative information), and then searched for concepts and conceptual relations that would interpret the patterns of observations found in those analyses.

Two Pathways for Each Alternative Path

The three paths of stage two really contain six alternative routes or **pathways**. There are six pathways because the researcher can start step 1 in any of three domains and can take either of two pathways for step 2. In effect, there are two alternative pathways underlying each of the three main paths or styles. These are shown in Table 3.1, and discussed in three extended examples in the next section.

For the experimental path, we have two subsets of designs. One involves **concept-driven designs**. For these, the researcher "starts with" (i.e., has primary interest in) the conceptual system in question and builds a design by drawing upon methods to fit that conceptual system. The other involves **method-driven designs**. For these, the researcher starts with and has primary interest in the methodological system and draws upon the conceptual system to fit the methodological choices.

Similarly, for the theoretical path, there are two ways to build a set of hypotheses. One involves **concept-driven hypotheses**. For these, the researcher begins with primary interest in the conceptual domain, as in the first case noted above, but instead of going to the methodological domain for step two (to build a design), he or she goes to the substantive domain to select substantive systems that will fit with the conceptual system of interest. The other theoretical pathway involves substantive or **system-driven hypotheses**. For these, the researcher starts in the

TABLE 3.1
The VNS System: Orientations and Pathways for Stage Two Research

Orientations	Main Paths	Pathways	1st Choice Domain	2nd Choice Domain	Step 1 Structure	3rd Choice Domain	Step 2 Product
Method	Experimental	MCS	M	C	Method-driven design	S	Implemented design
Basic	Experimental	CMS	C	M	Concept-driven design	S	Implemented design
Basic	Theoretical	CSM	C	S	Concept-driven hypotheses	M	Tested hypotheses
Applied	Theoretical	SCM	S	C	System-driven hypotheses	M	Tested hypotheses
Applied	Empirical	SMC	S	M	System-driven observations	C	Interpreted observations
Method	Empirical	MSC	M	S	Method-driven observations	C	Interpreted observations

NOTE: M = methodological domain; C = conceptual domain; S = substantive domain.

substantive domain with a system of special interest, and goes to the conceptual domain for material to fit the already selected substantive system instead of going to the methodological domain in step two, as is done in a case to be considered next.

Finally, for the empirical path, there are two ways to build a set of observations. One involves substantive driven or **system-driven observations**. For these, the researcher starts with a preferred system in the substantive domain, and goes to the methodological domain for methods to fit the selected substantive system instead of going to the conceptual domain in step 2. The other empirical pathway involves **method-driven observations,** in which the researcher begins with an interest in the methodological domain, and draws upon the substantive domain for material to fit the methodological system instead of going to the conceptual domain in step 2.

The different pathways affect which features of the research process will dominate the "fit" that is the primary validity question of stage two. For example, sometimes the choices made in step 1 within the initial (preferred) domain are choices that carry with them strong prescriptive and proscriptive assumptions that influence later steps of the process. That is, those first-domain choices specify stringent conditions that must be met if material selected from another domain is to be a proper "fit." These assumptions function as constraints on what can be selected from the other domains. For example, any conceptual model that includes a multiplicative combination of variables must be supported by methods of measurement that generate ratio scales.

Other choices from the first domain are permissive rather than proscriptive. That is, after having made a certain choice in step 1, you still have some flexibility regarding what material you can select from a second domain to fit that choice point in the instrumental structure of step 2. For example, some conceptual models require the combination of a weighted set of variables but do not delimit the form of the weights or of the combination.

Such a conceptual model can be fitted with any of a variety of measures and comparison techniques from the methodological domain.

In some cases, the second domain dominates for a certain choice point, rather than the first domain. If the first domain is permissive about a certain choice point, but the second is quite restrictive, the resulting structure will be as restrictive as the second rather than being as permissive as the first. For example, if a researcher with a conceptual model that uses a combination of weighted variables draws upon regression techniques, the resulting design—though concept-driven—will be restricted (in the combination and weights of variables) by the assumptions of that regression measurement model. *At each step, constraints on what choices can still be made are the union of all the restrictions contained in the assumptions of the preceding choices. The remaining possible choices are the intersection of the alternatives left free in each of those preceding choices.*

SOME EXTENDED EXAMPLES OF THE DIFFERENT PATHS AND PATHWAYS

We can illustrate some of these points by considering some extended examples in which all of the pathways are used in relation to the same focal problem. The first illustration shows how one methodological tool—factor analysis—can be used in work involving all six pathways. The second illustrates how work in a particular substantive topic—interaction processes in groups—can be studied by work following all six pathways, and how such use of multiple paths enriches the findings. The third example illustrates how a particular set of concepts—need achievement and related ideas—can be explored with work following all six pathways, and how such use of multiple pathways expands and sharpens applications of those concepts.

Example 1: Factor Analysis

We can illustrate some differences among the pathways by considering various uses of a single specific technique, factor analysis, within the research process. We presume that L. L. Thurstone's (1938) early work on factor analysis was "method-driven," so to speak. We further suppose that he began his stage two use of that methodology along what we have called the experimental path, with a method-driven design that already involved concepts such as simple structure. In one of his earliest applications, he implemented that method-driven design by using quite arbitrary (and indeed trivial) substantive material: He did a factor analysis involving many measurements on boxes of many different sizes and shapes. His choice was quite sensible under the circumstances. Thurstone did not really care about the factor structure of the boxes (we presume), but he used the material to show that his factor analytic technique would isolate, recover, and identify the "true" structure underlying a set of measurements. That set of procedures typifies the method-driven form of the experimental path.

In contrast, much subsequent work in the development of factor analysis—still "method-driven," we presume—followed the method-driven form of the empirical path. It first built a set of observations (e.g., factor analysis of scores from many tests of various kinds administered to a sample of individuals), with conceptual interpretation of the meaning of the factor analysis results left until later. The frequent use of factor analytic techniques in such a "concept-free" manner helped gain those methods a negative reputation among experimentalists.

To carry this example further, we might regard Guilford's work (for example, Guilford, 1954), that made use of factor analysis to explore the human abilities area, as exemplifying the system-driven form of the empirical path. He began (we presume) with primary interest in the substantive area of human abilities. He called upon methodological tools (mental tests on the one hand, and factor analysis on the other) as means to help pursue that interest. Conceptual interpretation of the meaning of resulting factors was thereby left for later consideration.

On the other hand, many researchers, who were basically following the theoretical path—either via a system-driven set of hypotheses or via a concept-driven set of hypotheses—have called upon factor analysis techniques as a way to process data that they collected in the service of testing those hypotheses (e.g., Osgood et al., 1957; Rokeach and Fruchter, 1956). Presumably, many modern uses of confirmatory factor analysis (e.g., Rokeach and Fruchter, 1956) follow this theoretical path in either its system-driven or its concept-driven form.

We can complete the illustration by imagining work beginning as the concept-driven form of the experimental path, involving factor analysis as part of its design plan, and then bringing in material from some substantive system to implement that design. This pathway is one that seems to have been followed by Thurstone's (1928, 1938) own work on the measurement of attitudes. Although that work did not crucially involve the use of factor analysis, it might have, and it illustrates the concept-driven design pathway very well. In that work (e.g., Thurstone, 1928; Thurstone and Chave, 1929), Thurstone seems to have followed the concept-driven form of the experimental path, beginning with a prime interest in the attitude concept, proceeding by developing some new data collection and data analysis tools, and then applying those concepts and tools to the study of "attitudes toward the church"—about which, we presume, he cared less than he cared about the attitude concept and the attitude measurement and analysis techniques involved in those studies.

This extended example shows that any one specific tool or technique or set of materials can play a part in research involving

any of the alternative paths and any of the underlying pathways. Moreover, the illustration suggests that such a specific technique or set of materials plays an entirely different role in research that follows different pathways. In Thurstone's work toward development of factor analysis, the assumptions and requirements of that technique entirely dominated choices throughout the research process. The research was done *for the purpose of development of that method.* In Guilford's (1954) use of factor analysis, it was presumably human abilities—and perhaps the mental tests themselves—that dominated the research process. Factor analysis was, one might say, a mere convenience. The same could be said of the role of factor analysis in the confirmatory studies described in the example as following one or the other of the theoretical pathways. Those researchers were interested in their hypotheses. If factor analysis had not been available, and in a viable form for computer analysis, they presumably would have chosen some other technique for processing and interpreting their evidence.

Thus, a given technique such as factor analysis can be used in research along any of the pathways. Its role in the research differs as a function of the pathway being followed. Therefore, it seems likely that research efforts that follow different pathways, even if they use some of the same techniques and materials, do not yield exactly comparable results. For example, there is no way that Thurstone's use of factor analysis on the boxes could have led him to the conclusion that there was no identifiable structure underlying the data. He selected the boxes *because they had an obvious structure that the measures would reflect.* But "no structure" is always a potential outcome for research that uses factor analysis and follows either the empirical path or the theoretical path. Thurstone's work in the development of attitudes could not lead to the "Bemian" conclusion that attitudes are inferences from behavior, or to the "Skinnerian" conclusion that intervening cognitive concepts such as attitude are excess intellectual baggage. Nor could that work lead to any of the "Coombsian" conclusions about unfolding, nonmonotonicity, and multidimensionality of preference functions, ordered metrics, and nonscale types. All of those possibilities were eliminated by constraints that were built into some of Thurstone's early choices (e.g., that the attitude concept was useful and primary, and that an attitude could be mapped to a single dimension). Yet all of those conclusions by Bem, Skinner, and Coombs are quite reasonable ones to draw if you are working with the same material (attitude concepts, attitude measures, and human preference responses) but along different routes or pathways.

Example 2: Group Interaction Process

Another extended example of the paths and pathways can be drawn from the early work on interaction process in groups. Bales's (1950) hallmark work presumably began as a concept-driven design. That is, Bales apparently was interested primarily in some conceptual matters about group process (e.g., equilibrium processes, instrumental and expressive functions), and he set out to develop methods to make observations pertinent to those concepts—his interaction process analysis (IPA) system. Only then did he concoct groups *of the kinds, and under the conditions, best suited for the application of his concept-driven design.* These conditions included: use of relatively small and ad hoc groups; work on a task assigned by the experimenter that required much verbalization and that could be completed within a manageable experimental session; work under conditions in which the experimental team could apply massive efforts to observe and code behavior of the members; and work under conditions in which the members would tolerate the obtrusiveness of the research activities. Given these research conditions, Bales and associates developed a number of powerful and general "findings" about interaction process in such "full-fledged" problem solving groups. Some of these findings had to do with a

postulated phase sequence for problem solving, an instrumental/expressive equilibrium process, a specialization of functional roles among group members, and the like.

Bion (1961), on the other hand, and especially Thelen and Stock, who tried to extend Bion's theoretical ideas into an empirical system (Stock and Thelen, 1958; Thelen, 1956; Thelen, Stock, and others, 1954), tried to study some of the same phenomena but via an entirely different route. Bion began from a "system-driven" orientation. Presumably, he cared about interpersonal behavior within therapy groups. He apparently pursued that interest via a system-driven form of the theoretical path. Thelen and colleagues tried to test Bion's theoretical notions by applying some data collection and analysis methods (in this case, a system differing from Bales's IPA). In other words, Thelen and colleagues took step 3 (testing a set of hypotheses) to follow up Bion's step 1 and step 2 work (developing a theory or set of hypotheses about therapy groups) on the theoretical path. The body of evidence from the Stock and Thelen work was in considerable contrast to Bales's findings. For example, they did not find the neat progression through problem solving phases that Bales and colleagues (e.g., Bales and Slater, 1955; Bales and Strodtbeck, 1951) had consistently found. Furthermore, when Talland (1955) used a system-driven empirical approach for observation of therapy groups, using Bales's IPA, his results agreed with the Stock and Thelen work in that it also did not find the postulated "phase sequence." The therapy groups were not full-fledged problem solving groups in Bales's sense of that term because they did not undertake and carry to completion a single task within a single group meeting. Instead, they pursued a single task (or a multi-task primary mission) throughout the life span of the group. Moreover, the task of therapy groups is in large part socioemotional, and the two categories become confounded. Thelen's therapy groups did not show a pattern of activities divided neatly between "task" and "social-emotional" behaviors, as Bales's ad hoc groups had done. On the other hand, Talland's groups did show a task/instrumental patterning. Talland's work used Bales's observation system, and use of that system virtually guarantees a task versus instrumental patterning of observations (McGrath, 1984).

The point of this example (and indeed of the whole chapter) is not that one of these is the "right" approach and the others are "wrong" approaches. The point is precisely to the contrary; namely, that none of these or any of the other pathways provides the "right" approach, or even a very useful approach, if used to the exclusion of all others. Research knowledge, that is, research information we hold with confidence, requires convergence of findings from research following differing paths.

Actually, the interest by Bion, Thelen, and Talland in interaction process in "real" groups led to a fuller understanding of the very phase sequence and equilibrium processes that Bales was interested in exploring and that Thelen's and Talland's data seemed to disconfirm. On the other hand, such applications of Bales's techniques and concepts to behavior in therapy groups in turn led to a better understanding of therapy groups, and of how their interaction patterns related to those of other kinds of groups. Psathas (1960) later showed that therapy groups did indeed follow the postulated "phase sequence," but did so over their entire life span and not within a single session—a conclusion that makes eminently good sense in terms of the original concepts that Bales set out to study. So, Bales's concept-driven experimental approach, coupled with the Bion/Thelen system-driven theoretical approach, Talland's system-driven empirical approach, and Psathas's concept-driven theoretical approach advanced our understanding of group interaction processes far more than could have been the case if an equivalent amount of research had all been done by any one of those four pathways (or any other single pathway for that matter).

Example 3: Need Achievement

Still another extended illustration can be drawn from past research on need achievement. McClelland and colleagues' early work on this problem (for example, McClelland et al., 1953) was clearly on the experimental path. It was driven by Murray's need and press concepts and concentrated on developing methods for measuring, scoring, and manipulating need achievement (N-Ach). That early work was based on substantive systems involving (male) college students doing artificial laboratory tasks. These researchers developed a highly elaborated (probably concept-driven) theory of need achievement (e.g., Atkinson and Feather, 1965). At the same time, some method-driven work on the empirical path and some method-driven work on the experimental path was done to develop and improve instrumentation both for the achievement motive and for need affiliation, need power, and other social motives. Work in the related area of women's fear of success (e.g., Horner, 1965; Hoffman, 1977) was much more system-driven. The researchers cared about the questions of women's social motivations. They initially used the N-Ach methodology to generate a system-driven set of observations. Subsequently, they developed system-driven sets of hypotheses about the causes and consequences of women's conflictful orientations toward task achievement (e.g., Horner, 1972). McClelland's (1965) later work, using achievement motivation as a broad concept virtually to account for the rise and fall of social systems, also could be regarded as a set of concept-driven hypotheses to which he brought various methods for measuring the concepts as suited the different cases he wished to study.

Thus, research involving need achievement can and has been done following all six pathways, just as was the case for uses of factor analysis and for the study of group interaction process. And as with those other areas, diversity of research approaches broadens and deepens our understanding of those focal problems.

PATHS, PATHWAYS, AND RESEARCH ORIENTATIONS

There has been much talk about the distinction between basic and applied research. A great deal of that talk has been misleading. The distinction involves both more and less than is usually reckoned in those discussions. The VNS offers, we think, a more systematic way to look at the distinction between basic and applied approaches to research than the somewhat piecemeal treatment it has usually received. In the VNS, the three-step form of stage two and the alternative paths and pathways for conducting stage two provide an entrée to a systematic analysis of the basic versus applied distinction. In this section, we will try to clarify the basic/applied distinction and its implications.

Basic Versus Applied Orientations to Research

We are all familiar with the basic versus applied research distinction. By that distinction we mean doing research that focuses on discipline-generated problems (variously called basic or fundamental, and considered to have generality and conceptual breadth) versus doing research that focuses on the solution of some real-world problems (variously called applied, relevant, or problem-oriented research and considered to have self-evident importance). Note that this distinction refers to the intentions of the researchers, rather than to the outcome of the studies. A study focusing on a discipline-generated problem can lead to useful applications. A study aimed at solving a system problem can yield basic knowledge. The basic/applied distinction refers to the researchers' initial focus or purposes, not to the study's consequences.

For the most part, people have used the basic/applied distinction to justify what they do. They often have done so by

condemning the weaknesses of what others do. There has been considerable acrimony in the debate and considerable misunderstanding of the other's viewpoint. Sometimes, for example, our graduate training programs treat applied research as if it were somehow different-in-kind from what is otherwise done at the same institution in the name of research. There is often the implication that the people who practice those alien arts of applied research are somehow a breed different from the people who do basic research. A similar but opposite kind of world view is often evident in the conference room and coffee room conversations of researchers involved in applied study of particular substantive systems.

There have been many recent discussions of basic and applied research approaches, and comparisons between the two, in the literature on research in the social and behavioral sciences (see, for example, Berkowitz and Donnerstein, 1982; Boehm, 1980; Calder et al., 1981, 1982, 1983; Cook and Campbell, 1979; Ellsworth, 1977; Lynch, 1982, 1983; Mackenzie and House, 1978; McGrath and Brinberg, 1983; Weick, 1967). A sizable part of the discussion of applied and basic research has been embedded in works ostensibly comparing different research strategies (e.g., lab versus field research)—a quite different set of issues. Some of the literature on basic and applied research approaches has favored one over the other; some has presented a more balanced view.

A few researchers have tried to use the distinction constructively, to specify different and complementary research approaches. In one of the most cogent recent presentations on this topic, Boehm (1980) discusses the "academic" or "traditional" model in comparison to the "real world" or "empirical" model for doing research in organizations. She notes how systematic separation of the two leads to mutual pejorative labeling ("dust bowl empiricists," "ivory tower types"). She lays out a series of important trade-offs between the two and argues cogently for progress by cross-fertilization of the two approaches. We have little to quarrel with in her analysis. We hope to take it several steps further, by using the VNS to identify more than two paths, discussing some of the differences between the paths in terms of their impact on the research process, and urging multiple path approaches as a necessary strategy for the field.

VNS and the Basic/Applied Distinction

Earlier in this chapter, we noted six separate pathways along which one can pursue research in stage two of the research process. Two of them lie on what we have called the experimental path, two on the theoretical path, and two on the empirical path. Cross-cutting those three paths or styles of research are several underlying approaches or orientations toward research, that are reflected in which of the three domains the researcher chooses as the starting place.

A researcher can start stage two in any one of the domains and be faced with two "pathways" along which to proceed in carrying out step 2 of that stage of the research process. The two pathways that start in the conceptual domain are **basic research** paths, as that term is often used. The two pathways that start in the substantive domain are **applied research** paths, as that term has been used. The remaining two pathways—the two that start in the methodological domain—are what we will call **technological research**. They have to do with the development of research technology. We will focus on applied versus basic research in this section because that pair of orientations has been the subject of much debate, but we will add some comments near the end of the section about the technological orientation.

When we talk about basic research, we usually have in mind research activities that follow the pathway here characterized as concept-driven design. The researcher begins with an interest in the conceptual domain, draws upon the methodological domain to construct a design, and then implements that design on some substantive system (e.g., Thurstone's development of attitude research, Bales's development of IPA, and McClelland's initial

work on N-Ach, discussed in the examples of the preceding section). When we talk about applied research, we usually have in mind research activities that follow the pathway here characterized as system-driven observations. The researcher begins with an interest in some substantive system, goes to the methodological domain to develop a set of observations, and then goes to the conceptual domain to interpret that set of observations (e.g., Talland's work on interaction in therapy groups, some of Horner's early work on fear of success, and Guilford's work on human abilities, discussed in the examples of the preceding section). These two pathways are distinctive, both in where they start (the conceptual and substantive domains, respectively) and in which domain they deal with last (the substantive and conceptual domains, respectively). For each of them, the other's preferred domain is left until last. *What distinguishes basic and applied researchers from one another, or at least what seems to set them in opposition to one another, is not so much that each prefers a different domain as a starting point, but rather that each typically seems to avoid the other's preferred domain until the last.* What makes the basic researcher sometimes anathema to the applied researcher is not so much that the basic researcher has a central interest in conceptual matters, but rather that the basic researcher seems to have a disdain for substantive systems. To the applied researcher, the basic researcher seems to go about research in a way that says: "any old substantive system will do, just so it fits my design." Conversely, the applied researcher is sometimes held in disdain by the basic researcher not so much because he or she has a central interest in some substantive system, but rather because the applied researcher seems to have a disdain for and to avoid conceptual systems. To the basic researcher, the applied researcher seems to go about research in a way that says: "Any old concepts will do, just as long as I can use them to label my data."

But these two quite separate pathways are not the only possibilities. A basic researcher could just as well start with the conceptual domain and go to the substantive domain—to build a concept-driven set of hypotheses rather than a concept-driven design (e.g., McClelland's later work on achievement motivation and social system productivity, and Rokeach and Fruchter's confirmatory use of factor analysis, discussed in the examples of the preceding section). Similarly, the applied researcher could just as well start in the substantive domain and go to the conceptual domain—to build a system-driven set of hypotheses rather than a system-driven set of observations (e.g., Bion's work on group process in therapy, and Horner's later work on a theory of female motivational conflict, discussed in the examples of the preceding section). If both did so, that would put the basic and applied researcher on the same theoretical path, and on a closely linked pair of pathways, and both would have a quite similar next step: bringing in material from the methodological domain to test the respective sets of hypotheses.

Those two pathways are alike in two important respects, even though they involve different first steps. First, for both of them, step 2 produces a structure that is a set of hypotheses (not a design and not a set of observations). Second, for both of them, step 3 requires combining methods with that structure. Those similarities point the way to a potential rapproachment of basic and applied efforts and interests.

We do not want to overstate the similarities of these two pathways on the theoretical path. A concept-driven set of hypotheses is not the same thing as a system-driven set of hypotheses on the same focal problem, even though both involve the same step 2 structure (a set of hypotheses) and even though step 3 for each involves a fitting of material from the methods domain. The starting domain dominates the matching that occurs in step 2. When one builds a system-driven set of hypotheses, one makes sure that concepts are selected and organized to reflect properly the aspects of the substantive system in which one has an interest (and because of which one is doing the research in the first place). When one builds a concept-driven set of hypotheses, one makes sure that aspects of the substantive system are drawn,

differentiated, and organized in a way that reflects properly the sets of concepts and relations in which one has an interest. Parallel statements could be made, of course, with respect to step 2 of the other pathways.

To build a concept-driven set of hypotheses involves moving from the more abstract world of concepts toward the more concrete. To build a system-driven set of hypotheses requires moving from the more concrete world of the substantive domain toward the more abstract. It is likely, though not logically necessary, that in any given case a system-driven set of hypotheses will still be more concrete and less abstract than a concept-driven set of hypotheses. This differential in abstractness/concreteness is likely to be reflected at all three levels: elements, relations, and embedding system. So the constructs and postulates of the system-driven set of hypotheses probably will be more "substance laden," whereas the constructs and postulates of the concept-driven set of hypotheses will reflect a more abstract and formal "theoretical system." We can see this to some extent in the Bales system and the Bion/Thelen schema for studying group interaction process. The Bales system is much more abstract and formal in its formulation and structure (having begun as a concept-driven enterprise), whereas the Bion/Thelen system is much more closely tied to substantive features of the therapy groups for which it was designed (having begun as a system-driven endeavor). Each of the interaction schematas reflects the starting place—the preferred domain—of its protagonists. In other words, *the first domain—which represents the researchers' fundamental interest in the ongoing research endeavor—structures all later choices; subsequent choices are made to ensure that that central interest gets well-served.*

Note, however, that these two theoretical pathways (system-driven hypotheses and concept-driven hypotheses) are not sharply distinguished in regard to either rigor or relevance, which are the two attributes that have been the hallmarks of the stereotypes regarding basic and applied orientations. The "typical" basic researcher, following the concept-driven design pathway, is regarded as "rigorous" because that experimental path or style emphasizes methods and research designs that offer potential rigor when implemented (at a cost, to be sure, in "realism"). Similarly, the "typical" applied researcher, following the system-driven observations pathway, is regarded as conducting highly relevant research because that empirical path emphasizes methods and data that can include, from the outset, direct records of those parts of the substantive domain that are of special interest or relevance to the researcher. But when the applied and basic researcher both choose their own form of the "theoretical path" (i.e., system-driven and concept-driven hypotheses, respectively), neither the rigor nor the relevance of their efforts has yet been established (or precluded) at that second, structure-building step. When either one of them is carried further—by bringing in methods with which to test these sets of hypotheses—that step 3 effort can offer rigor and relevance in about equal measure on either the system-driven or the concept-driven pathways. When both applied and basic research proceed along the theoretical path, there is no strong reason why either one should look rigorous and the other sloppy, or why either one should look relevant and the other "pointless." Both should seem relevant because both have given the substantive system relatively high priority. Both should seem logically rigorous because both have given the conceptual domain relatively high priority. Just how empirically rigorous they will seem depends on what each does in the final step of stage two, when methods are selected with which to test the set of system-driven, or the set of concept-driven hypotheses.

When the third domain is brought in, elements and relations within it are selected and organized to fit the structure already built from the first two domains. No matter which domain is the preferred starting place for a research endeavor, there is more constraint on choices of the elements and relations from the second domain, and still more constraint for the third. Choosing elements and relations from the substantive domain to implement a design is a much more restricted set of choices than picking some substantive system elements and relations as a *starting point* for research. Likewise, picking concepts to interpret a set of observations is a much more restricted set of choices than is selection of some conceptual elements and relations as the subject for study. And, of course, picking methods to help test hypotheses is a much more restricted set of choices than selecting methodological elements and relations for a method-driven study.

The researcher will adjust what gets selected from the third domain to fit the structure built by the combination of the first two domains, just as he or she will adjust what is selected from the second domain to fit the choices already made in relation to the first—the domain of primary interest. By step 3, bringing in the third domain, researchers' choices are often quite constrained. All stage two research endeavors cover all three domains, and all stage two research ends with a body of evidence that reflects a three-domain combination. But the specific pattern of that evidence, and its interpretation, depends crucially, we believe, on which of the six possible pathways the researcher used to reach that end point.

In summary, we thus far have made several major points about basic and applied research, as viewed within the VNS. First, we identified the basic researcher as one who is primarily interested in (prefers) the conceptual domain, and the applied researcher as one who is primarily interested in (prefers) the substantive domain. Second, each of these two approaches or research orientations has two pathways available for stage two research, and those two pathways link them to two different research styles or paths. The most frequent pathway chosen for each—the concept-driven design for the basic researcher and the system-driven set of observations for the applied researcher—puts them on quite different and apparently "opposite" paths. Furthermore, for those typical pathways, each seems to display a negative preference for the other's preferred domain. This seemingly negative preference for the other's domain underlies much of the negative tone of the discourse about basic versus applied research in the literature. These typical pathways also give some credence to the stereotypes by which applied and basic research come to be regarded as antithetical to one another.

But each of those research orientations has another potential pathway—concept-driven hypotheses and system-driven hypotheses, respectively. That set of choices puts the applied and basic researcher not on opposite paths but on a parallel and closely linked pair of pathways that follow the same overall path or style: the theoretical path.

The Third Orientation: Technological Research

Thus far, we have discussed two orientations toward research: basic research and applied research. The VNS implies a third research orientation, one we have labeled the *technological research* orientation, in which the researcher's primary interest is in some particular portion of the methodological domain. We think this implication points to a feature of our field that is important, though given little attention heretofore.

There surely are researchers in the behavioral and social sciences who seem primarily interested in building tools by which we can gather and analyze information about the phenomena of our field. Such "tool building" is the hallmark of the stage one researcher who specializes in the methodological domain, just as there are stage one system experts who specialize in the conceptual and substantive domains. But sometimes such interest in tool development carries over into stage two activities, and manifests itself in the form of "method-driven" approaches to stage two research. Just as some researchers begin their stage two activities by focusing on the conceptual domain (as the basic researcher does) or the substantive domain (as the applied researcher does), so some others begin stage two by focusing on the methodological domain (the technological researcher). And, just as the basic and applied researchers can fulfill their central purpose via either of two different pathways, so, too, is the technological researcher faced with two different pathways. The technologist may take either the experimental path (by building a method-driven design and then implementing it) or the empirical path (by building a method-driven set of observations and then interpreting it). The first of these puts the technologist close to the basic researcher. The second puts the technologist close to the applied researcher.

When the technologist follows the experimental path (i.e., method-driven design), there will be a tendency to pay less attention to the substantive domain (as Thurstone did in his factor analysis of the dimensions of boxes, noted in an example earlier in this chapter). When the technologist follows the empirical path (i.e., method-driven sets of observations), there will be a tendency to pay less attention to the conceptual domain. Examples of that latter pathway would include many of the "dust bowl empiricist" uses of mental tests or factor analysis, alluded to earlier, or public opinion surveys (that also show the tendency to give less attention to conceptual interpretation). Examples of the method-driven empirical pathway also would include sophisticated uses of such bodies of data as the Consumer Expenditure Survey, the Human Relations Area Files, and various survey data archives.

Pathways arising from the technologist orientation, like those arising from the other orientations, court trouble when used as the sole pathway for the study of a certain focal problem, but enrich the research process when used in combination with other pathways. Like the other two orientations, the technological orientation reflects the researcher's central purposes. And it affects which stage two validity issues will and will not be encountered, the topic of the next chapter.

CLOSING COMMENTS

The basic/applied distinction too often has been used in our literature as a shibboleth for justifying the exclusive use of particular research orientations. That, in turn, has at times led to a kind of intellectual snobbery, prejudice, and "segregation" that has not served us well. In this chapter we have tried to show that use of only one of the possible study paths and research orientations hurts the full development of the field. At the same time we have tried to show some ways in which the very real differences in approaches to research can give us new leverage and the opportunity to understand additional features of the material we choose for study from the substantive, methodological, and conceptual domains.

The problem suggests an analogy to someone exploring a territory with the aid of a map. Pathways are different routes to some destination. You can take a different pathway each time you go to that destination, or you can take the same path each time. If you use the same path repeatedly: (a) you will learn well how to go from A to B; (b) you will develop a well-marked route (a paradigm) that others may follow; (c) you will probably attract many followers along that available route; but (d) you will learn nothing at all about the region except the part that lies alongside your favorite road.

So it is with the study of complex problems in various areas of behavioral and social sciences. We need to give special attention to different pathways and to the use of multiple pathways. Research in our field involves very complex problems, hence much territory to be explored. The very complexity of the research implies that we need to make much progress in regard to all three domains. We need conceptual models sufficiently complex to be a reasonable match for the material to be modeled. We need methodological tools adequate to handle that complexity. We need new insights regarding the nature and functioning of those substantive systems to which our research is intended to apply. To forego potentially fruitful research gains in any one of those domains because of misguided insistence on the importance of one particular pathway would be especially damaging for any discipline in the behavioral and social sciences. All paths are useful, and ultimately all are necessary, if we are to explore fully and hence come to understand fully, those substantive, methodological, and conceptual systems that we study.

Using multiple paths is important for yet another reason. Steps 2 and 3 on each pathway can be regarded as "sites" at which different validity issues arise. When stage two research is done along any one of the pathways, the researcher encounters—and thus potentially can deal with—certain aspects of the correspondence validities of stage two. But any one pathway does not even encounter most of the "sites" at which those validity issues arise. Hence, research along any one pathway cannot deal with most issues of validity—that is, cannot *reduce uncertainty* with respect to most of them. And, if all research on a given focal problem follows the same pathway, that body of research information will contain more uncertainty than needs to be the case. The next chapter will explore those stage two validity issues in detail.

GLOSSARY OF TERMS

Experimental Path:

Step 2 of this path consists of the combination of elements and relations from the conceptual and methodological domains to form an instrumental structure that we call a **study design**. The combination of elements from these two domains form **operations**. The combination of relations from these two domains form a set of **planned comparisons**.

Step 3 of this path involves **implementing a design**. When the phenomena (i.e., elements selected from the substantive domain) are combined with the operations, the result is a set of **variables**. At the relations level, the combination of the patterns of phenomena with planned comparisons results in a **set of findings**. The embedding system for the variables and the set of findings is a **body of evidence**.

Pathways of the experimental path:

(1) A **concept-driven design**. For this pathway, the researcher has primary interest in studying certain elements and relations from the conceptual domain (step 1). In step 2, the researcher selects elements and relations from the methodological domain to fit the concepts and relations already selected for study. The third step is to select elements and relations from the substantive domain to

implement the design. When a pathway starts in the conceptual domain, we describe it as involving a **basic research orientation.**

(2) A **method-driven design.** For this pathway, the researcher has primary interest in studying certain elements and relations from the methodological domain (step 1). In step 2, the researcher selects elements and relations from the conceptual domain to fit the modes of treatment and the comparison techniques already selected for study. The third step is to select elements and relations from the substantive domain to implement the design. When a pathway starts in the methodological domain, we describe it as involving a **technological research orientation.**

Theoretical Path:

Step 2 of this path consists of the combination of elements and relations from the conceptual and substantive domain to form an instrumental structure that we call **a set of hypotheses.** The combination of elements from these two domains form **constructs.** The combination of relations from these two domains form **hypotheses.**

Step 3 of this path involves **testing a set of hypotheses.** When the modes of treatment (i.e., elements selected from the methodological domain) are combined with the operations, the result is a set of **variables.** At the relations level, the combination of the patterns of phenomena with planned comparisons results in a **set of findings.** The embedding system for the variables and the set of findings is a **body of evidence.**

Pathways of the theoretical path:

(1) A **concept-driven set of hypotheses.** For this pathway, the researcher has primary interest in studying certain elements and relations from the conceptual domain (step 1). In step 2, the researcher selects elements and relations from the substantive domain to fit the concepts and relations already selected for study. The third step is to select elements and relations from the methodological domain to test the set of hypotheses. When a pathway starts in the conceptual domain (as in the concept-driven design pathway), we describe it as involving a **basic research orientation.**

(2) A **system-driven set of hypotheses.** For this pathway, the researcher has primary interest in studying certain elements and relations from the substantive domain (step 1). In step 2, the researcher selects elements and relations from the conceptual domain to fit the phenomena and patterns of phenomena already selected for study. The third step is to select elements and relations from the methodological domain to test the set of hypotheses. When a pathway starts in the substantive domain, we describe it as involving an **applied research orientation.**

Empirical Path:

Step 2 of this path consists of the combination of elements and relations from the methodological and substantive domain to form an instrumental structure that we call **a set of observations.** The combination of elements from these two domains form **indices.** The combination of relations from these two domains form **sets of arrays.**

Step 3 of this path involves **interpreting a set of observations.** When concepts (i.e., elements selected from the conceptual domain) are combined with the indices, the result is a set of **variables.** At the relations level, the combination of sets of arrays with conceptual relations results in a **set of findings.** The embedding system for the variables and the set of findings is a **body of evidence.**

Pathways of the empirical path:

(1) A **method-driven set of observations.** For this pathway, the researcher has primary interest in studying certain elements and relations from the methodological domain (step 1). In step 2, the researcher selects elements and relations from the substantive domain to fit the modes of treatments and planned comparisons already selected for study. The third step is to select elements and relations from the conceptual domain to interpret the set of observations. When a pathway starts in the methodological domain (as in the method-driven design pathway), we describe it as involving a **technological research orientation.**

(2) A **system-driven set of observations.** For this pathway, the researcher has primary interest in studying certain elements and relations from the substantive domain (step 1). In step 2, the researcher selects elements and relations from the methodological domain to fit the phenomena and patterns of phenomena already selected for study. The third step is to select elements and relations from the conceptual domain to interpret the set of observations. When a pathway starts in the substantive domain (as in the system-driven set of hypotheses), we describe it as involving an **applied research orientation.**

Chapter 4

STAGE TWO
Validity as Correspondence

In Chapter 1, we presented the general meaning of validity in stage two as **correspondence** between the elements and relations selected from each of the domains. In Chapter 3 we described six different pathways for conducting research in stage two. In this chapter, we will argue that the particular forms of the correspondence validity issues depend on which of the six pathways is followed in a study. The starting point for each of the pathways is one of the three domains. Much of our discussion on correspondence issues is organized by domain. The first section treats the meaning of validity in stage two, and the **features of relations** that are involved in the **matching process**. The second section discusses the matching process in terms of **information gain** and uncertainty reduction. The third section deals with how researchers cope with various types of mismatches that can occur in studies that follow different pathways. The final section provides a summary of the validity issues treated in the chapter and a discussion of some of their implications for research.

VALIDITY AS CORRESPONDENCE

When we work in stage two, we would like to bring together elements and relations from the three domains so that they fit together perfectly. Step 1 of stage two involves specifying a set of elements and relations from one domain—the domain of most interest to the researcher. Step 2 involves forming a design, developing a set of hypotheses, or building a set of observations by combining elements and relations from a second domain with those already chosen for study in step 1. Step 3 involves implementing that design, testing that set of hypotheses or explaining that set of observations by combining elements and relations from the third domain with that instrumental structure built in step 2. Correspondence issues arise at steps 2 and 3 of this stage on each of the alternative research paths because those steps involve a *cross-domain fitting* of elements and relations.

Correspondence is not just a fitting of duplicate sets of elements and relations from each of two domains. Rather, it is the *meshing together of two complementary networks of relations.* The nature of that complementarity differs depending on which two of the domains are being combined in step 2, or which of the instrumental structures are being combined with a third domain in step 3. That is, the nature of the fit depends on which of the alternative paths is being followed. For example, step 2 of the pathway that builds a concept-driven design fits a network that consists of a set of comparison techniques to a network that consists of a set of conceptual relations. The result of that combination is a network that consists of a set of planned comparisons—what is often called a study design.

One central premise of VNS, already noted, is that research is concerned primarily with relations. Stage two of the research process involves matching networks of relations from different domains. But what is matched ultimately is a set of features for the relation between each pair of elements in the network, for each of two domains (or, in step 3, for an instrumental structure and a domain).

The relations from the three domains are different in kind. In the conceptual domain, relations are abstractions, *symbolic representations.* In the methodological domain relations are *procedures* for measuring and comparing. In the substantive domain, relations are *processes* of interest.

The domains also differ in the functions they serve within the research process. The function of materials from the conceptual domain is to *specify* what levels of various features are expected in a relation. The function of materials from the methodological domain is to distinguish among levels of features of the relations to be examined. The function of materials from the substantive domain is to display whatever level of features of the relations that might occur in the "real-world" substantive system from which the material is sampled. Thus, when the researcher attempts to combine elements and relations from two of the domains (or from an instrumental structure and the third domain), he or she is faced with the problem of combining elements and relations that differ in kind and in function.

Below, we will characterize the relations within each of the domains in terms of seven features: *presence, temporal order, logical order, direction, functional form, stochasticity, and temporal stability.* In the rest of this section, we will present those features of relations, and then describe tasks involved in the matching process, in relatively formal terms and in considerable detail. Then, in the next section, we will discuss the matching process in terms of potential information gain and uncertainty reduction. In the third section of the chapter, we will examine how researchers can cope with mismatches of various types. In that section, we will try to present the ideas in less formal language, and will use examples to clarify key points.

Features of Relations

Correspondence validity involves the fit on a number of features intrinsic to each relation between a pair of elements, i and j. That fit needs to be examined with respect to each pair of elements, ij, in the network of relations drawn from each domain.

Some features intrinsic to such relations are as follows:

(1) Statements of the *presence* or *absence* of a relation for any pair of elements, i and j.
(2) Statements of the *temporal order* of the pair of elements, i and j. These can include statements that i precedes j, and j precedes i, or that i and j are simultaneous.
(3) Statements of the *logical order*, or logical direction, of the relation between i and j. These can include statements that i leads to j, that j leads to i, or that i and j each affect the other.
(4) Statements of the *direction of any functional relation* between i and j. These can include statements that i and j are positively related, or that i and j are inversely related.
(5) Statements of the *form of any functional relation* between i and j. These can include statements that i and j are related in a linear (and therefore monotonic) form, or in a nonlinear but monotonic form, or in a nonmonotonic but single-peaked form, or in a multiple-peaked form of more complexity, such as a recurring cycle.
(6) Statements of the *deterministic* or *stochastic* character of the ij relation.
(7) Statements of the *temporal stability* of the ij relation. These can include statements that the ij relation is stable over time, or that it is stable but with minor fluctuations, or that it changes in a stable (regular) pattern, or that it changes in a variable pattern.

These seven classes of statements refer to features intrinsic to relations between two elements, i and j. They have representations in each of three domains. Several additional features of any ij relation involve elements extrinsic to that ij pair but intrinsic to the network. That is, they are features that tie each ij relation to (all) other elements and relations in the network under consideration. These extrinsic features of ij relations include the following:

(8) Statements of the relation of each other element, g, to element i of the ij relation. These elements, g, represent *antecedents* of the ij relation. These gi relations are to be assessed on the seven intrinsic features of relations listed above.
(9) Statements of the relation of each other element, k, to element j of the ij relation. These elements, k, represent *alternative antecedents* (alternative to i) of the ij relation. These kj relations, also, are to be assessed on the seven intrinsic features of relations listed above.
(10) Statements of the relation of each other element, l, to both i and j or to the ij relation itself. These elements, l, represent *moderator variables* that may alter or influence or mediate or modulate the ij relation. These l-ij relations, also, are to be assessed on the seven intrinsic features of relations listed above.

These extrinsic features, 8, 9, and 10 above, contain all of the "extra-relation" effects on i, j, and ij. These are the locations of many of the effects, anticipated and otherwise, that arise in research: "third variables," mediators, suppressors, and the like. In fact, these extrinsic relations are the locations for many of the effects that Campbell and Stanley (1966) called *plausible rival hypotheses,* and that Blalock (1982) called *sets of auxiliary conditions*. They are also the locations for experimenter effects; for carryover and context effects in factorial designs; for subject response styles, response sets, and other reactivity effects.

For example, VNS would characterize a "history" effect within an experimental design as an "alternative antecedent" (feature 9) of a relation, ij. The historical event, k, may have had a differential impact on j for cases receiving different levels of i. We would construe a "social desirability" effect as a "moderator" (feature 10) of an ij relation in which the level of i (the particular contents of the question) influences the level of l (social desirability), which in turn influences the level of j (the response to the question). Thus, many of the issues or problems that have been associated with the validity of study designs and measures are regarded, in VNS, as features extrinsic to the relation ij that is being studied—but features intrinsic to the *network* of relations being explored.

It is important to keep a network frame of reference in considering the features of relations, especially the extrinsic features.

When we are considering only a single relation, we must take into account these extrinsic features that affect that relation, and treat them as "outside" that relation. But when we are considering evaluation of all features of all relations in a network, we can construct all of the information by assessing only the seven intrinsic features for each relation, but doing so exhaustively for all pairs of elements in the network.[1]

Tasks of the Matching Process

The first tasks in the matching process require the researcher to select a network of elements and relations from the domain of his or her primary interest, and then to specify features of all pairs, ij, of elements in that network. These two tasks are task a and task b in Table 4.1. They coincide with step 1 of stage two of the research process.

The matching process in step 2 entails four tasks (c, d, e, and f in Table 4.1). The researcher first selects a network of elements and relations from a second domain (task c); then pairs each element in that second domain with an element from the first domain (task d). Decisions about which elements match up with which are always decisions by fiat on the part of the researcher. What concept we will use to label a particular measurement scale, for example, or what method we will use to detect a phenomenon, or what event we will pick as the embodiment of a particular concept, are all matters of judgment. They are to be evaluated in terms of what has been called *face validity*.

The researcher then specifies features of all of the relations, ij, in the network from the second domain (task e). At this point, features of relations are specified in each domain for a matched set of relations from the two domains. Now, the researcher needs to assess the degree of correspondence or fit in the cross-domain match of all of the features of all of the paired relations, ij, in the two domains (task f). The nature of that assessment and the implications of cross-domain fit or lack of it are the focus of the rest of this chapter.

TABLE 4.1
Tasks in the Matching Process for Stage Two

Step One of Stage Two: Specify a network of elements and relations from the preferred domain.	
Task a:	Select a network of elements and relations from domain 1.
Task b:	Specify features of all pairs of elements, ij, in the network from domain 1.
Step Two of Stage Two: Match the network from the first domain with a network of elements and relations from a second domain, to build an instrumental structure.	
Task c:	Select a network of elements and relations from domain 2.
Task d:	Pair every element in domain 1 with a matching element from domain 2.
Task e:	Specify features of all relations, ij, in the network from domain 2.
Task f:	Compare each feature of relations, for domains 1 and 2, for each relation ij, to determine its degree and form of fit.
Step Three of Stage Two: Fit the instrumental structure from step two with a network of elements and relations from domain 3.	
Task g:	Select a network of elements and relations in domain 3.
Task h:	Pair every element from domain 3 with a matching element in the instrumental structure built from domains 1 and 2 in step two.
Task i:	Specify features of all relations ij, in the network from domain 3.
Task j:	Compare each feature of relations, for the instrumental structure built from domain 1 and 2 and the network from domain 3, for each relation ij, to determine its degree and form of fit.

The matching process in step 3 entails four more tasks (g, h, i, and j in Table 4.1). These are parallel to the four tasks of step 2 (tasks c, d, e, and f in Table 4.1) except that steps h and j involve fitting material from the third domain to the instrumental structure already built in step 2, rather than the fitting of material from two domains.

As noted above, materials from different domains serve different functions in the research process. They therefore pose different kinds of questions regarding features of relations and regarding the cross-domain matching process.

For feature 1, the feature dealing with presence or absence, the questions we ask for each of the domains are these: Does material from the conceptual domain *specify* that the ij relation is present or is not present? Can material from the methodological domain *distinguish* whether ij is present or is not present? Can material from the substantive domain *display* either presence or absence of an ij relation?

For feature 2, temporal order, the key questions are these: Which temporal order does the conceptual domain specify for ij—i precedes j, j precedes i, i and j are simultaneous? Can the methodological tools chosen distinguish the three temporal orders from one another? Can the substantive materials sampled display any of the three temporal orders?

For feature 3, logical order, the key questions are these: Which logical order does the conceptual materials specify for ij—i leads to j, j leads to i, i and j are interdependent? Can the comparison techniques drawn from the methodological domain distinguish among these logical orders? Can the substantive system, as sampled, display any of these logical orders?

For feature 4, direction of function, the key questions are these: Which direction does the conceptual domain specify, positive or inverse, for the ij relation? Can the methodological tools selected for use distinguish positive from inverse relations? Can the substantive system as sampled display either a positive or inverse relation?

For feature 5, form of function, these are the key questions: Which functional form does the conceptual domain specify for ij—linear, nonlinear but monotonic, non-monotonic but single peaked, or multiple peaked? Can the comparison technique chosen from the methodological domain distinguish among all of these functional forms? Can the sample of the substantive system chosen for inclusion in the study display any of these functional forms?

For feature 6, dealing with the deterministic or stochastic character of the ij relation, the key questions are these: Does the conceptual domain specify that ij is deterministic or stochastic? Can the methodological tools chosen for use distinguish among those patterns of stochasticity? Can the substantive system as sampled display any of them?

For feature 7, temporal stability, these are the key questions: Does the conceptual system specify that ij is stable over time, stable but with minor fluctuations (i.e., error), changing over time but in a regular way, or varying over time in an irregular way? Can the comparison technique chosen from the methodological domain distinguish these different patterns of stability and change from one another? Can the portion of the substantive system chosen for study display any of these patterns of temporal stability and change?

INFORMATION, UNCERTAINTY, AND THE MATCHING PROCESS

Potential Information Yield in Each Domain

Information gain from a study depends on the number of alternative outcomes that were plausible before the study and are excluded by results of the study. Such information gain also can be described as the reduction of uncertainty in the features of the relations under study. So the amount of potential information that can be derived from a study about a particular relation, ij, is a function of the degree to which the choices from each of the three domains let the researcher deal with (i.e., differentiate among) all of the possible values of each of the seven features of that relation. "Deal with" means different things for the different domains.

In the methodological domain, from the point of view of potential information yield, the comparison techniques selected for use *ideally* ought to distinguish among all possible values of all features for each of the ij relations being studied. Such an ideal comparison technique would differentiate among all possible types of functional forms, of temporal stability, and the like. Such an ideal comparison technique could be said to contain maximum potential information. In any given actual case, however, the comparison techniques used in a study are likely to be more limited in their ability to discriminate among values of some features of some ij relations. For example, use of a comparison technique that involves contrasting the means of two groups does not permit differentiating among different functional forms of that ij relation.

In the conceptual domain, from the point of view of potential information yield, the conceptual relation, ij, selected for use ideally ought to specify a particular value that is expected to occur for every feature of every relation, ij, in the conceptual network being considered. To do so is, at the same time, to specify that all other values of every feature of each relation, ij, *will not* occur. Therefore, if the relation ij in the conceptual domain is specified as present, as i leads to j, as positive in direction and linear in form, that specification also implies that ij is *not* of the form "j leads to i" (and other possible alternative forms), is *not* inverse in direction, and is *not* nonmonotonic, monotonic nonlinear, step function, and so on. Such a conceptual specification, too, would contain maximum potential information. In any given actual study, however, it is likely that the network of relations selected in the conceptual domain will not specify a single value for each feature of all relations. Therefore, there will be more than one plausible interpretation of the outcome of such a study. That is, the study will reduce *some* uncertainty, but not all of it, for features of relations that are *underspecified* (in the sense just described). It will yield some information, but less than the maximum potential.

In the substantive domain, from the point of view of potential information yield, the patterns of phenomena selected for use ideally ought to be selected so that they *can display any of the values* (of features of all relations, ij) *that do, in fact, occur in the real-world substantive systems from which they are sampled.* That is, for maximum potential information in the substantive domain, the selection of materials from that domain must *not constrain outcomes* such that certain values (of certain features of some relations) *cannot occur* (i.e., cannot be displayed in the study sample) *even if those values were "true" in "nature."* In any given study, however, it is likely that materials have been drawn from the substantive domain in ways that limit what values of

some of the features of relations can be mainfested in the observations. For example, if only a narrow temporal slice of the substantive system is selected for study, that selection will constrain what values of the temporal order and temporal stability features could be displayed in the study's observations, *no matter what temporal order or state of temporal stability was actually the case in that substantive system.* In this example, the manner in which the researchers select the substantive systems for study prevents them from viewing various temporal orders. Researchers cannot gain information about temporal orders *that cannot occur* because of researcher actions, not because of the state of nature. In such a case, the study would yield less than maximum potential information because it would not permit the researcher to tell whether or not certain values of some features of some ij relations existed. Thus, the researcher could reduce some, but not all, of the uncertainty regarding those features of those ij relations.

Potential Information Yield and the Matching Process

One or more of the domains is likely to contain less than maximum potential information on a given feature of a relation ij (i.e., is likely to be *constrained* at less than the ideal case). If at least one domain is thus constrained on a given feature, then the result of cross-domain matching will also be restricted in its potential information on that feature of that relation. The potential information yield for the outcome of a study, with respect to a given feature of a given relation, is determined by the level of potential information in the domain on which that feature of that relation is most constrained.

Potential information yield and remaining uncertainty depend on how many values of features are dealt with in a study. A study can gain information only about the alternative values that are dealt with in all three domains. A study can have maximum potential information (a) by specifying one and only one value in the conceptual domain (hence, by implicitly specifying that all other values will not occur); (b) by being able to identify any value of that feature in the methodological domain; *and* (c) by being able to observe any value of that feature in the substantive domain. A study cannot "deal with" some values of features of a relation ij (a) if its specification in the conceptual domain does not eliminate all but one value; (b) if the methodological tools cannot distinguish among at least some values; or (c) if the substantive observations cannot display at least some values. Uncertainty will remain in regard to alternative values that were thus constrained (i.e., that were not dealt with) in *any* of the domains.

COPING WITH NONCORRESPONDENCE

As we noted in Chapter 3, each of the six pathways for conducting research has its starting point in one of the three domains. The researchers' activities in stage two may be viewed as using the features of the relations selected from the domain of primary interest, as a template or pattern onto which a similar template of features of relations from the second, and later the third, domain are overlayed.

The particular pathway chosen for use reflects the researchers' priority of interest in regard to the three domains, which in turn affects how researchers will deal with lack of fit at steps 2 and 3 of stage two. Generally, features of the relations from the first domain are treated as immutables, not to be modified to fit with those later selected from the second and third domains. The researcher thus constrains what set of elements and relations can be chosen from the second domain. Furthermore, what can be selected from the third domain, to integrate with the structure formed by the combination of the two domains (i.e., a design, a set of hypotheses, or a set of observations) is even more constrained; it must fit into a network of relations already constructed from the first and second domains. For instance, suppose a study follows the pathway involving a concept-driven design. Primary interest is in the relations among some set of

concepts (e.g., attitudes, social pressure, and intention). The researcher is unlikely to modify those concepts and conceptual relations of interest if they do not fit with some set of methods and comparison techniques. Instead, he or she will probably select methods that provide a better fit to the concepts. The researcher is even less likely to modify the concepts to accommodate a lack of fit of elements and relations from some substantive domain in step 3 of that pathway. The researcher following this pathway is more likely to build a design by "shopping around" for methods that will fit the characteristics of the concepts and relations of interest, and then to compare the study by "shopping around" for a convenient set of substantive phenomena on which to implement that design.

Suppose, instead, that a study follows the pathway that involves a system-driven set of observations. Primary interest is in understanding some pattern of phenomena from the substantive domain. In step 1 the researcher needs to lay out all of the features of some particular set of relations that are to be studied. In step 2 the researcher needs to select measures and comparison techniques by which to observe those relations. In step 3 the researcher needs to select elements and relations from the conceptual domain by which to interpret this set of observations.

As in the previous example, the researcher is unlikely to modify material from the first domain—in this case, the substantive phenomena that will be examined. Rather, the measures and comparison techniques from the methodological domain will be selected to fit the network of features of the substantive domain already chosen. The fitting of the concepts and relations to the set of observations in step 3 is even further constrained by the network of features from both the substantive and methodological domains.

A number of researchers have identified separate sets of noncorrespondence issues for the three research paths. Campbell and Stanley (1966) and Cook and Campbell (1979) discuss a number of such noncorrespondence issues (e.g., history, maturation, testing, instrumentation) in their description of experimental research (i.e., the experimental path in the VNS). Bagozzi (1984) describes some noncorrespondence issues involved in the development and testing of theories (i.e., the theoretical path in the VNS). Glaser and Strauss (1967) discuss noncorrespondence issues in their "grounded theory" treatment of empirical research (i.e., the empirical path in the VNS). The VNS treats all three paths for conducting stage two research. Validity issues are fundamentally the same throughout stage two; they are fundamentally a matter of match or correspondence among features of relations across the three domains. But those correspondence validity issues arise in somewhat different guises for work done following the different paths and pathways. Even more important, the researcher is likely to cope with noncorrespondence differently on each of the paths. The discussion to follow presents a concrete example of what happens in the matching process, but does so only for the concept-driven design pathway.

An Extended Example: A Concept-Driven Design to Study Attitudes, Social Pressure, and Behavior

Suppose you wish to study the relations among a set of three concepts: attitudes, social pressure, and behavior. Your step 1 tasks are to lay out all of the elements and relations in your set, and specify the features of those relations. There are three relations here: The attitude-behavior relation (labeled Cij); the social pressure-behavior relation (labeled Ckj); and the social pressure-attitude relation (labeled Cki). Mainly, we will deal with the first of these relations, the attitude-behavior (Cij) relation. We will bring in the third variable (social pressure, Ck) and the other two relations (Cik and Ckj) when they help to make a point.

Suppose further that you postulate that attitude (Ci) is related to behavior (Cj), such that changes in attitude precede and tend to lead to, or cause, changes in behavior in the same direction. We can represent that supposition in terms of the seven features of the Cij relation as follows:

(1) Presence or absence: attitude is related to behavior (Ci is related to Cj);

(2) Temporal order: attitude precedes behavior (Ci precedes Cj);
(3) Logical order: attitudes lead to behaviors (Ci leads to Cj);
(4) Direction of function: positive; favorable attitude will go with favorable behavior (Ci is positively related to Cj);
(5) Form of function: the more favorable the attitude, the more favorable the behavior (Cj is a monotonic function of Ci);
(6) Deterministic/stochastic character of function: favorable attitudes tend to lead to favorable behavior (Ci "tends to" lead to Cj); and
(7) Temporal stability: the supposition implies, although it does not state, that the attitude-behavior relation tends to be stable over time.

You need to specify the features of two more relations: social pressure-behavior (Ckj), and social pressure-attitude (Cki). Suppose you postulate that social pressure *is* related to behavior, such that changes in social pressure precede and tend to lead to changes in behavior in the same direction. We can represent that postulation, in terms of the features of the relation, as follows:

(1) Ck is related to Cj.
(2) Ck precedes Cj.
(3) Ck leads to Cj.
(4) Ck and Cj are positively related.
(5) Ck and Cj have a monotonic relation.
(6) Ck has a stochastic relation with Cj.
(7) By implication, the Ckj relation is stable over time.

Suppose you further postulate that social pressure and attitude (Cki) may or may not be related; that they may occur in any temporal order; that their logical order is unspecifiable; and that the direction, form, stochasticity, and stability of the relation are unspecifiable. This postulation for the conceptual network does not differentiate among any values of any of the features of the ki relation; hence the *conceptual domain* contains no potential information about any of the seven features of the ki relation. (That is, if we start with such a nonspecification of Cki in the conceptual domain, we can gain no information from the conceptual domain about Cki, although we could potentially gain information from the other two domains.)

This completes the tasks (a and b) of step 1 for a concept-driven pathway. To follow a concept-driven design pathway, step 2 involves choosing elements and relations from the methodological domain; specifically, a method or mode of treatment for each of three elements (Mi, Mj, and Mk) and a comparison technique for assessing each of three relations (Mij, Mkj, and Mki). In many areas of the social and behavioral sciences, the researcher makes those choices from a relatively well-stocked library of possibilities. Theoretically, the researcher makes those choices so as to provide the best possible fit for the concepts already selected. Actually, those choices are constrained by the researcher's experience, training, and preferences.

Suppose that two sets of semantic differential scales are chosen to measure attitudes and social pressure (scales Mi and Mk); and that some kind of formal records (such as sales records) are chosen as the source of measures of the behavior (Mj). Suppose, too, that both linear correlation and multiple regression are chosen as comparison techniques for assessing the set of three relations (Mij, Mkj, and Mki). The behavior records, Mj, would undoubtedly be assigned as the criterion variable in the multiple regression, and attitudes and social pressure scales, Mi and Mk, as the predictor variables. You now need to specify the features of the relations in the methodological domain. We will focus on correlation as the comparison technique.

(1) Presence: Correlation will detect presence and degrees of *linear* relation for all three relations.
(2) Temporal order: Correlation will *not* distinguish one temporal order from another for any of the relations unless measures are taken on more than one occasion.
(3) Logical order: Correlation will *not* distinguish one logical order from another for any of the relations.
(4) Direction of function: Correlation can distinguish positive and negative direction of linear relation for all three relations.

(5) Form of function: Correlation will impose linearity on all three relations, will test the degree of linearity, but will not assess the presence or degree of any other form of function.
(6) Stochasticity: Correlation can be used to distinguish deterministic from nondeterministic relations.
(7) Temporal stability: Correlation cannot be used to distinguish among some patterns of stability and change over time, unless measures are taken on more than one occasion.

The next task (d) for step 2 is to match this set of features to the features of those relations called for in the conceptual network. For the ij relation, the conceptual network specifies presence of a relation, and the methodological network can detect presence or absence of that relation. Feature 1, therefore, has potential information gain. But the conceptual network specifies both a temporal and a logical order, whereas the comparison technique cannot detect either. Both features 2 and 3, therefore, contain no potential information. The second domain (methodological) *restricts* what you can learn about those features of that relation, even though those features are specified in the conceptual domain.

On direction of function, the conceptual network specifies a positive direction, and the comparison technique can detect which direction occurs. On form of function, however, the conceptual network specifies a monotonic form (of which a linear relation is a special case), whereas the methodological tool (correlation) *imposes* a linear relation, and cannot distinguish among other functional forms. Hence, the study outcome is restricted in the amount of potential information it can yield on this feature.

There is a match on stochasticity. The conceptual network specifies nondeterministic, and correlation can deal with either deterministic or nondeterministic relations. Finally, the conceptual network implicitly specifies that the Cij relation has temporal stability, whereas correlation cannot distinguish patterns of stability or change unless measures are taken for multiple occasions. So this feature yields no information. Overall, then, there is maximum potential information gain on three of the features, some potential information gain on one of them, and no information gain on three features for the ij relation. Note that the relation of social pressure to behavior (kj), has the same profile of features as does ij, in each domain. Therefore ij and kj have the same pattern of matches and mismatches and the same pattern of potential information yield.

The relation ki, however, is quite a different matter. In the example, none of the features of the social pressure-attitude relation, Cki, were specified in the conceptual network. But the comparison technique from the methodological domain, correlation, will generate indices indicating presence, direction, degree of linearity, and stochasticity of the Cki relation. Those four features will yield restricted potential information. The lack of differentiation in the conceptual network restricts what can be learned about those features of ki. The remaining three features—temporal and logical order, and temporal stability—yield no information; the conceptual network does not specify these features and correlation cannot distinguish among values on those features. The study will yield little potential information on the ki relation.

In step 3 for the pathway involving a concept-driven design, the researcher needs to select material from the substantive domain and identify the features of the relations in that material. Suppose the researcher decided to examine the relation between people's views about buying cars and actual car purchases, gathered from people in Baldwin, Michigan in the coming year. By the end of step 2, the researcher already knows that the study can only gain full information about three features of ij and kj, restricted information about one of the features, and no information about three features. So he or she is likely to select material from the substantive domain to deal mainly with those features that can still yield much potential information by the end of step 2. That would imply selecting material from the substantive domain so that it could display presence or absence, and positive or negative direction, of a linear relation. But the substantive material need not be able to display any other functional form because correlation

would not distinguish them in any case. Furthermore, because correlation does not permit distinguishing either temporal or logical order or temporal stability, the substantive material can be selected without concern for the timing or order of observation of events (e.g., dealing only with substantive events in a single slice of time).

However, correlation could distinguish temporal stability and change, and temporal order, if measures were taken on Si, Sj, and Sk at more than one time. You might want to select substantive material so that Si, Sj, and Sk each can be observed on more than one occasion, and use Mi, Mj, and Mk in that fashion, thereby permitting exploration of the temporal order and temporal stability features. In this light, the researcher might choose, for the substantive material, to gather the views of people in the village of Baldwin about buying cars and about their perceptions of other's ideas about buying cars during April and again during August, and to gather automobile purchase records for all months from that April until the next. If that were the manner in which substantive material was selected for the study (and if the methodological tools were used in a multiple-occasion fashion), then the relation between i and j in the substantive network can be described on the seven features as follows:

(1) Presence: the substantive sample can display either presence or absence of the attitude-behavior (Sij) relation;
(2) Temporal order: the substantive sample was selected so as to permit display of temporal order in the attitude-behavior (Sij) relation;
(3) Logical order: because the methodological tools used in the study already imposed constraints that make it impossible to assess logical order, the substantive material was not selected so as to permit display of various logical orders;
(4) Direction of function: The substantive material can display either a positive or negative direction for the attitude-behavior (Sij) relation;
(5) Form of function: the substantive material was sampled so as to permit display of any functional form for s_{ij} but the prior choice in the methodological domain (m_{ij}) imposes a linear functional form on the results;

(6) Stochasticity: The substantive network can display either a deterministic or stochastic relation for the Sij relation.
(7) Temporal stability: The substantive domain can display either stability over time or change over time, if measures are taken on more than one occasion.

The outcome of step 3 is a body of evidence. The study is constrained, however, in what features of the relations it can gain information about. For two of the features, the logical order and functional form of the relation, the study is constrained in what it can yield information about because the comparison technique cannot distinguish among the various logical orders, and it cannot distinguish any functional form except linear.

A study can gain information concerning only those features that match across the three domains. For the features of the relations selected for this example, the study could gain information about the presence or absence of the ij and kj relations, their direction, their stochasticity, and possibly their temporal order and temporal stability (depending on whether measures were taken on multiple occasions). Such a study could not gain information about the logical order of those relations, and could learn only about whether the functional form is linear.

For the pathway involving the concept-driven design, mismatches that are due to features of the comparison technique (the second domain), are likely to be coped with by selecting different comparison techniques, in order to gain information about the features of the relations in the domain of primary interest (that is, the conceptual domain). In this example, the features of the relations selected from the third domain, the substantive domain, did not further constrain the features of the network of relations beyond what had already been constrained by the methodological domain. In other instances, however, the third domain can further constrain the information that can be gained from a study.

Some Comments on Mismatches

Constraints from the three domains differ because the domains serve different functions. Constraints from the conceptual domain arise mainly from failure to specify which values of some features of relations among the concepts being studied are expected to occur. (Such constraints due to nonspecification were the case for the ki relation in our example). Constraints from the methodological domain arise mainly because some comparison techniques impose one particular level of some features (as for the linearity of the correlation technique in our example) or cannot differentiate among levels of some features (as for logical order in the correlation example). Constraints from the substantive domain arise mainly from the way in which the researcher selects, or samples, portions of the "real world" that are to be included in the study. When the researcher selects intact systems for study, that selection is liable to bring with it many constraints regarding which levels of various features of the relations can be displayed. For example, ongoing systems tend to obscure causal directionality in sets of relations that are interconnected with feedback loops. On the other hand, when researchers sample only selected components of intact systems, or when they concoct systems for study, that selection is liable to bring with it just as many constraints—though different ones—regarding which levels of various features of the relations can be displayed. For example, relations between elements that are studied out of context tend to fix the values of features of those contextual relations (that is, extrinsic features of the ij relation under study), thus denying the researcher the chance to learn about them.

SOME FINAL COMMENTS

In our previous presentations of the VNS (e.g., Brinberg and McGrath, 1982; McGrath and Brinberg, 1983, 1984), we named 12 subtypes of validities within stage two—one for step 2 and step 3, of each path, at the element and relation level. They included some familiar terms (e.g., construct validity, predictive validity), but also included a number of terms that we invented for the schema (e.g., pattern validity, process validity). We now regard those designations of types of validity to be incorrect and misleading for two reasons.

First, earlier we had conceived of them as different *types* of validities. We now think they can more usefully be construed as different *locations* within stage two of the research process at which validity issues arise. In one sense, they are all the same type of validity issue; each is a cross-domain matching of features of relations among elements. In another sense, these locations pose different validity issues. Different things are being "fit" at different points in the process; hence the constraints differ and so do the consequences.

Second, in the past we had separated validity issues associated with the element and relation levels. We now consider the validity issues—the questions of fit—to be associated with various features of the relations among elements.

Because Campbell and colleagues have done the hallmark work that dominates current thinking about the validity of research studies, it is worth noting how the VNS relates to that work. Our analysis of validity in stage two, as correspondence, in large part parallels (or at least seems to be in agreement with) Cook and Campbell's (1979) treatment of internal validity of designs. Their notion of internal validity has to do with logical rather than empirical correspondence. Their category of *statistical conclusion validity*, on the other hand, subsumes some of the issues regarding empirical correspondence of relations. Furthermore, their category of *external validity* subsumes some of the issues we treat within stage three, as matters of robustness, although our concept of stage three validity seems to be considerably broader than their external validity notion (see Chapter 5).

It is their category of *construct validity* that we seem to have sliced up in our treatment. Some aspects of construct validity are matters of definition—of fiat—as when the researcher maps elements across different domains in the first task of step 2 of stage two. On the other hand, much of construct validity involves parts of the matching process. Step 2 of the pathways that involve fitting the conceptual to the methodological domain, or fitting the conceptual to the substantive domain, and step 3 of the pathways that involve fitting the conceptual domain to a prior structure (a set of observations) are matching operations that are in many ways efforts toward what has been called construct validity. (The term construct validity has been used both for correspondence at the element level and at the relations level. The latter has also been referred to as nomological validity.)

Our treatment of stage two validity issues embeds one very crucial distinction that is seldom made in the methodological literature of the social and behavioral sciences. In delineating validity issues, we have emphasized the *role of logical possibilities, rather than the role of empirical outcomes.* We have made correspondence validities rest on whether a method can distinguish among—and a substantive system can display various values of—features of the relations under study, and whether the conceptual network has specified or postulated a certain value of that feature. We have not made validity rest on whether or not a particular sample of substantive system actually does display, and the methodological network actually identify, the particular values of those features that are specified in the conceptual domain. In other words, we have made validity rest on whether we can find out in our study that the values specified in the conceptual domain either are or are not correct, rather than upon whether those conceptual predictions turn out, empirically, to be correct. In so doing, we have excluded from validity as correspondence such concepts as "predictive validity"—the degree to which a set of postulated relations actually hold in some sample of cases. Such empirical outcomes are one underlying idea about the meaning of validity in stage three: validity as robustness.

NOTE

1. We recognize, of course, that these ten features of relations are not the only way to present the matter. It is possible, for example, to combine feature 1 with features 4 and 5;

that is, presence with direction and form of function. Moreover, it is possible to specify other values of some of the features. For example, we could add functional forms such as step function, quadradic, and so on, for feature 5. Furthermore, we make no claim that these ten features exhaust the set of interesting and important features of relations. We merely present these as one set of features that are very general in their applicability and very powerful in their implications for information gained from research. We will use these ten features as the basis for discussion of the matching process within stage two research.

GLOSSARY OF TERMS

Correspondence or Fit: The degree to which there is a match between the values (of features of relations) that contain potential information in one domain, and the values (of those features of those relations) that contain potential information in another domain(s). The greater the correspondence, the greater the amount of potential information in the resulting findings. The functions of the domains differ, and therefore the impact of (lack of) correspondence differs across domains.

Features of Relations: Associated with every relation are a set of features (characteristics) that can be used to describe that relation. We have presented seven such features intrinsic to any relation and three features that we view as extrinsic to any single relation (although intrinsic to the network of relations being examined). Matches across domains among these features (or some similar set) reflect the potential information that may be gained from the study.

Matching Process: A set of tasks within the steps of stage two by which a researcher can assess the degree of correspondence, or fit, on values of features of relations across domains, and therefore determine the amount of potential information in a study.

Potential Information Gain: The outcome of stage two is a set of empirical findings. The amount of information that potentially can be gained from these findings, however, depends on the degree of cross-domain correspondence or fit of the features of the set of relations under study. There is maximum potential information gain for a given relation when there is a conceptual specification for all

values of each feature of that relation; when the comparison techniques can distinguish among all possible values for each feature of that relation; and when the observations of the phenomena are sampled so that they can display all possible values of each feature of that relation.

Chapter 5

STAGE THREE
Validity as Robustness

The outcome of stage two activities is a set of empirical findings. These findings will only gain meaning and importance, however, when they reduce our uncertainty about the focal problem. In this chapter, we give a description of the activities of stage three of the research process, viewed as systematic efforts to build research information and establish confidence in it. Considering the relatively strong emphasis that psychology and other behavioral sciences have given to stage two activities, those fields have given relatively little attention to understanding the processes by which research information is transformed into knowledge; that is, how we gain confidence that we "know" what the findings we have gathered mean.

This chapter has five sections. The first describes the meaning of validity within stage three of the research process. The next three describe the quest for "external validity" within each of the three domains: substantive, methodological, and conceptual. The final section contains a discussion of some of the consequences of stage three activities for reducing uncertainty about research findings.

THREE ASPECTS OF VALIDITY AS ROBUSTNESS: REPLICATION, CONVERGENCE, AND DIFFERENTIATION

The outcome of a single study, by itself, contributes little to our body of knowledge. Only when the results of a single study have been compared with other studies that examine the same focal problem do we increase our knowledge about that problem. Brinberg and McGrath (1982) present three issues that need to be considered when assessing the generalizability of a finding: (1) If the study were repeated, would the same finding occur (i.e., would the findings replicate)? (2) If the study was repeated, but certain facets were allowed to vary (e.g., type of respondent), would the finding triangulate (i.e., would it show convergence) across those differences? (3) Under what conditions will the finding not hold (i.e., what are the boundaries, or limits, associated with that finding)?

The first issue, replication, is a special case of the second, convergence, for which no factors are intentionally varied. The second and third issues are related. They are the two possible outcomes from a robustness analysis. That is, if the findings do not converge, the researcher has identified a boundary associated with those findings. Boundary search and convergence analysis always are done simultaneously. They are different perspectives on the same question, and they reflect opposite sets of expectations by the investigator. All *unequivocal* outcomes of stage three activities represent evidence about one or the other. We have chosen to separate issue 2 (convergence analysis) and issue 3 (boundary search) to highlight the need for researchers to examine both the scope *and* limits of their findings.

Both the effort to replicate a finding and attempts to show its convergence over selected differences are attempts to reduce uncertainty by failing to disconfirm some hypothesis. When replicating a study, as well as when conducting a convergence analysis, the anticipated outcome is a failure to disconfirm the hypothesis that the findings of the original study and of the replicate are the same (i.e., to find no differences). When a researcher does not find differences, the uncertainty associated with the prior finding is reduced.

When a replication is attempted and differences do occur between the original and the subsequent study, however, researchers usually assume that there is now increased uncertainty associated with the prior finding. We argue that when a convergence analysis finds differences (and when the researcher can be confident in that finding), whatever facet was changed represents one boundary of the prior finding. Finding such differences *also contributes to a reduction of uncertainty because it identifies conditions under which one can (in the future) confidently state when that (stage 2) finding will not hold.*

To focus on only one of these issues limits what the researcher is able to learn about the focal problem. Research in the behavioral and social sciences has placed a heavy emphasis on the replicability of research findings, and some emphasis on exploring the scope of those findings. It has addressed in only very limited form the issue of assessing the boundaries of those findings.

Replication in the VNS

One important activity for the stage three researcher is to determine if the set of findings can be reproduced when the same pathway and the same set of elements, relations, and embedding system from each of the domains are used again. In principle, an exact replication is not possible; two occasions are always different (Runkel and McGrath, 1972). One standard technique for assessing replicability of a measure, for instance, is a test-retest procedure, in which the retest is treated as a replicate. But the retest differs from the original test with respect to time and with respect to amount of prior exposure to the material. The researcher using this procedure needs to act *as if* the test and the retest are identical sets of conditions, so that results can be used to assess replicability of the instrument. Within the VNS, the researcher needs to treat the elements, relations, and embedding systems selected from each of the domains *as if* they are replicates, in order to assess whether a finding is reproducible. At the element level, researchers typically refer to the validity issues involved in such replication as the question of the *reliability* of a measure. At the relations level, the validity issues involved are part of what Cook and Campbell (1979) call *statistical conclusion validity*. We refer the reader to several excellent sources (e.g., Stanley, 1967; Nunnally, 1967; Cook and Campbell, 1979) for a detailed discussion of the factors influencing the reliability of a measure and the reliability of a relation.

Triangulation or Convergence in the VNS

In addition to assessing replicability, though, the researcher also needs to determine under what broader range of conditions the finding will and will not hold; that is, the scope and the boundaries of that finding. The principle of robustness is an integral component of the work of many highly regarded treatments of research methodology (e.g., Cronbach, 1972, 1982; Campbell, 1981; Popper, 1959). Within the Popperian view of knowledge acquisition, for example, researchers acquire knowledge either by failing to disconfirm some expectation or hypothesis or by confidently disconfirming such a hypothesis.

At its core, robustness includes both the principle of convergence and the principle of differentiation. **Convergence or triangulation** holds that we gain confidence in a research finding (i.e., we eliminate threats to the validity of that finding and reduce our uncertainty vis a vis that finding) only when we have agreement of substantive outcomes derived from the use of different and independent models, methods, and occasions. Campbell and Fiske (1959) introduced the principle of triangulation in their presentation of the convergent and discriminant validity of measures. Wimsatt (1981: 126) also uses the triangulation or convergence principle when describing a robustness analysis of research findings:

> 1. To analyze a *variety* of *independent* derivation, identification or measurement processes; 2. to look for and analyze things which are *invariant* over or *identical* in the conclusions or results of these processes; 3. to determine the *scope* of the processes across which they (the findings) are invariant and the *conditions* on which this invariance depends; and 4. to analyze and explain any relevant failures of invariance.

In the VNS, the sets of elements and relations from the conceptual, methodological, and substantive domains each are a potential source of ambiguity when interpreting an empirical finding. The stage three researcher tries to reduce the uncertainty associated with each potential source of ambiguity by addressing the three issues raised earlier in this chapter (replicability, convergence, and boundary search) with respect to each of the three domains.

In the conceptual domain, for example, Feyerabend (1970) draws upon the idea of convergence in discussing the need for researchers to develop competing models. He argues that theoretical pluralism is assumed to be an essential feature of all knowledge that claims to be objective. The work of several scholars that focused on the critical experiment is consistent with our principles of convergence. For example, Garner et al. (1956) apply the principle of triangulation or convergence within the conceptual domain when they argue that a discipline is likely to advance only when it compares alternative theories. Furthermore, Platt (1964), in his influential article on "strong inference," presents the notion that a researcher should compare competing theoretical explanations of a focal problem to determine which is the better explanation of that problem.

The idea of triangulation or convergence is most often associated with the work of Campbell within the methodological domain (Brewer and Collins, 1981). Campbell (1981) and colleagues have argued for triangulation across measures and regard such triangulation as the method of choice for reducing certain sources of potential invalidity in the measurement and manipulation of variables. Both Fiske (1982) and Campbell and O'Connell (1982) offer recent reviews of some work in that tradition. Although that work is essentially concerned with the validity of measures (element level within the methodological domain), the triangulation approach also has been advocated with respect to validity at the relations and embedding systems levels in the methodological domain. McGrath et al. (1982), for example, argue that successive studies of the same problem should make use of maximally different strategies (e.g., laboratory experiments, field studies, sample surveys) and different research designs. Diversity of approaches will increase our confidence in a set of findings if the findings converge. The researcher can offset the weaknesses of any one strategy or design by using other strategies or designs that also are flawed, though differently.

The triangulation idea also applies in the substantive domain. The behavioral and social science community has long recognized that any interpretation of findings must be considered in terms of the limits of the samples (of actors behaving in context, of patterns of phenomena, of higher order systems) on which any study is based. For example, some critics of behavioral science have long been concerned that psychology relies too much on evidence gained from research on college students under artificial conditions. Such a concern implies a call for triangulation of research findings across different facets of the substantive domain (as well as across facets of the methodological domain). But the need for such exploration of the scope and limits of a finding, with respect to facets of the substantive domain, is too often honored in the breach rather than in the observance. We routinely insert a standard caveat near the end of our reports of empirical investigations, such as, "Future research is needed using other samples and stimulus materials, to determine the robustness . . . " But that caveat is often juxtaposed with conclusions that violate it.

Our view of convergence is similar to Cronbach's (1972, 1982) discussion of generalizability theory. In his view, the researcher is interested in determining the extent to which a finding will generalize to some underlying population (in his terms, populations of units, treatments, and observations). The facets of

interest are determined by the researcher. In the VNS, we argue that not only are there such general sets of populations that a researcher needs to consider within the substantive domain, but also within the conceptual and methodological domains as well. Within each of these domains, there is a wide variety of elements and relations that, if examined, can provide useful information concerning the interpretation of findings.

Differentiation or Boundary Search in the VNS

Triangulation or convergence has a complement, the idea of **differentiation or boundary search**. Differentiation is intrinsic to the research activities of stage three, although it is often discussed in negative terms. The dual concepts of convergence and differentiation—of the search for the scope and limits of a finding—are expressed by Wimsatt (1981) as invariance and failures of invariance. The concept of "*failures of invariance*" (Wimsatt, 1981) makes the very important point that the researcher is not only looking for the conditions under which the findings will fit the hypothesis, but also should be trying to identify and explain the conditions under which the findings are inconsistent with (i.e., disconfirm) the hypothesis. Lynch (1982) has talked about such nonconfirmatory outcomes, failures of invariance, as the lack of construct validity in a nomological validity sense. Our position is that both "invariance" and "failures of invariance," both the scope and limits of the findings, or in Lynch's terms, both "construct validity" and "lack of construct validity" can yield useful information.

In the conceptual domain, for example, the researcher is interested in determining whether the concepts (i.e., elements) selected provide a unique explanation for the empirical findings or whether alternative concepts are equally effective in explaining those findings. Suppose, for example, that a researcher selects the attitude concept to explain certain choice behavior. If a different concept (e.g., situational pressure) is equally effective in explaining that choice behavior, then the finding is not differentiated across the two alternative constructs; that is, neither of them offers a *unique* interpretation of the empirical finding.

In the substantive domain, we also attempt to identify the boundaries associated with an empirical finding. For instance, a researcher will often attempt to determine whether a finding will hold for different types of respondents (e.g., college students versus housewives). If the findings are different across types of respondents, then the researcher has identified one boundary (limit) of that finding. By identifying such boundaries, the researcher is better able to predict the conditions under which the empirical findings will and will not hold.

We also need to assess the limits (boundaries) of a finding with respect to the methodological domain. For instance, we could use two distinct methods (e.g., a self-report and a trace measure) to measure one variable in a particular relation. If the findings are different for the two measures, we have identified one boundary of that finding, one set of conditions under which the finding does not hold.

The dual concepts of triangulation or convergence, and differentiation or boundary search, are directly connected to Popper's (1959) falsification principles. Triangulation or convergence is the acquisition of knowledge by failing to disconfirm the hypothesis that two things (e.g., a hypothesis tested with each of two methods) are the same. Differentiation is the acquisition of knowledge by confidently disconfirming the hypothesis that two things are the same. To focus on only one of these strategies is to restrict our ability to reduce uncertainty in any set of empirical findings. It limits our ability to know what we know.

Typically, researchers focus their stage three efforts on attempts to replicate research findings. When they go beyond that, it is usually to carry out triangulation with respect to one facet of one domain, most often the population (actors) facet of the substantive domain. Researchers seldom deliberately search for the boundaries associated with their findings. When outcomes of a stage three study lead to nonconfirmation of the convergence of a set of findings, these are usually regarded, pejoratively, as negative results (that is, as "failures of invariance" in Wimsatt's terms, or "lack of construct validity" in Lynch's terms). *We contend that these so-called negative results can provide researchers with information just as useful as would positive results because those negative results also can reduce the uncertainty associated with a research finding.*

For both a convergence analysis and a boundary search, however, uncertainty is reduced only if the researcher is able to account for (i.e., to make sense of) the results. For a convergence analysis, uncertainty is reduced when the researcher is able to explain (in terms of concepts that are meaningful within the current paradigm) why two or more findings are similar. For a boundary search, uncertainty is reduced when the researcher is able to explain (in terms of such meaningful concepts) why two or more findings are different.

Much of the discussion concerning stage three activities in the social and behavioral sciences has been phrased in terms of the external validity of empirical findings (e.g., Berkowitz and Donnerstein, 1982; Cook and Campbell, 1979; Cronbach, 1982; Lynch, 1982, Mook, 1983). Within that discussion, there has been considerable debate about whether external validity is *necessary* for various kinds of experimental research (e.g., the dialogue between Calder et al., 1981, 1982, 1983 and Lynch, 1982, 1983). In our view, the primary concern with external validity is *not* just a matter of generalizing to a population or infusing "mundane realism" into research. Rather, it is a concern for the robustness of an empirical finding, which in turn is a concern for reducing uncertainty vis a vis that finding. The robustness of a finding needs to be assessed with regard to each of the three domains. We will argue that all types of research (i.e., basic, applied, technological) need to be concerned with the robustness (external validity) of a set of findings generated from that research approach. Only by exploring the scope and limits of a set of findings is the researcher likely to reduce uncertainty about it.

As noted earlier, there are three domains with respect to which a researcher needs to reduce uncertainty. The replication issue is not specialized by domain, nor is it particular to any facet of any domain. When a replication is conducted, a researcher tries to reproduce the conditions of a previous study *in all particulars*. In fact, replication may be considered a special case of convergence analysis (although most researchers treat replication and convergence as separate issues). In the following sections, we will discuss the research activities of stage three, with regard to both convergence and boundary search, for each of the three domains.

ROBUSTNESS ANALYSIS IN THE SUBSTANTIVE DOMAIN

Convergence Analysis

We refer to the research activities of stage three, in the substantive domain, as the search for **ecological validity**. Convergence analysis in the substantive domain is what most people mean when they talk about external validity. Futhermore, much of the literature on external validity of research findings has focused on convergence on just one facet of the substantive domain; namely, the type of respondent used in the research studies (e.g., college students versus housewives versus blue-collar workers). Generally, researchers have explored the external validity of their findings by assessing the extent to which a particular research finding "generalizes to" (Cook and Campbell, 1979) some designated population. Sometimes this debate has seemed to hinge on the question of which type of sample is the "proper" kind on which research findings ought to be built. What seems more cogent, to us, is not which type of respondents is used in any one study, but rather that samples of different types of respondents are used in different studies of the same problem, so that we can examine the robustness of research findings across variations in that facet. When findings from studies using different types of subjects converge, the uncertainty associated with those findings is reduced.

Convergence also needs to be assessed for variations across types of behaviors. For example, findings bearing on individual or group task performance effectiveness need to be explored across a wide range of tasks. The uncertainty associated with any set of findings can be reduced to the extent that similar findings occur across studies involving different classes of behaviors.

Similarly, the convergence of a set of findings needs to be assessed across a wide range of context within which the actors are behaving. Context, here, includes spatial, temporal, and situational aspects of the environment within which the behavior occurs. Much research on person-by-situation interactions (e.g., Endler and Magnussen, 1976) suggests that the context within which a behavior occurs modifies its meaning. We can further reduce uncertainty about our findings by assessing their convergence across different contexts.

To summarize, research exploring the convergence of findings across different facets of the substantive domain has focused mainly on the samples of respondents in the study, and to a lesser extent on the range of behaviors to which the finding is expected to apply. There has been little effort to assess convergence of findings with respect to facets of the physical, temporal, and situational context.

Boundary Search

As noted earlier, boundary search is a complement to convergence analysis as a means to reduce the uncertainty associated with a set of findings. Researchers can deliberately seek the limits of a finding (that is, to disconfirm a finding with confidence) by testing it under conditions for which it is expected *not* to hold. Runkel and McGrath (1972) assert that knowledge is knowledge of differences. For a finding to be useful, a researcher needs to identify not only its scope but also its limits. When a finding has no limits, it also provides no useful information because no distinctions are made. Lynch (1982) discusses boundary searches as attempts to identify background factors that interact with a finding. By incorporating these background factors into a more complex statement of the empirical finding, the researcher reduces the uncertainty associated with that finding by specifying both its scope and its boundaries (i.e., those conditions under which the particular finding should or should not be expected to hold).

The search for boundaries in the substantive domain is helped by systematically considering the facets of that domain. First, there is the question of the types of respondents for which a particular finding does not hold. For instance, Jaccard et al. (1979) used a subjective probability model to predict voting behavior and applied it to samples of both highly educated and poorly educated respondents. The model predicted behavior for both groups, but did so more effectively for the highly educated respondents than it did for the poorly educated group. Those authors, therefore, identified one limiting condition, or boundary, of the findings of their model. Yet to use that information to reduce uncertainty, these researchers need to consider how the influence of education level and related attributes can be incorporated into the model's predictions.

Similarly, a researcher can try to identify the range of behaviors over which a research finding does or does not hold. Some consumer behavior researchers, for instance, (e.g., Engel and Blackwell, 1982) argue that different decision processes occur for low-involvement and high-involvement behaviors. If so, then these theorists need to specify why various types of decisions can and cannot be predicted accurately from a given model.

The boundaries of a finding also can be examined with respect to context. Cronbach (1975) discussed what we would call boundary search on the context facet in terms of the influence of higher-order interactions. For example, he analyzed a situation in which a certain dose of a sedative was effective for 30 minutes in one context (i.e., when a certain type of wood chip had been used for the animal's bedding) and for only 15 minutes in another context (i.e., when a different set of wood chips had been used). Identifying such an interaction amounts to specifying a context boundary. To do so can reduce the uncertainty associated with

the findings about effects of drug dosage—but only if those findings can be incorporated into a theory of the processes underlying the focal problem.

In sum, we can reduce uncertainty about our findings by identifying the boundaries associated with them across various facets of the substantive domain. Simply determining the convergence or scope of a finding is not sufficient; the limits of the finding also must be identified and interpreted.

ROBUSTNESS ANALYSIS IN THE METHODOLOGICAL DOMAIN

In the preceding discussion of validity in the substantive domain, we assumed that both the methods and the concepts were held constant, and that only selected elements and relations from the substantive domain were allowed to vary. But in addition to determining the convergence and boundaries of a finding with respect to various facets of the substantive domain, researchers also need to assess the convergence and boundaries of that finding with respect to various facets of the methodological domain. We refer to stage three research activities in the methodological domain as the search for **methodological validity**.

Convergence Analysis

In the methodological domain, we are used to thinking about convergence analysis because of Campbell's (see Campbell and Fiske, 1959) influential work on convergent validity. When examining this issue, researchers typically select different measurement techniques to measure the same variable(s) and use some variant of the multitrait-multimethod paradigm to determine whether the methods converge. Some researchers have even extended this convergence logic to the use of multiple indicators for the independent variables (Runkel and McGrath, 1972).

We can assess convergence in the methodological domain by examining a set of findings that vary, systematically, in obtrusiveness or reactivity of the measures. For example, we can contrast a set of findings derived from use of a semantic differential with a set of findings based on trace measures. If the findings converge across types of measures, we can be more confident in them.

We also can examine the convergence of a finding across comparison techniques. For instance, Birnbaum (1982) demonstrated how between-subject versus within-subject factorial designs can result in different findings because of differences in the perception of the stimuli in those designs.

We also can consider convergence of findings across various research strategies (see Chapter 2). If the evidence on a focal problem comes from only one strategy (e.g., all from laboratory experiments), there is uncertainty concerning the meaning of those findings because they are confounded with properties of that strategy. This uncertainty can be reduced if researchers use different strategies in studies of the same focal problem and their findings converge across those studies. There has been much debate as to the relative merits of laboratory versus field studies. As with the debate over what types of subjects would be "best" for research, much of the debate about strategies has seemed to be an argument over which strategy is the "right" one for research (compare Berkowitz and Donnerstein, 1982; Mook, 1983). We have argued, instead, (see Chapter 2) that every strategy is flawed, and that only through the use of multiple strategies can we offset those flaws and thereby reduce the uncertainty associated with those findings.

Boundary Search

The other side of the convergence question, here as in the substantive domain, is the question of limits of variations in methods, designs, and strategies beyond which a finding will *not* hold.

We can gain information about a finding (i.e., reduce uncertainty) by identifying classes of measures over which that finding does *not* hold. For instance, suppose a finding holds for a reactive self-report measure but not for a nonreactive measurement technique such as a trace measure. One possible reaction is to point out the limitations of the particular method for which the finding did not hold. An alternative approach, and the one we support, would be to explore what basic features differentiate these two measurement instruments (e.g., the apparent social desirability of the response) and to incorporate that information as part of the "meaning" of the (now elaborated) research finding. If a researcher is confident about the effect of the measurement instrument's properties themselves on the findings, then those instrument-related differences in the findings may provide insights concerning the limits or boundaries of that finding. To treat differences in findings merely as artifacts of the different measures is to ignore potentially useful (i.e., uncertainty reducing) information.

The same point of view concerning boundary search also applies to study designs and research strategies. If the findings from two different comparison techniques (e.g., a between- versus a within-subject design) are *predictably* different, the researcher can incorporate the basis of that prediction into a more meaningful interpretation of the finding. For instance, Grice and Hunter (1964) showed that quite different results were obtained in simple reaction time data for eyelid conditioning when within-subject and between-subject designs were used. Their results specify one of the limits of the finding, thereby reducing the uncertainty associated with that finding insofar as they can incorporate that finding into their guiding theory. Similarly, if a finding from the lab does not replicate in a field setting (see for example Hovland et al., 1949) this is potentially useful information about the limits (boundaries) of that finding, but only if it can be incorporated into the interpretation of that finding. Furthermore, such "failures to replicate" can lead to revision of the theory guiding such replication attempts.

ROBUSTNESS ANALYSIS IN THE CONCEPTUAL DOMAIN

Stage three researchers need to know whether elements and relations from the conceptual domain *uniquely* and *adequately* account for the set of findings under study. One aspect of uncertainty that is associated with a finding is whether alternative concepts and relations could account for the empirical findings equally as well as, or more effectively than, the concepts being studied. Stage three researchers can address these questions by examining the findings with respect to different sets of concepts and relations from the conceptual domain. Doing so will help indicate the extent to which the concepts and relations that were originally selected for use do accurately and uniquely account for the set of findings, compared to alternative concepts and relations. Such an approach is consistent with Feyerabend's (1970) discussion of theoretical pluralism; that is, the need to assess the meaningfulness of a model (i.e., a set of concepts and their relations) by testing it against alternative models.

When examining alternative elements and relations from the conceptual domain, the researcher often wishes for the outcome to be a confident disconfirmation of the hypothesis that the original model is no more accurate than the alternatives (that is, to find the original model to be more effective than alternative models). In contrast, in the other two domains, the researcher usually hopes for a failure to disconfirm (that is, to have the findings hold over the range of facets of substance and methods being examined). The roles that convergence and differentiation play are reversed in the conceptual domain. Therefore, we will deal with both of these issues in the opposite order. We refer to the research activities of stage three, in the conceptual domain, as **explanatory validity.**

Boundary Search

In the conceptual domain, a robustness analysis compares the original concepts and relations with an alternative set. If the

original elements and relations uniquely and adequately account for the findings when compared with an alternative set, that would increase confidence in the "validity" of the original concepts and relations, and would establish a limit (boundary) to the original set. For instance, much research on the relation between intention and behavior (e.g., Ajzen and Fishbein, 1980) has asked whether additional concepts are needed to explain behavior. For the most part, such research has established that intention is an effective predictor of behavior. Furthermore, the intention-behavior relation is a more parsimonious basis for explanation of behavior than models that involve additional predictor concepts (e.g., Brinberg, 1979).

The robustness of the functional relations among a set of concepts may be assessed by contrasting the original relations with some specific set of alternatives. For instance, Campbell (1963) presented a model using a threshold function to relate attitudes and behavior as an alternative to the linear relation typically tested. Such an analysis reduces the uncertainty concerning the form of the relation between attitude and behavior to the extent that it allows the researcher the opportunity to eliminate one set of potential alternative relations.

Convergence Analysis

The outcome of a robustness analysis may be treated as information concerning either the scope or the limits of a finding. As we noted earlier in this chapter, we have chosen to treat convergence analysis and boundary search separately in the other domains because of the way that lack of convergence (i.e., the limits of a finding) has been treated in the literature. In the conceptual domain, though, it is convergence that has been ignored and differentiation that has been given all the attention.

Finding out that some other set of concepts works as well as the selected set also can be valuable information, helping to reduce uncertainty. For example, Birnbaum and Mellers (1979) postulated (and found) that a single factor model of stimulus recognition could not be rejected as a descriptor of the relations between affect and exposure. A one-factor model appears to be an adequate, and more parsimonious, description of the empirical finding, compared to the two-factor model that dominates the literature in that area.

SOME IMPLICATIONS OF STAGE THREE ACTIVITIES

In stage three, the researcher's aim is to acquire knowledge by reducing the uncertainty associated with a set of empirical findings. Our view is similar to both Cronbach's (1972, 1982) discussion of generalizability theory and Lynch's (1982) discussion concerning the effects of background factors interacting with a set of findings. Both these scholars argue that researchers need to assess the robustness of a finding. Cronbach uses this principle of robustness in his discussion of the relation of a particular facet (e.g., a sample of units, treatments, observations) to its underlying population. Lynch, too, uses this principle of robustness when he argues that researchers need to examine additional (background) factors to determine their impact on a set of findings. If no interactions occur, Lynch argues, we can be more confident in the construct validity of those findings.

Although we have used the term construct validity in a manner different than Lynch, we agree with the spirit of his position. That is, incorporating information concerning the scope and limits of a finding enhances our understanding of those findings. Our view differs from Cronbach's generalizability theory and Lynch's position in that we use the VNS, and its domains, to specify a broader range of facets a researcher needs to consider in order to reduce the uncertainty associated with a finding. Furthermore, we specify three sets of stage three activities: replication, convergence analysis, and boundary search. These three sets of activities are each a necessary part of a discipline's efforts to reduce the uncertainty associated with its findings.

The social and behavioral sciences have acquired a wealth of empirical information, but much of that information is wrapped in uncertainty. Rather than spend valuable (and limited) resources generating new (and equally uncertain) information concerning the same focal problems, we urge researchers to spend more of their future time and effort pursuing robustness analyses as a means for reducing the uncertainty associated with the findings they already have.

GLOSSARY OF TERMS

Robustness: We use this term to describe the meaning of validity and the sets of research activities in stage three. Three sets of activities allow the researcher to assess the robustness of their findings.

Replication. If the researcher repeats the study and selects the same set of elements and relations from each of the three domains, would the same set of empirical findings occur? That is, are the findings reliable?

Convergence or triangulation. If the researcher repeats the study but allows (or causes) certain facets to vary systematically (e.g., type of respondent, type of behavior, type of measure) would the same set of empirical findings occur? For example, suppose a researcher was examining the relation between certain attitudes and behaviors and examined that relation using both college students and senior citizens (i.e., the researcher varied systematically the type of respondent in the study). If the same empirical findings occur (i.e., if results **converge** or **triangulate**), the researcher would be more confident in his or her understanding of the attitude-behavior relation under study.

Differentiation or Boundary Search. If the researcher repeats the study but allows certain facets to vary systematically (e.g., type of respondent, type of behavior, type of measure) would a *different* set of empirical findings occur? For example, suppose a researcher was examining the relation between certain attitudes and behaviors and examined that relation using both college students and senior citizens (i.e., the researcher varied systematically the type of respondent in the study). If *different* empirical findings occur (i.e., **differentiation** or **boundaries**), *and the researcher's conceptual schema can explain those differences*, that researcher would be more confident in his or her understanding of the attitude-behavior relation under study.

Stage Three Validity Issues. We describe the research activities in stage three, in the substantive domain, as the search for **ecological validity**. Ecological validity is the extent to which a researcher can specify the scope and limits of a set of empirical findings with respect to the elements and relations selected from the substantive domain.

We describe the research activities in stage three, in the methodological domain, as the search for **methodological validity**. Methodological validity is the extent to which a researcher can specify the scope and limits of a set of empirical findings with respect to the elements and relations selected from the methodological domain.

We describe the research activities in stage three, in the conceptual domain, as the search for **explanatory validity**. Explanatory validity is the extent to which a researcher can specify the scope and limits of a set of empirical findings with respect to the elements and relations selected from the conceptual domain.

Chapter 6

VNS AND THE RESEARCH PROCESS

In the previous five chapters, we have made many distinctions concerning the research process. We have described all research as drawing on three analytically separate, but logically related domains: conceptual, methodological, and substantive. Each of the domains has three levels: elements, relations, and an embedding system. The research process has three distinct stages: a prestudy or generative stage, a central or execution stage, and a follow-up or interpretative stage. We have given validity three distinct meanings, one associated with each of the three stages: as value, as correspondence, and as robustness. We have described three paths or styles for conducting the central stage of research: experimental, theoretical, and empirical. We have discussed three alternative research orientations: basic, applied, and technological.

The VNS presented in this book is our attempt to describe the relation of validity to the research process. We have made explicit the components of the research process and have linked the various forms and meanings of validity to those components. The most fundamental contribution of the VNS is that it integrates the many meanings of validity, and it does so by tying those validity concepts to the research process.

We think the VNS can aid our scholarly endeavors in a number of ways. In this chapter, we will discuss some possible uses, implications, and limitations of the VNS. First, the social and behavioral sciences need techniques to help integrate what we know in particular problem areas. There are a number of systematic approaches to such integrations (e.g., Rosenthal, 1978; Glass, 1976; Hunter et al., 1982; McGrath and Altman, 1966). These all involve ways to aggregate and quantify relations. There is disagreement about whether and how best to partition the body of evidence in the process of integrating it. We think the VNS is a very useful tool to do that systematically and elegantly. Hence, we see it as a tool for *qualitative meta-analysis*, so to speak. In the first section of this chapter, we discuss how the VNS—especially the stage three ideas discussed in the preceding chapter—can be used as a *guide to the partitioning* of bodies of literature and to the search for the scope and limits of sets of findings.

Second, the VNS has implications for how paradigms develop and change in the behavioral and social sciences. These implications are the topic of the second section of the chapter. Third, the VNS can help the researcher plan and guide their studies and analyze the research process. These potential uses of the VNS are the topic of the third section of this chapter, along with a discussion of some of the limitations of the schema.

VNS AS AN AID TO AGGREGATION OF RESEARCH INFORMATION

Recently, several researchers (e.g., Rosenthal, 1978, 1982; Glass et al., 1981; Hunter et al., 1982) presented quantitative techniques for summarizing a body of evidence. These techniques are referred to as meta-analysis. With them, the researcher can address three questions: (1) Is there a relation between two (or more) variables (e.g., between gender and influenceability)? (2) What is the strength of that relation? and (3) Under what conditions (e.g., subjects, behaviors, measures, and contexts) will the findings hold or fail to hold? The latter is the key question of stage three of the VNS.

In the VNS, the estimation of strength and presence of a relation is just the start of the researcher's stage three task. The search for scope and limits of a set of findings, however, can provide the researcher with considerable information about that focal problem. Our discussion of meta-analysis will focus on ways in which the VNS offers a strategy for addressing this third question.

The issue of determining what studies will be treated as testing the same hypothesized relation is fundamental for any meta-analysis. Meta-analysts have debated the conditions under which a set of studies should be treated as replicates for a single integration or partitioned into several subsets of studies to be integrated separately within each subset (compare Glass and Kliegl, 1983; Wilson and Rachman, 1983; Eysenck, 1978). From the point of view of the VNS, the question of partitioning is pivotal. In addition to assessing the replicability of findings, the meta-analyst also needs to determine to what range of conditions an empirical relation will generalize, and conversely, the boundaries associated with that relation. The issues facing the meta-analyst are the issues facing the stage three researcher; that is, determining the replicability of a finding, its scope, and its limits. A full exploration of those questions requires the partitioning of the body of literature across levels of the important facets in all three domains.

The set of facets we used in Chapter 2 to describe the contents of the three domains provides an overall framework for partitioning any given set of empirical findings. The set of facets also may be used by the researcher to identify areas of research in which there has been insufficient work; that is, to identify areas needing further research.

Partitioning on Facets in the Substantive Domain

The three primary levels within the substantive domain are phenomena (states and actions of entities), patterns among those phenomena, and a higher-order system within which those phenomena and patterns are embedded. Facets at all these levels can be used as partitions in meta-analysis. Many research areas (e.g., consumer choice, complaining behavior) have used samples from a variety of populations (e.g., college students, housewives, community leaders). Dividing the set of studies by using the type of respondent as a partition provides the meta-analyst with an opportunity to gain information concerning the scope and limits of the relation under study.

Research in the area of judgment and choice has compared several models that describe the relation between different states of a single individual (e.g., Jaccard and Becker, in press). For instance, there has been considerable research concerning the type of algebraic model that contains an integration rule that mirrors what individuals do when combining information (e.g., Lynch, 1979; Anderson, 1982). A researcher reviewing the literature that examines one particular algebraic model (e.g., a differential weighted averaging model; Lynch, 1979; Jaccard and Becker, in press) can determine the scope and limits of that model by partitioning the set of studies on the basis of the content of the judgments. If the findings when using a particular model converge across the different content areas (e.g., birth control decisions, altruistic decisions), confidence in that model as a general model of human judgment is increased. Conversely, if the findings do not converge, but the researcher is able to interpret this lack of convergence, confidence in that model also may increase.

A good example of research exploring variations across embedding systems is comparative cross-cultural research. Much of the research done from the Human Relations Area Files, for instance, involves trying to assess the same set of relations within each of a selected set of cultures. This assessment often involves selecting cultures so as to evaluate both the scope and the limits of those relations with respect to culture as an embedding system. In the cross-cultural area, the VNS ideas of scope and boundary search are often expressed in terms of cultural generality and specificity.

Partitioning on Facets in the Methodological Domain

In Chapter 2, we described three levels within the methodological domain: (1) modes of treatment (e.g., measures and manipulations), (2) comparison techniques, and (3) research strategies. The meta-analyst can use facets at each of these levels to partition the body of literature being examined. For example, the body of literature examining the relation between involvement and choice may be partitioned on type of measure, with some studies using physiological measures and others using self-report measures (e.g., Krugman, 1965; Houston and Rothschild, 1978).

At the relation level, different types of planned comparisons (e.g., Campbell and Stanley, 1966; Cook and Campbell, 1979; Judd and Kenny, 1982; Runkel and McGrath, 1972) can vary across several facets such as: (1) the rule used to assign subjects to conditions (i.e., random assignment rule, nonrandom but known assignment rule, unknown assignment rule); (2) treatment of the independent variable (i.e., manipulated, observed); and (3) number of times one respondent receives each mode of treatment (i.e., repeated measures versus single observation designs).

For instance, research in the area of choice set composition and Bayesian inferences (e.g., Ofir and Lynch, 1984; Payne, 1982) contains work using both between- and within-subject factorial designs. The type of design could be used as a facet to partition the set of studies. In research examining the Fishbein theory of reasoned action, researchers have both observed (see Ajzen and Fishbein, 1980) and manipulated (e.g., Lutz, 1977; Miniard and Cohen, 1979) the key independent variables of attitude and subjective norms. Partitioning on mode of treatment of the independent variables (i.e., manipulated or observed) may provide the researcher with the chance to gain new insights concerning the relations among the variables in the model.

At the embedding system level, the set of research strategies also may be used by the meta-analyst to partition a set of studies. For instance, studies of advertising effectiveness include both laboratory and field strategies (e.g., Gardner and Raj, 1983).

Using the research strategy as a variable on which to partition provides the researcher with an additional basis for assessing the scope and limits of findings.

Partitioning on Facets in the Conceptual Domain

The three levels within this domain are properties of phenomena (states and actions of entities), relations among these properties, and a set of conceptual assumptions (i.e., the conceptual paradigm). Each of these levels can be used to partition a body of literature to provide useful information to the meta-analyst.

For instance, various conceptual models of judgment (e.g., Fishbein and Ajzen, 1975; Jaccard and King, 1977; Lynch, 1984) may be compared to determine which model best fits a set of observations. This type of competitive theory testing is analogous to Platt's (1964) argument for strong inference. The meta-analyst summarizing the literature on various choice models can determine the scope and limits of one model, relative to competing models, by examining the set of studies using the type of model as a partition.

The meta-analyst also may gain information about the empirical relation being considered by partitioning studies on the basis of variations in features of the hypothesized relation, such as the form of the function. For instance, research on the attitude-behavior relation could contrast a threshold function postulated by Campbell (1963) with the linear relation postulated in most attitude research.

One prominent conceptual assumption in the behavioral and social sciences is that organisms strive toward homeostasis. Cognitive consistency theories (e.g., Heider, 1958; Festinger, 1957) use this principle as the underpinning for their theories. Research in the area of variety seeking, novelty seeking, and risk taking (e.g., Hagerty, 1983), however, brings this assumption into question. By using this assumption as a basis for partitioning a set of studies, the meta-analyst could gain useful insights for areas of

research that could incorporate either of these assumptions (e.g., information acquisition).

Summary

If the studies in a meta-analysis are treated as homogeneous, that will restrict what the reviewer can learn about the body of literature. We have presented a number of facets on which a meta-analyst could partition such a body of literature. Few, if any, bodies of literature will have studies that vary across all the levels of all of the facets discussed within the conceptual, methodological, and substantive domains. The meta-analyst, however, ought to be alert to use any such partitions that the body of literature permits. Furthermore, the major facets on which the literature cannot be partitioned are clues to areas requiring further (stage three) research.

IMPLICATIONS OF THE VNS FOR INNOVATIONS AND PARADIGM SHIFTS

Kuhn's (1962) now classic work touched off considerable interest in the question of paradigms within the social and behavioral sciences. We do not have sufficient space here to offer an extensive and general critique of that literature. Rather, we will confine our discussion of paradigms to issues on which the VNS system may shed some light.

For the VNS, questions about what is the dominant paradigm, how that came to be, and how change in that paradigm comes about, all have to do with stage one of the research process. Stage two may be viewed as carrying out what Kuhn has called "normal science." Furthermore, the VNS implies that if paradigms and shifts in them are features of stage one, then there must be three of them: a conceptual paradigm, a methodological paradigm, and a substantive paradigm. Kuhn and others who have written on this topic seem to have focused on the conceptual, and to some degree, on the methodological domain. We also will talk about what it means to have a shift in the substantive paradigm.

Paradigm shifts imply existence of a prior paradigm, and some kind of "breakthrough" or key event that leads to a change in that paradigm. Questions of when and why a new paradigm comes about must be asked in conjunction with questions of how the prior paradigm got established, what forces support it and resist changes in it, and what conditions will lead to a breakdown of that resistance and subsequent change in the old paradigm.

In the VNS, we would argue that part of the stage one researcher's task is to seek new materials at all levels (elements, relations, embedding systems) in each of the three domains. This premise opens two questions worth examining here: First, how do changes at each level relate to the idea of paradigm change? Second, if there is a part of the research process devoted to a search for changes, why are paradigms so persistent (at least as viewed by most scholars who have commented on these matters in recent years)? A third question is implicit in the discussion thus far: Do changes in the different domains have different kinds of effects on the overall paradigmatic status of a field?

Before launching our discussion, however, we need to state several assumptions upon which that discussion is based. First, we will take the position that, although *innovation* at the element and relations levels may expand the set of possibilities available at any given time, a *paradigm shift* requires a change at the *embedding system* level. Second, a successful paradigm shift not only involves the *legitimization of a new set of paradigmatic assumptions at the embedding system level,* but also involves the *delegitimization*, the overthrowing, of a major part of the previous paradigm. Third, although for convenience we treat these issues separately for the three domains, a successful paradigm shift will ultimately overthrow or replace portions of the prior paradigm in all three domains. We will begin our consideration of these issues by dealing first with the methodological domain, the realm within which we can most easily discuss these matters.

Paradigms and Innovations in the Methodological Domain

A paradigm in any field represents the more or less coherent set of possibilities that are established for that field at the embedding system level. So, for example, in the methodological domain, we might regard the *dominant paradigm as that set of ideas and possibilities that underlie the set of strategies* noted earlier in this chapter. No one strategy is *the* methodological paradigm for the field, but each is a possibility within that paradigm.

We asserted eight classes of strategies earlier in Chapter 2. That set of strategies is not the paradigm; rather, it is the specifiable set of possibilities that is the paradigm's current embodiment. We do not want to press the case that the strategy circumplex presented in Chapter 2 defines the methodological paradigm for the field. But we do want to make the point that some such matrix of possibilities, along with their associated assumptions at the embedding system level, does constitute the dominant paradigm of the field at any given time. The matrix of possibilities at the embedding system level need not be a closed system, although the strategy circumplex is so. In any case, it is not a completed system. New possibilities could be invented that would become important variants of one or another of the existing set. But such new possibilities would probably not break the bounds of the existing schema; that is, they would elaborate but not overthrow the existing paradigm.

Alternatives that lie outside the current matrix of possibilities are not likely to be invented by researchers who have been trained within the current paradigm. A change in paradigm may be most likely to result from importation of methodological ideas from outside the field, hence outside the existing paradigm.

In general, innovations at the element or relation level in a domain—in the present context, innovations in modes of treatment or in comparison techniques in the methodological domain—extend the present set of available elements and relations but lie within the current matrix of possibilities at the embedding system level. For example, the semantic differential was a new method for obtaining self-reports systematically—as had been Thurstone's technique for measuring attitudes and Likert's general method for obtaining ratings—but all of these fit easily within the currently dominant set of research strategies and supporting assumptions. Methods of the same general class—self-reports in the form of questionnaires and interview protocols—were already in use. Techniques using them in comparisons were already available, and they could be put to use within most of the research strategies already in the matrix of possibilities. In general, *we seldom invent or import methods that we do not already have a place for in our currently dominant paradigm.*

The same is true of innovations at the level of comparison techniques. They tend to expand the currently available set of possibilities, but to do so within the set of possibilities defined by the embedding system already in place. Two examples of such innovations in comparison techniques are Tucker's (1965) three-mode factor analysis and some of the innovations introduced by Coombs (1964) in his theory of data. In both cases, the new comparison techniques undoubtedly brought with them further innovations in data collection and analysis procedures (e.g., the invention of some techniques for nonmetric factor analysis was one consequence of Coombs's scaling ideas). But all of those changes could be carried out within at least some of the potentially available strategies of the field (e.g., judgment studies), and all built on the set of paradigmatic assumptions underlying those strategies.

In the past, there has been much importation of methods into the behavioral and social sciences, largely from the longer-established physical and biological sciences. In many ways, our current methodological paradigm consists of a more or less coherent collection of all of those imports. But because the current matrix of possibilities for the behavioral and social sciences already includes much of what could be imported from those other sciences, we probably should not expect many more innovations from those fields. Perhaps we might get them from the arts and the humanities, whose methodological paradigms are

dramatically different from our own. Or, perhaps we will get them as secondary effect from shifts in either the conceptual or the substantive paradigms of our own field. Or, perhaps we will not get them at all!

What are the conditions under which we do shatter the old and erect the new paradigms in the methodological domain? In the VNS, that would have to come about by a reconstruction of the matrix of possibilities at the strategic, or embedding system, level. What do shifts of that kind look like, and can we identify any that have come about in recent decades?

The Watsonian revolution—out with introspection, in with the measurement of overt behavior—may constitute such a case. To some extent, Watson's innovations involved the popularization and legitimization of strategies that were already available in physical science, but they had not been much used in the social and behavioral sciences at that time. What is important is the dramatic shift in the set of strategies that were regarded as potentially available: Not just the *legitimization* and popularization of some new or little-known ones, but the *delegitimization* of those that had previously been dominant ones.

In contrast, the stirrings in social psychology, beginning in the middle 1960s, that urged less work in laboratory settings and more in natural settings, was not such a shift in the methodological paradigm. Strategies for conducting studies in natural settings were already available and quite widely known—though perhaps too infrequently used—both in social psychology and in nearby areas of social and behavioral science. In fact, some of the classic studies done prior to that time (e.g., Mead, 1949, Newcomb, 1943; Sherif et al., 1961; Whyte, 1943) were done in natural settings. Furthermore, these stirrings did not in fact lead to a substantial shift in the set of potentially available strategies. They may have led to minor increases in the use of natural settings in some areas of social psychology, but they certainly did not delegitimize the laboratory. And, in fact, many of the studies that were billed as involving the new paradigm—that is, natural settings—

in fact were attempts to gain, by compromise, the advantages of both the lab and the field. So, even if this could have led to a paradigm shift had it been successful, it did not do so because the innovation did not "take" to any substantial degree.

Paradigms and Innovations in the Conceptual Domain

In the conceptual domain, innovations at the element level (properties) happen frequently. Innovations at the relation level are rather rare. Innovations at the embedding system level are extremely rare, and it is here that the apparent persistence and unshakability of dominant paradigms is most clear cut. The dominant conceptual paradigm of the social and behavioral sciences is very robust; as with the embedding system level for the methodological domain, researchers have difficulty considering ideas that lie outside the currently available set of possibilities.

At the element level in the conceptual domain, new properties are invented every day in the social and behavioral sciences. Some of these often seem to be minor restatements of other concepts already in the available set. [Does Milgram's (1965) obedience differ substantially from Asch's (1951) or others' conformity? Is Bandura's (1962) modeling really different from Miller and Dollard's (1941) imitation?] But other "new" concepts seem to have been genuine innovations. For example, although Thibaut and Kelley's (1959) comparison level (CL) is similar to the earlier concept of adaptation level, their comparison level for alternatives—CL(alt)—seems to have been a really new concept when they proposed it. That concept subsumes the economist's notion of opportunity costs, but entails more than that. Still, element level concepts do not break the bounds of the currently dominant conceptual paradigm, with its constraining assumptions (e.g., equilibrium). As with the element level in the methodological domain, *we seldom invent concepts for which we do not already have a place in the currently dominant conceptual paradigm.*

Innovations of any kind at the relations level, in the conceptual domain, seem relatively rare. For example, we seem to persist in

conceptualizing relations as linear, or else as some simple form of curvilinear relation, or at the most as a symmetrically nonmonotonic one. This persistence occurs in spite of considerable evidence that more complex formulations (such as cyclical patterns) would offer a better fit to the substantive phenomena involved. Only rarely does a nonmonotonic relation become a central part of a conceptual formulation (e.g., the widely touted "inverted U" theories of the relation between stress and performance).

Our field's preference for linear formulations undoubtedly occurs in part because we are most familiar with, and most likely to have access to, techniques for dealing with linear relations in the methodological domain. But it may go deeper than that. There seems to be some evidence that humans have trouble appreciating the full implications of nonlinear, and especially of nonmonotonic, relations. They certainly have trouble using such relations in combination with linear ones in predicting complex outcomes (e.g., Brehmer, 1976). In any case, changes from linear forms come slowly, if at all.

We can regard the Gestalt "revolution" as a revolt against the then dominant conceptual paradigm, partly at the relations level but mainly at the level of the embedding conceptual system. The Gestalt theorists made some new assumptions about the nature of the material under study—humans and human behavior—that challenged important parts of the then dominant paradigm. For example, they argued for an ahistorical or systemic form of causal relations among system parts, in contrast to the then dominant historical and quasi-mechanical view of causal forces. We would regard the Gestalt movement as a successful innovation; that is, it successfully introduced some new conceptual possibilities. But, whether it is to be regarded as a successful paradigm shift is a matter of definition. The Gestalt movement expanded the set of alternative assumptions available for use, but it did not replace the previously dominant assumptions that it tried to overturn.

We can perhaps regard the Freudian movement as a major revolution in the conceptual paradigm of psychology (and, subsequently, in many of the social sciences, in the arts and humanities, and the whole lay culture as well). Freud attempted to overthrow the prior paradigm, in which mental illness was treated, perforce, as the consequence of sin and therefore a moral matter. He attacked at all three levels: what properties were appropriate for study (e.g., libido, superego), what kind of relations held among those properties (e.g., repression, projection), and in terms of the nature of the embedding systems and associated conceptual assumptions (e.g., underlying life and death forces, fixed developmental stages, interpretative treatment of outcroppings of material at the unconscious level).

Not all of the properties, relations, and conceptual assumptions of the prevailing paradigm were attacked. For example, Freud placed heavy emphasis on the idea of equilibrium, as had most social scientists before him and most since then. Freud also accepted, but reinterpreted the meaning of, properties such as task competence, affection, and symptoms of physical disorders. The Freudian view prevailed, both in providing a new set of ways to look at matters and, in large part, in delegitimizing the old ways. Hence, it can be regarded as an important (and rare) instance of a conceptual paradigm shift in psychology. The Freudian paradigm, some argue, has now itself been overthrown in clinical psychology by a variation of Watsonian behaviorism. If so, that would be further evidence that the behavioristic revolution was a major paradigm shift in the field.

In recent years, there has been some discourse within social psychology (and other areas of social and behavioral science) on what is sometimes referred to as a dialectics movement. Insofar as this approach questions the basic conceptual assumptions of the field, it can be regarded as a revolt against the dominant conceptual paradigm. In at least one version (Altman et al., 1981), it is clearly such an attack. Altman and colleagues argue for substituting the ideas of dialectic conflict and change for the prevailing ideas of equilibrium and teleological progression. Because the new conceptual assumptions may well be incompatible with the current ones, this effort may lead to a paradigm shift (or fail to produce such a shift) rather than to a mere expansion of the old set of possibilities. It is still far too soon to tell.

Paradigms and Innovations in the Substantive Domain

Most discussions of paradigms in the behavioral and social sciences deal with conceptual and methodological paradigms, or perhaps with the two merged together. The structure of the VNS suggests that, because there is a third domain there is a third form of, or location for, a shift in paradigm. This suggestion, of course, may merely reflect the Procrustean forces at work within such a schema. But we will set aside that possibility and proceed on the assumption that the concepts of paradigms, innovations, and paradigm shifts are also appropriate for application to the substantive domain.

The idea of innovation in the substantive domain poses some basic philosophical problems. To say that we have *invented* some new substantive phenomena, or potential patterns of them, or substantive systems, is to imply that "there is something new under the sun" so to speak, and we have created it. On the other hand, to say that we have *discovered* some new substantive phenomena, patterns, and systems is to imply that *we have some means of discovery that lies outside the knowledge-gaining framework (science) that would require use of both the conceptual and the methodological domains as well as the substantive domain.* We do not wish to take a position on these very fundamental philosophical questions, nor do we think it necessary to do so to address this topic. Instead, we will talk about the initial step as "identifying" some new elements, relations, or systems—new in the sense that they were not perceived as being in the set of available possibilities before—without specifying the nature of the identification process (e.g., generation, discovery, or construction from parts already there).

The identification of new elements, relations, and systems in the substantive domain, within the social and behavioral sciences, often involves specifying either a new system level or a new behavioral process—new in the sense that they have not been studied before *in that field.* For example, within social psychology in the early 1950s, there was a movement to study small groups (and to do so in controlled laboratory settings). That movement constituted a recognition that groups might be a unit of analysis that encompassed different phenomena than those to be studied at the individual level. Studies of groups then became a veritable flood, with the floodtide ebbing and finally receding in the latter half of the 1960s.

During the high tide of group research, and because of its domination of social psychology, many behavioral phenomena were studied within groups, some of which did not at all need to be studied in a group context. For example, studies of individual learning and problem solving were often done in the context of a task-oriented group, even when the group's task simply involved the same task for each of the individual members, done in coaction rather than interaction. But during that time of ascendancy of group research, a small group was the most prominant or popular substantive system for social psychological study. It had legitimacy.

There are several hypotheses about why the group research "movement" came and went, and did or did not come back again (e.g., Steiner, 1974, 1983; McGrath and Altman, 1966; McGrath and Kravitz, 1982; McGrath, 1984). For whatever reasons, the course of these changes reflects shifts in what was considered an appropriate substantive system for study.

It has been assumed that the drop in popularity of the "group" as a unit for study was accompanied automatically by a rise in the use of "individual" as the unit of study. That assumption is probably correct to some extent, but the matter is somewhat more complex. The demise of interest in study of small groups was accompanied by a shift in what behavioral processes were to be the focus of study—the so-called "cognitive revolution" in social psychology. Much of the work during the heyday of groups dealt with affective and conative or behavioral processes, as they occurred within groups. Relatively little of that work would now be regarded as studying cognitive processes.

The shift to a focus on cognitive processes was a shift not from groups to individuals as units of study, but from intra-group processes to certain intra-individual processes as the focus of study. With that—by no means coincidentally—came a shift from groups to individuals as the embedding system for that study.

Such shifts, in either focus of study or the embedding systems of that study, sometimes are more apparent than real. One upshot of the drop in popularity of small groups as the substantive system of choice for social psychologists (and also the drop in popularity of interpersonal affective and behavioral processes as the preferred focus of study) was a shift by many small group researchers to research dealing with large work organizations. But this shift turned out to be anything but a paradigm shift, or even a crucial shift in embedding system. Some of the organization research done by such newly shifted small group researchers consisted of studies of interpersonal affective and behavioral processes *in groups that happened to be located within organizations*—hardly a paradigm-shattering change.

One can make the case, too, that both real and apparent paradigm shifts in the substantive domain tend to be cyclical rather than directional, that what has gone will come back again (perhaps in somewhat disguised form). Groups became the dominant substantive system for social psychological study, then faded in popularity, then perhaps returned. A near exclusive emphasis on cognitive processes replaced a dominance by affective and motivational processes, but recently social cognition researchers have begun to bring affect back into the picture. These cycles may highlight another underlying feature of these matters: Substantive paradigms may be most resistent of all to major and dramatic change that actually overthrows what was there before, or replaces it in anything but temporary popularity.

Interrelations Across Domains

The reader by now recognizes that the separation of the research process into the three domains represents a markedly Procrustean act. The three domains are closely related to one another. Although analytically separable, they always *exist* together. Any substantial change in paradigm within any one of the domains has consequences—reverberations—throughout the other levels of that domain and throughout the other domains as well.

For example, when the Freudian revolution brought new concepts and relations into consideration, techniques for gathering evidence about "the id" and "the unconscious" and the process of repression were needed. Also, means to identify substantive units, relations, and embedding systems for these operations was needed. When Watson called for the measurement of actual behavior, not the reporting of introspective impressions, investigators had to develop concepts that would make meaningful the resulting measurements (e.g., concepts like reinforcement, "shaping," and extinction) and to give up other concepts that had been tied to the introspective method.

When social psychologists' attention turned from affective and behavioral processes within groups to cognitive processes within individuals, they needed to invent new methods for measuring those intangible and internal processes (some of which turned out to be a reincarnation of the long-ostracized methods of introspection!) and new conceptual tools to use in their interpretation. In contrast, the current dialectical movement has not yet resulted in a paradigmatic shift. Its proponents have not yet proposed a viable set of methods for adequate assessment of the more dynamic and fluid processes that follow from its shifts in paradigmatic assumptions (from linear to dialectic, from emphasis on equilibrium to emphasis on change, and so forth). Although the paradigm in some one domain always shifts first, no paradigm shift can really prevail unless it produces a thorough-going change at all levels in the paradigms of all three domains.

SUMMARY AND LIMITATIONS OF THE VNS

The VNS offers a view of the nature of the research process that is an alternative to the view usually presented in describing social

and behavioral science research. Although the VNS is limited in ways different than that typical view, we believe it can provide researchers with a more integrated, and a more complex, treatment of the research process. The VNS, we believe, has several advantages as a framework for understanding the research process:

(1) The VNS provides a basis for organizing and interpreting various forms and meanings of validity.
(2) The VNS makes salient several research issues that have largely been ignored. For example, the VNS insists on the need to establish *both* the scope and limits of a finding; and it represents work in stage one (developing, clarifying, and refining the materials within each of the domains) as an integral part of the research process.
(3) The VNS describes various research styles and orientations, and links them to the research process and to the researchers' purposes.
(4) The VNS provides a basis upon which to interpret the activities of the researcher as part of the research process.

The VNS also can be used as a framework from which to develop ideas. In our earlier discussion of innovation, we noted that innovations often come about by "importing" new elements and relations into one of the three domains from some other field. If such new elements and relations in one of the domains are paired with material typically used in a second domain, this new pairing will result in a new design, a new hypothesis, or a new set of observations. These new structures may provide the researcher with an opportunity to understand the focal problem better. For instance, Lewin borrowed the concepts of valence and force fields from physics and applied them to a substantive problem in the social sciences (i.e., group decision making) to try to develop a better understanding of that focal problem.

The VNS also can offer a set of guidelines for conducting and evaluating research and for understanding the research process. A fundamental premise of the VNS is that there are multiple pathways for conducting research and that adherence to any single pathway, to the exclusion of the other pathways, limits what a researcher (and a discipline) is able to learn. Each pathway has its starting point in a different domain, and researchers following the different pathways are likely to cope with mismatches in different ways. The use of multiple strategies can reduce the likelihood that the findings about a focal problem will be constrained by a single orientation and a single strategy for coping with mismatches.

Stage two researchers can use the VNS to make more salient the types of criteria that are in each of the domains. Those criteria have an impact on what kinds of elements and relations are available, hence what get used in stage two research.

Stage two researchers also can use the features of the VNS matching process to highlight potential sources of noncorrespondence. Specifically, they should be aware that: (1) lack of specification of the features of the relations in the conceptual domain will lead to mismatches; (2) the inability of comparison techniques to differentiate among the levels of the features of relations will lead to mismatches; and (3) the selection of material from the substantive domain so that it cannot display various levels of the features of relations will lead to mismatches. By addressing these three sources of noncorrespondence, researchers are likely to increase the match among features of the relations across domains and thereby increase the information they can gain from a study.

The VNS also views a finding from a single study on any focal problem as quite limited in the information that it contains. Only when researchers have examined the scope and limits associated with that finding (i.e., carried out the research activities of stage three) will that finding increase our knowledge about the focal problem.

The VNS also highlights the interdependence of research activities in the three stages of the research process. For a field to advance, none of these stages can be neglected. To date, there has been a heavy emphasis in the behavioral and social sciences on research activities in stage two. We believe that the behavioral

and social sciences need to give greater attention to research activities in stage one and greater emphasis on establishing the replicability, scope, and limits of findings (i.e., research activities of stage three).

The VNS also can be used to provide some insight into the argument concerning the "critical experiment" (e.g., Garner et al., 1956). Researchers have been urged to design research that critically compares two (or more) theoretical explanations. At one level, we are in complete agreement with that position; that is, knowledge is advanced when researchers can identify a set of concepts that uniquely and adequately describes a set of observations. No single study, however, can provide sufficient evidence to support or refute any theory (any set of comparisons in the conceptual domain) because such a procedure does not address the uncertainty associated with the other two domains—substantive and methodological. Only after research has examined the scope and limits of a set of findings concerning a focal problem, across many facets of all three domains, can we be confident about that finding.

We think that the VNS also can provide some insights into certain epistemological issues within the research process. The three stages represent different forms of knowledge acquisition. The very nature of the scientists' work tends to differ for the different stages. Moreover, we think those differences are related to some of the continuing controversies in the philosophy of science. It is in stage two, the central or execution stage, that the scientist seems to be following a "policy" that mirrors the rules and practices of logical empiricism. Here we see "scientist as impersonal investigator," as impartial tester with no vested interest in any particular outcome. It is the stage two scientists' *intent* to be objective toward, if not aloof with regard to, the material being studied. In stage one (the prestudy or generation stage) on the other hand, the scientist seems to be following a policy that mirrors the rules and practices of "constructionism." Here we see "scientist as active participant," as creative generator. The stage one scientist is subjective in that he or she is intellectually, if not emotionally, involved with the materials being studied.

Stage three, the followup or generalization stage of the research process, is harder to characterize in such terms, in part because it has been given far less attention by behavioral and social scientists. In stage three, the scientist seems to be following a policy that mirrors the rules and practices of "interpretative pragmatism." Here, we see "scientists as accreditor," as explorer of the credibility of findings. The scientist has a commitment, not to the findings as such, but to their interpretation and use.

There are costs, however, with any schema that imposes structure on a complex set of relations. The major strength in the VNS is also its major weakness. It makes many distinctions, and imposes an elaborate structure on the material. We have used a set of terms in describing these distinctions, that overlaps minimally with familiar terms that have become laden with surplus meaning. For example, we have avoided formal use of the terms theory, model, and data, because all of them have surplus meaning in our field. The cost of avoiding the surplus meaning associated with previously used terms unfortunately is an awkwardness of language that comes from use of new and unfamiliar terms. Thus, the VNS is very jargon-laden.

The structuredness of the VNS also can mislead. We have described stages of research, levels, pathways, and so forth as if these were clearly distinct. We realize that there are numerous feedback loops among the stages and between the steps within each of the pathways. To reduce the level of complexity in our presentation of the VNS, however, we have acted *as if* the separation between the stages and among the steps were clear cut. This orderliness gives the VNS a strong doctrinaire flavor that we do not intend.

Furthermore, the very complexity of the system places a burden on its exposition and, thereby, probably hinders understanding and use. But as we suggested in Chapter 2, there is a dilemmatic trade-off between comprehensiveness, which we sought, and parsimony, which we had to sacrifice.

Such a strongly structured system also induces strong Procrustean forces in the developers and users of that system. The tendency to make virtually any material fit the schema is very strong,

and the reader is warned to anticipate that we have probably forced some fits that might not be as "obvious" to others as they seem to us. We urge the reader to turn that problem to a potential gain, by viewing any feature of the VNS that seems contrived or inappropriate as a spot in which the system likely needs expansion and revision.

Although we have tried, in this chapter, to show some ways in which the VNS can be used, basically the VNS is not just a *tool to be used* but also a *point of view* about research. It is a framework for considering the fields of social and behavioral science, and the research processes within them. It forces us to ask questions we have not asked before—about the literature bearing on certain focal problems, about how we go about doing research. It is a *schema for raising consciousness about research.* That is its value!

References

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior.* Englewood Cliffs, NJ: Prentice-Hall.

Altman, I., Vinsel, A. M., & Brown, B. B. (1981). Dialectic conceptions in social psychology: An application to social penetration and privacy regulation. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14). New York: Academic Press.

Anderson, N. H. (1982). *Foundations of information integration theory.* New York: Academic Press.

Asch, S. (1951). Effects of group pressure upon the modification and distortion of judgment. In H. Guetzkow (Ed.), *Groups, leadership and men.* Pittsburgh: Carnegie Press.

Atkinson, J. W., & Feather, N. T. (1965). *A theory of achievement motivation.* New York: John Wiley.

Bagozzi, R. (1984). A prospectus for theory construction in marketing. *Journal of Marketing, 21,* 11-29.

Bales, R. F. (1950). *Interaction process analysis: A method for the study of small groups.* Cambridge, MA: Addison-Wesley.

Bales, R. F., & Slater, P. E. (1955). Role differentiation. In T. Parsons, R. F. Bales & others (Eds.), *The family, socialization, and interaction process.* Glencoe, IL: Free Press.

Bales, R. F., & Strodtbeck, F. L. (1951). Phases in group problem solving. *Journal of Abnormal and Social Psychology, 46,* 485-495.

Bandura, A. (1962). Social learning through imitation. In M. R. Jones (Ed.), *Nebraska symposium on motivation, 1962.* Lincoln, NB: University of Nebraska Press.

Berkowitz, L., & Donnerstein, E. (1982). External validity is more than skin deep: Some answers to criticism of laboratory experiments. *American Psychologist 37,* 257-275.

Bion, W. R. (1961). *Experiences in groups, and other papers.* New York: Basic Books.

Birnbaum, M. H. (1982). Controversies in psychological measurement. In B. Wegener (Ed.), *Social attitudes and psychophysical measurement.* Hillsdale, NJ: L. Erlbaum.

Birnbaum, M. H., & Mellers, B. A. (1979). Stimulus recognition may mediate exposure effects. *Journal of Personality and Social Psychology, 37,* 391-394.

Blalock, H. M. (1982). *Conceptualization and measurement in the social sciences.* Beverly Hills, CA: Sage Publications.

Boehm, V. E. (1980). Research in the "Real World"—A conceptual model. *Personnel Psychology, 33*, 495-503.

Brehmer, B. (1976). Social judgment theory and the analysis of interpersonal conflict. *Psychological Bulletin, 86*, 307-324.

Brewer, M. B., & Collins, B. E. (1981). *Scientific inquiry and the social sciences*. San Francisco: Jossey-Bass.

Brinberg, D. (1979). An examination of the determinants of intention and behavior: A comparison of two models. *Journal of Applied Social Psychology, 6*, 560-575.

Brinberg, D., & McGrath, J. E. (1982). A network of validity concepts within the research process. In D. Brinberg & L. Kidder (Eds.), *New directions for methodology of social and behavioral science: Forms of validity in research* (No. 12). San Francisco: Jossey-Bass.

Calder, B. J., Phillips, L. W., & Tybout, A. M. (1981). Designing research for application. *Journal of Consumer Research, 8*(2), 197-207.

Calder, B. J., Phillips, L. W., & Tybout, A. M. (1982). The concept of external validity. *Journal of Consumer Research, 9*(3), 240-244.

Calder, B. J., Phillips, L. W., & Tybout, A. M. (1983). Beyond external validity. *Journal of Consumer Research, 10*(1), 112-114.

Campbell, D. T. (1963). Social attitudes and other acquired behavioral dispositions. In S. Koch (Ed.), *Psychology: A study of science* (Vol. 9). New York: McGraw-Hill.

Campbell, D. T. (1981). Comment: Another perspective on a scholarly career. In M. B. Brewer & B. E. Collins (Eds.), *Scientific inquiry and the social sciences*. San Francisco: Jossey-Bass.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 30*, 81-105.

Campbell, D. T., & O'Connell, E. J. (1982). Methods as diluting trait relationships rather than adding irrelevant systematic variance. In. D. Brinberg & L. Kidder (Eds.), *New directions for methodology of social and behavioral science: Forms of validity in research* (No. 12). San Francisco: Jossey-Bass.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand-McNally.

Cook, T. D., & Campbell, D. T. (1979). *Design and analysis of quasi-experiments for field settings*. Chicago: Rand-McNally.

Coombs, C. H. (1964). *A theory of data*. New York: John Wiley.

Cronbach, L. J., and others. (1972). *The dependability of behavioral measurement*. New York: John Wiley.

Cronbach, L. J. (1975). Beyond the two disciplines of psychology. *American Psychologist, 30*, 116-127.

Cronbach, L. J. (1982). *Designing evaluation of educational programs*. San Francisco: Jossey-Bass.

Derrick, F. W., & Lehfield, A. K. (1980). The family life cycle: An alternative approach. *Journal of Consumer Research, 7*, 214-217.

Ellsworth, P. C. (1977). From abstract ideas to concrete instances: Some guidelines for choosing natural research settings. *American Psychologist, 32*, 604-615.

Endler, N. S., & Magnusson, D. (1976). *Interactional psychology and personality*. Washington DC: Hemisphere Publishing.

Engel, J. F., & Blackwell, R. D. (1982). *Consumer behavior*. New York: Dryden Press.

Eysenck, H J. (1978). An exercise in meta-silliness. *American Psychologist, 33*, 517.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Feyerabend, P. K. (1970). Against method: An outline of an anarchistic theory of knowledge. In M. Radner & S. Winoker (Eds.), *Minnesota studies in the philosophy of science* (Vol. 4). Minneapolis: University of Minnesota Press.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.

Fiske, D. W. (1982). Convergent-discriminant validity in measurements and research strategies. In D. Brinberg & L. Kidder (Eds.), *New directions for methodology of social and behavioral science: Forms of validity in research* (No. 12). San Francisco: Jossey-Bass.

Gardner, M. P., & Raj, S. P. (1983). Responses to commercials in laboratory versus natural settings: A conceptual framework. In R. Bagozzi & A. Tybout (Eds.), *Advances in consumer research* (Vol. 10). 142-146.

Garner, W. R., Hake, H. W., & Eriksen, C. W. (1956). Operationism and the concept of perception. *Psychological Review, 63*, 149-159.

Glaser, B. G., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.

Glass, G. V (1976). Primary, secondary and meta-analysis of research. *Educational Researcher, 5*, 3-8.

Glass, G. V, & Kliegl, R. M. (1983). An apology for research integration in the study of psychotherapy. *Journal of Consulting and Clinical Psychology, 51*, 28-41.

Glass, G. V, McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage Publications.

Grice, R. G., & Hunter, J. J. (1964). Stimulus intensity effects depend upon the type of experimental design. *Psychological Review, 71*, 247-256.

Guilford, J. P. (1954). *Psychometric methods* (2nd ed.). New York: McGraw-Hill.

Hagerty, M. R. (1983). Variety seeking among songs which vary in similarity. In R. Bagozzi & A. Tybout (Eds.), *Advances in consumer research* (Vol. 10). 75-79.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: John Wiley.

Hoffman, L. W. (1977). Fear of success in 1965 and 1974: A follow-up study. *Journal of Counseling and Clinical Psychology, 45*, 310-321.

Horner, M. S. (1968). Sex differences in achievement motivation and performance in competitive and non-competitive situations. *Dissertation Abstracts International, 30*, 407B. University Microfilms No. 69-12-135.

Horner, M. S. (1972). Toward an understanding of achievement related conflicts in women. *Journal of Social Issues, 28*, 157-175.

Houston, M. J., & Rothschild, M. L. (1978). Conceptual and methodological perspectives on involvement. In S. Jain (Ed.), *Research frontiers in marketing: Dialogues and directions, 1978 educators' proceedings*. Chicago: American Marketing Association.

Hovland, C. I., Lumsdale, A. A., & Sheffield, F. D. (1949). *Experiments on mass communication*. New York: John Wiley.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: Cumulating research findings across studies*. Beverly Hills, CA: Sage Publications.

Jaccard, J., & Becker, M. A. (in press). Attitudes and behavior: An information integration perspective. *Journal of Experimental Social Psychology*.

Jaccard, J., & King, G. W. (1977). A probabilistic model of the relationship between beliefs and behavioral intentions. *Human Communications Research, 3*, 332-342.

Jaccard, J., Knox, R., & Brinberg, D. (1979). Prediction of behavior from beliefs: An extension and test of a subjective probability model. *Journal of Personality and Social Psychology, 37*, 1239-1249.

Judd, C. M., & Kenny, D. A. (1982). Research design and research validity. In D. Brinberg & L. Kidder (Eds.), *New directions for methodology of social and behavioral science: Forms of validity in research* (No. 12). San Francisco: Jossey-Bass.

Krugman, H. E. (1965). The impact of television advertising: Learning without involvement. *Public Opinion Quarterly, 29*, 349-356.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Lewin, K. (1953). Studies in group decision making. In D. Cartwright & A. Zander (Eds.), *Group dynamics: Research and theory*. (1st ed.). Evanston, IL: Row, Peterson.

Lutz, R. (1977). An experimental investigation of causal relations among cognitions, affect and behavioral intentions. *Journal of Consumer Research, 3*, 197-208.

Lynch, J. G. (1979). Why additive models fail as descriptors of choice behavior. *Journal of Experimental Social Psychology, 15*, 397-417.

Lynch, J. G. (1982). On the external validity of experiments in consumer research. *Journal of Consumer Research, 9*(3), 225-239.

Lynch, J. G. (1983). The role of external validity in theoretical research. *Journal of Consumer Research, 10*(1), 109-111.

Lynch, J. G. (1984). Adventures in paramorphic modeling: Models of consumers' processing of negative information. Working paper, Gainesville, University of Florida, Center for Consumer Research.

Mackenzie, K. D., & House, R. (1978). Paradigm development in the social sciences: A proposed research strategy. *Academy of Management Review, 3*(1), 7-23.

McClelland, D. C. (1965). Toward a theory of motive acquisition. *American Psychologist, 20*, 321-322.

McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. (1953). *The achievement motive*. New York: Appleton-Century-Crofts.

McGrath, J. E. (1984). *Groups: Interaction and performance*. Englewood Cliffs, NJ: Prentice-Hall.

McGrath, J. E., & Altman, I. (1966). *Small group research*. New York: Holt.

McGrath, J. E., & Brinberg, D. (1983). External validity and the research process: A third-party comment on the Calder/Lynch dialogue. *Journal of Consumer Research, 10*(1), 115-124.

McGrath, J. E., & Brinberg, D. (1984). Alternative paths for research: Another view of the basic vs. applied distinction. In S. Oskamp (Ed.), *Applied social psychology annual*. Beverly Hills, CA: Sage Publications.

McGrath, J. E., & Kravitz, D. A. (1982). Group research. *Annual review of psychology, 33*, 195-230.

McGrath, J. E., Martin, J., & Kulka, R. A. (1982). *Judgment calls in research*. Beverly Hills, CA: Sage Publications.

Mead, M. (1949). *Male and female*. New York: Morrow.

Milgram, S. (1965). Some conditions of obedience and disobedience to authority. *Human Relations, 18*(1), 57-75.

Miller, N. E., & Dollard, J. (1941). *Social learning and imitation*. New Haven, CT: Yale University Press.

Miniard, P. W., & Cohen, J. B. (1979). Isolating attitudinal and normative influences in behavioral intention models. *Journal of Marketing Research, 16*, 102-110.

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist, 38*(4), 379-387.

Newcomb, T. M. (1943). *Personality and social change*. New York: Dryden Press.

Newcomb, T. M. (1961). *The acquaintance process*. New York: Holt, Rinehart & Winston.

Nunnally, J. C. (1967). *Psychometric theory*. New York: McGraw-Hill.

Ofir, C., & Lynch, J. G. (1984). Context effects on judgment under uncertainty. *Journal of Consumer Research, 11*(2), 668-679.

Osgood, C. E., Suci, G. I., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.

Payne, J. (1982). Contingent decision behavior. *Psychological Bulletin, 92*, 382-402.

Platt, J. R. (1964). Strong inference. *Science, 146* (Whole No. 3642), 347-353.

Popper, K. R. (1959). *The logic of scientific discovery*. New York: Basic Books.

Psathas, G. (1960). Phase movement and equilibrium tendencies in interaction process on psychotherapy groups. *Sociometry, 23*(2), 177-194.

Rokeach, M., & Fruchter, B. (1956). A factorial study of dogmatism and related concepts. *Journal of Abnormal and Social Psychology, 53*, 356-360.

Rosenthal, R. (1978). Combining results of independent studies. *Psychological Bulletin, 85*, 185-193.

Rosenthal, R. (1982). Valid interpretation of quantitative research results. In D. Brinberg & L. Kidder (Eds.), *New directions for methodology of social and behavioral science: Forms of validity in research* (No. 12). San Francisco: Jossey-Bass.

Runkel, P. J., & McGrath, J. E. (1972). *Research on human behavior: A systematic guide to method*. New York: Holt, Rinehart & Winston.

Sherif, M., Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1961). *Intergroup conflict and cooperation*. Norman, OK: Institute of Group Relations.

Stanley, J. C. (1967). *Improving experimental design and statistical analysis*. Chicago: Rand-McNally.

Steiner, I. D. (1974). What ever happened to the group in social psychology? *Journal of Experimental Social Psychology, 10*, 93-108.

Steiner, I. D. (1983). Whatever happened to the touted revival of the group? In Blumberg, H. H., Hare, A. P., Kent, V., & Davies, M. F. (Eds.), *Small groups and social interaction* (Vol. 2). New York: John Wiley.

Stock, D., & Thelen, H. A. (1958). *Emotional dynamics and group culture: Experimental studies of individual and group behavior*. New York: New York University Press.

Talland, G. A. (1955). Tasks and interactions process: Some characteristics of therapeutic group discussion. *Journal of Abnormal and Social Psychology, 50*, 105-109.

Thelen, H. A. (1956). Emotionality of work in groups. In L. D. White (Ed.), *The state of social sciences*. Chicago: University of Chicago Press.

Thelen, H. A., Stock, D., & others. (1954). *Methods for studying work and emotionality in group operation*. Chicago, IL: University of Chicago, Human Dynamics Laboratory.

Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. New York: John Wiley.

Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology, 33*, 529-554.
Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.
Thurstone, L. L., & Chave, E. J. (1929). *The measurement of attitude*. Chicago: University of Chicago Press.
Tucker, L. R. (1965). Experiments in multi-mode factor analysis. In C. W. Harris (Ed.), *Proceedings of the 1964 invitational conference on testing problems* (pp. 46-57). Princeton, NJ: Educational Testing Service.
Weick, K. E. (1967). Organizations in the laboratory. In V. R. Vroom (Ed.), *Methods of organizational research*. Pittsburgh: University of Pittsburgh Press.
Whyte, W. F. (1943). *Street corner society: The social structure of an Italian slum*. Chicago: University of Chicago Press.
Wilson, G. T., & Rachman, S. J. (1983). Meta-analysis and the evaluation of psychotherapy outcome: Limitations and liabilities. *Journal of Consulting and Clinical Psychology, 51*, 54-64.
Wimsatt, W. C. (1981). Robustness, reliability and overdetermination. In M. B. Brewer & B. E. Collins (Eds.), *Scientific inquiry and the social sciences*. San Francisco: Jossey-Bass.

Author Index

Subject Index

About the Authors

David Brinberg is Associate Professor of Marketing at Baruch College, City University of New York. He received his Ph.D. in psychology from the University of Illinois, Urbana. His research interests include attitude models and their use; the analysis of interaction processes in couple decision making; resource exchange in interpersonal relations; and the study of forms of validity in the research process.

Joseph E. McGrath is Professor of Psychology at the University of Illinois, Urbana. He received his Ph.D. from the social psychology program at the University of Michigan. His areas of research interests include research on small group interaction and performance; studies of negotiation; research on stress in individuals, groups, and organizations; research on the social psychology of time; and studies of validity and other aspects of research methodology.