HANDBOOK of IMPROVING PERFORMANCE IN THE WORKPLACE

Instructional Design and Training Delivery | Selecting and Implementing Performance Interventions | Measurement and Evaluation

Volume 1: Instructional Design and Training Delivery

EDITED BY Kenneth H. Silber and Wellesley R. Foshay

International Society for Performance Improvement
WHERE KNOWLEDGE BECOMES KNOW-HOW

Handbook of Improving Performance in the Workplace

Volume One

Instructional Design and Training Delivery

Join Us at
Pfeiffer.com

Pfeiffer®
An Imprint of
WILEY

Register at **www.pfeiffer.com/email**
for more information on our publications,
authors, and to receive special offers.

ABOUT ISPI

The International Society for Performance Improvement (ISPI) is dedicated to improving individual, organizational, and societal performance. Founded in 1962, ISPI is the leading international association dedicated to improving productivity and performance in the workplace. ISPI reaches out to more than 20,000 performance improvement professionals in over 40 countries through publications and educational programs.

ISPI's mission is to develop and recognize the proficiency of our members and advocate the use of Human Performance Technology. This systematic approach to improving productivity and competence uses a set of methods and procedures and a strategy for solving problems for realizing opportunities related to the performance of people. It is a systematic combination of performance analysis, cause analysis, intervention design and development, implementation, and evaluation that can be applied to individuals, small groups, and large organizations.

Website: www.ispi.org
Mail: International Society for Performance Improvement
1400 Spring Street, Suite 400
Silver Spring, Maryland 20910 USA
Phone: 1.301.587.8570
Fax: 1.301.587.8573
E-mail: info@ispi.org

International Society for Performance Improvement

WHERE KNOWLEDGE BECOMES KNOW-HOW

About Pfeiffer

Pfeiffer serves the professional development and hands-on resource needs of training and human resource practitioners and gives them products to do their jobs better. We deliver proven ideas and solutions from experts in HR development and HR management, and we offer effective and customizable tools to improve workplace performance. From novice to seasoned professional, Pfeiffer is the source you can trust to make yourself and your organization more successful.

Essential Knowledge Pfeiffer produces insightful, practical, and comprehensive materials on topics that matter the most to training and HR professionals. Our Essential Knowledge resources translate the expertise of seasoned professionals into practical, how-to guidance on critical workplace issues and problems. These resources are supported by case studies, worksheets, and job aids and are frequently supplemented with CD-ROMs, websites, and other means of making the content easier to read, understand, and use.

Essential Tools Pfeiffer's Essential Tools resources save time and expense by offering proven, ready-to-use materials—including exercises, activities, games, instruments, and assessments—for use during a training or team-learning event. These resources are frequently offered in looseleaf or CD-ROM format to facilitate copying and customization of the material.

Pfeiffer also recognizes the remarkable power of new technologies in expanding the reach and effectiveness of training. While e-hype has often created whizbang solutions in search of a problem, we are dedicated to bringing convenience and enhancements to proven training solutions. All our e-tools comply with rigorous functionality standards. The most appropriate technology wrapped around essential content yields the perfect solution for today's on-the-go trainers and human resource professionals.

Pfeiffer
www.pfeiffer.com

Essential resources for training and HR professionals

DEDICATION

To the pioneers of instructional design (whose names have been lost or are unknown to the current practitioners of our craft) who, in the late 1960s and early 1970s, shaped and nurtured the ID field in its early days, and taught us, the editors and authors of this volume, the theory and practice of ID.

The complete list of people this book is dedicated to is everyone mentioned in Mike Molenda's Chapter Three, Origins and Evolution of Instructional Systems Design. However, these are the influences we knew personally in own early journeys and would like to make special mention of here:

People: John Barson, Dale Brethower, Les Briggs, Don Ely, Gene Faris, Jim Finn, Bob Gagné, Joe Harless, Roger Kaufman, Bob Mager, Sue Markle, Dave Merrill, Tom Schwen, and Len Silvern

Institutions: Brigham Young University, Florida State University, Indiana University, Michigan State University, Pennsylvania State University, University of Pittsburgh, Syracuse University, University of California Los Angeles, University of Southern California; IBM, Northrup Grumman Corporation, and the U.S. Air Force

Journals and Workshops: Audiovisual Communication Review, Audiovisual Instruction, Educational Technology, Journal of Instructional Development; Criterion Referenced Instruction, and Instructional Development Institute

Without them, the excellent authors of this volume would never have found this field, nor written such brilliant chapters.

Ken Silber
Rob Foshay
September 2009

Handbook of Improving Performance in the Workplace

Volume One

Instructional Design and Training Delivery

Edited by

Kenneth H. Silber and Wellesley R. Foshay

Co-Published by the International Society for Performance Improvement

International Society for Performance Improvement
WHERE KNOWLEDGE BECOMES KNOW-HOW

Pfeiffer
A Wiley Imprint
www.pfeiffer.com

Copyright © 2010 by International Society for Performance Improvement. All Rights Reserved.

Published by Pfeiffer
An Imprint of Wiley
989 Market Street, San Francisco, CA 94103-1741 www.pfeiffer.com

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400, fax 978-646-8600, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, 201-748-6011, fax 201-748-6008, or online at http://www.wiley.com/go/permissions.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

Readers should be aware that Internet websites offered as citations and/or sources for further information may have changed or disappeared between the time this was written and when it is read.

For additional copies/bulk purchases of this book in the U.S. please contact 800-274-4434.

Pfeiffer books and products are available through most bookstores. To contact Pfeiffer directly call our Customer Care Department within the U.S. at 800-274-4434, outside the U.S. at 317-572-3985, fax 317-572-4002, or visit www.pfeiffer.com.

Pfeiffer also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Cataloging-in-Publication Data
Handbook of improving performance in the workplace.
p. cm.
''Co-Published by the International Society for Performance Improvement.''
Includes bibliographical references and index.
ISBN 978-0-470-19068-5 (v. 1 : cloth) – ISBN 978-0-470-19069-2 (v. 2 : cloth) – ISBN 978-0-470-19067-8 (v. 3 : cloth)
1. Performance technology. 2. Employees–Training of. I. International Society for Performance Improvement.
HF5549.5.P37H364 2010
658.3'14–dc22

2009026946

Acquiring Editor: Matthew Davis
Marketing Manager: Brian Grimm
Production Editor: Michael Kay
Editor: Rebecca Taff
Indexer: Sylvia Coates
Editorial Assistant: Lindsay Morton
Manufacturing Supervisor: Becky Morgan

Printed in the United States of America
Printing 10 9 8 7 6 5 4 3 2 1

CONTENTS

LIST OF EXHIBITS, FIGURES, AND TABLES

EXHIBITS

FIGURES

TABLES

INTRODUCTION TO VOLUME ONE

This introduction is designed to set the tone, scope, and philosophy for Volume One of the *Handbook of Improving Performance in the Workplace*. It will address the following topics:

- Goal of the volume
- Ways to read and use this book
- How instructional design (ID) relates to human performance technology
- How this handbook relates to the other handbooks in this series
- The research-based philosophy of the book (This topic will be covered in more detail in our Introduction to Part Three.)
- How our views about ID have changed
- New perspectives on established practice

GOAL OF THE VOLUME

The goal of this volume is to help you become an informed practitioner of instructional design (ID).* In publishing this series, *Handbook of Improving*

*Although this field is called by many different names, we will consistently use the term ''instructional design'' and the abbreviation ID. Those interested in the name controversy can find a detailed history in Chapter Three.

Performance in the Workplace, ISPI and Pfeiffer have partnered with the intent of:

> Underscoring the importance of the link between training design and delivery and improving workplace performance. As most books available for the field provide theories, models, or one individual's process or method, this series stands alone as a rigorous, research-based approach to bring together the body of knowledge, standard principles, and evidence-based best practices from several related disciplines and fields of practice.

This one-of-a-kind set collects best-in-field researchers, academicians, thinkers, and practitioners together across several professions, disciplines, and geographical boundaries under the banner of the International Society for Performance Improvement's ''performance landscape'' to provide a unified and comprehensive compendium of standard principles and best practices for core topics.

The purpose of this set of volumes is to provide training and performance improvement professionals and students with comprehensive and authoritative references on three core areas of study and practice. Sponsored by ISPI, the leading international association dedicated to improving productivity and performance in the workplace, each volume represents the most current research, knowledge, and practice in the field and covers established theories, cutting-edge research, fresh developments, and proven applications.

Donald Schön (Capella, 2006) describes the type of outcome we would like for this volume—for readers to become ''practitioner-scholars'' who emphasize:

> Application of theory and knowledge to real-world problems. As learners become **practitioner-scholars**, they develop systems and strategies for analyzing and resolving problems; they build their ability to synthesize theory and application, with an emphasis on application. As Schön notes, ''[students] tend to think differently about the theories offered by researchers when they realize that they hold comparable tacit theories of their own'' (1987, p. 324). Practitioner-scholars analyze the theories of researchers and compare them to their own knowledge and experience. (Capella, 2006, p. 8)

We hope that you, the reader, will gain ''how to'' knowledge from every chapter and also come to understand ''why'' you are doing what you are.

The editors and authors have balanced the material in all chapters so that this volume is useful to you as:

- A ''how do I do X in ID'' reference—no matter what your ID question, this volume contains techniques to use, supported by forms and jobs aids and examples.

- A ''why am I doing this'' professional development tool—providing a way to understand the theory and research behind what you do so that you can generalize the information to new situations and problems you face.

WAYS TO READ AND USE THIS BOOK

This volume is offered both in a print version and as a chapter-by-chapter download online. This means you may be reading this introduction in print or online, and before or after reading other parts of the book. We will focus here on some suggested ways to approach the chapters in this volume.

As a Long-Time, Expert ID Practitioner

Those of you in this role are probably aware more than other readers that the ID field does not stand still. The techniques you learned in graduate school or workshops and have been practicing for five or ten years (or more, as we have), while still somewhat valid, have grown and changed. As we point out in this introduction, the field has moved from being a process to being a set of principles, the focus has moved from concentrating on algorithmic problem solving and single concepts to high-level problem solving and knowledge structures, the instructional strategy mix has grown, and the learning and research base of the field has changed dramatically.

We recommend that you focus on the following chapters:

- Part One, Chapter Two: A Principle-Based Model of Instructional Design to give you a new view of what you learned as ADDIE
- Part Two, Chapter Seven: Cognitive Task Analysis
- Part Three, Chapters Eight, Nine, Ten, Eleven for sure, and Twelve and Thirteen if your organization is moving to online delivery and problem-solving training. Also read Chapter Fourteen if you do psychomotor training.
- Part Five, Chapter Eighteen, without a doubt, and Seventeen and Nineteen, depending on your role in the organization

Looking for Techniques to Do a Specific Part of ID

If you are regularly doing ID on your job, you will probably approach this volume by turning first, or only, to the chapter(s) that address the ID question you have at the moment. This approach is in keeping with just-in-time training concepts, and the volume is designed to make such use both easy and meaningful.

As you read a chapter, you will notice, both in the chapter and in our introduction to the chapter, references to other related chapters in the volume. Since ID is an interconnected set of principles (an idea we will expand upon in a bit), you can be sure that reading those other chapters will help you address your need and expand upon your understanding of the chapter you are currently reading.

Introducing Yourself to the ID Role

If you are working as a trainer or in a line position in an organization and have been newly assigned to the ID role, we would suggest you start with the following chapters:

- Part One, Chapter Two: A Principle-Based Model of Instructional Design to help you gain an overview
- Part Two, Chapters Four, Five, Six, and Seven to help you understand the importance of, and techniques for doing, analysis before you design
- Part Three, Chapters Eight through Fourteen to help you understand the most common instructional strategies instructional designers use in the situations you are likely to find yourself in as a beginner
- Part Five, Chapter Eighteen to help you understand the collaborative nature of the ID process

Learning ID as a Graduate Student (or Teaching ID to Graduate Students)

If you are learning ID for the first time in a university setting (or helping others do so), you are probably going to be more interested in theory and research as well as practice, as you strive to become the ''scholar practitioner.'' We realize that the nineteen chapters in this volume are too many for anyone to try to ''learn'' in one semester—or, worse yet, in one quarter. (See Chapters Nine and Ten about cognitive load.)

Therefore we would suggest the following chapters and sequence:

- Part One, Chapters One, Two, and Three to provide a theoretical, principle, and historical foundation of the field
- Part Two, Chapters Four, Five, and Six to be able to do the basic kinds of analysis required of entry-level IDs
- Part Three, Chapters Eight, Nine, and Ten (and Chapter Eleven for the overachievers) to learn what the research says about the instructional strategies you will be using in most of the ID work you do in courses and as a beginner in the field

- Part Four, Chapters Fifteen and Sixteen to be clear both on how ID views assessment of learners, and how it confirms that the lessons that you produce actually result in learning
- Part Five, Chapter Eighteen to understand the collaborative nature of ID

HOW ID RELATES TO HPT

This handbook is part of a series co-published with the International Society for Performance Improvement (ISPI). ISPI's focus is human performance technology (HPT), which is defined in the third edition of the *Handbook of Human Performance Technology* as:

> Human performance technology is the study and ethical practice of improving productivity in organizations by designing and developing effective interventions that are results-oriented, comprehensive and systemic. (Pershing, 2006, p. 6)

It is now generally agreed, even by the Association for Educational Communications and Technology, that Educational Technology is a subset of HPT:

> As is advocated in the related field of human performance technology (HPT), there are many different sorts of interventions that may be used in the workplace to improve performance, such as tools, incentives, organizational change, cognitive support, and job redesign, in addition to instruction (Pershing, 2006). Since it encompasses all these sorts of interventions, HPT is a broader concept than educational technology. (Januszewski & Molenda, 2008, p. 7)

It is also generally agreed (Silber, 2007) that ID is a subset of educational technology.

Visually this relationship can be represented as concentric circles, as shown in Figure I.1.

So ID is part of human performance technology. It is the part that is called into play when a performance analysis (see Chapter Five) determines that the cause of the performance problem we are trying to solve is ***lack of skills/knowledge on the part of the learner.***

For those interested in the separate, but intertwined, histories of the fields of ID and HPT, we recommend reading not only Molenda's excellent history of ID in Chapter Three, but also Ferond's (2006) history of HPT, Chapter Seven in the third edition of the *Handbook of Human Performance Technology* (Pershing, 2006).

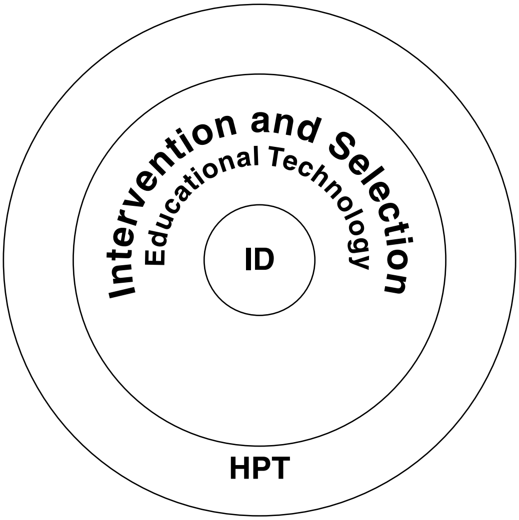


Figure I.1 ID in Relation to HPT

HOW THIS HANDBOOK RELATES TO THE OTHER TWO HANDBOOKS IN THIS SERIES

The intent in publishing three handbooks was to cover the following areas:

- ***Volume One: Instructional Design and Training Delivery.*** This volume covers instructional models and application; alternatives to instructional solutions; analyzing needs, learners, work settings; establishing performance objectives; delivering instruction effectively; managing ID projects; and so forth.
- ***Volume Two: Intervention Selection and Implementation.*** This volume addresses interventions at work, worker, and workplace levels (games and simulation, distance training, mentoring, motivation, performance support systems, OD, knowledge management, communities of practice, and so on). We address what to do to make interventions stick, to create mechanisms for follow-through, to bring about ownership, and other topics.
- ***Volume Three: Measurement and Evaluation.*** Task analysis, performance analysis, instructional analysis, content analysis, needs analysis; qualitative/quantitative methods, observation methods; models for measurement and evaluation, ROI, BSC, and other topics.

Even the casual reader will note that this handbook has a section entitled Evaluation, even though there is an entire handbook devoted to the topic of evaluation. The two chapters in Part Four of this handbook cover the basics of evaluation that an instructional designer must employ: assessing learning at the end of the lesson, and conducting formative, summative, and confirmative evaluation to demonstrate that the lesson that has been designed actually is effective in producing results. Volume Three: Measurement and Evaluation goes into the latter area and into many other aspects of evaluation in much greater depth.

THE RESEARCH-BASED PHILOSOPHY OF THE BOOK

Every trainer acts on a set of beliefs, often not well articulated, about how people learn and how best to teach them. Often, these beliefs are based on personal experience, and whatever skill one has developed in training comes from following the model of admired practitioners. It is fairly rare for practitioners to base their work on a coherent, theoretically sound body of knowledge about learning and instruction. Unfortunately, many popular theories and models of instruction are based only weakly on research. The field is replete with the trendy and the clever, often based more on ideology and personal experience than on sound empirical evidence painstakingly accumulated from methodologically sound research.

As Molenda points out in Chapter Three, from its inception the premise of the ID field has been to build a systematic technology (methodology) for design of instruction through careful application of a coherent body of theory grounded in available empirical research on learning and instruction. The emphasis of the ID field has been on cumulative knowledge building and on the incorporation of new knowledge into the existing practice of effective design. This means that the field tends to resist the ''trendy,'' particularly when advocates of new media or instructional practices fail to recognize how this cumulative knowledge of effective learning and instruction design principles applies to them. Thus, this book does not include chapters on design of instruction for each flavor of e-learning, platform instruction, distance learning, or others or chapters on favorite training techniques such as role play, discovery learning, small group/collaborative learning, and so forth. Our underlying premise is that the same principles of learning and instruction apply across all media and all teaching practices and all learning contexts and all learners (Merrill, Drake, Lacy, Pratt, & the ID2 Research Group, 1996). Design of learning environments using any combination of media and techniques is based on these principles.

We do not mean to argue the naïve thesis that ID practice follows directly, and solely, from empirical research on learning and instruction. ID is a field of practice, with its own design practices—a praxeology (Schön, 1984, 1987; Silber, 2007). The praxeology of ID is informed and guided by research, but it also is guided by reflective practice. Whenever we examine the effectiveness of our work products and reflect on the success of our design practices, we are conducting a design experiment (Merrill, Drake, Lacy, Pratt, & the ID2 Research Group, 1996). These design practices change and evolve through this kind of reflective practice. While our praxeology is grounded in the research-based theory of learning and instruction, it is tested and validated through our experience. It is the intent of this book to present the praxeology of instructional design, and to make explicit the ways in which that praxeology is supported and informed by research and theory, and validated by practice.

HOW OUR VIEWS ABOUT ID HAVE CHANGED

As Molenda explains in Chapter Three, the theory and practice of ID has evolved. The field has advanced by drawing on experience of practitioners, advances in cognitive theory of learning and instruction, and advances in the theory of design and in models of project management. More specifically:

- Our design experience has led to development of a great many techniques for analysis, design, quality management, evaluation, and efficiency of ID. Our sense is that exemplary practice in this field is both more efficient and more effective, across a wider range of needs, than was true at the field's inception.
- Progress in the technologies of information and communication has led to new ways for design teams to work efficiently and has given us a great many more strategies for learners to use in gaining information, and in communicating and collaborating with each other (and with the design team). As a result, consistent with the principles of HPT, we have learned how to more efficiently and effectively target our training.
- Advances in the theory of cognitive learning and instruction (described in Chapters Two, Eight, Nine, and Ten) have led to many insights about how to build deep understanding and expertise in learners, across both low-level knowledge and procedural skills and high-level problem solving and strategic thinking skills. As a result, we now know much more about how to build the highest-value capabilities in the workforce.
- Advances in the theory of design have moved ID away from process-based models, toward principle-based decision making (as described by Silber in Chapter Two). This is a paradigm shift (discussed by Molenda in Chapter Three). At its inception, ID sought to develop design *processes* (typically variations on analysis-design-development-implementation-evaluation or ADDIE) that, if followed meticulously, would lead to development of

effective and efficient instruction. This paradigm has been largely replaced by an emphasis on design principles that may come to bear at any point in the process. Significantly, this paradigm of design recognizes that information and project parameters are discovered gradually over the course of the project, and thus data collection is a continuous part of design, not just a confirmatory afterthought.

- Correspondingly, ID has followed the lead of other design-oriented fields, such as architecture and software engineering, in replacing the linear ''water fall'' work flow process with a non-linear, iterative process that emphasizes successive approximation toward the goal, early and heavy involvement of learners and stakeholders in the design, and a multi-specialty collaborative team structure.

In sum, master ID practitioners with current expertise now know how to work much more effectively, with reduced risk and more targeted effort, thus leading to greater efficiency than was the case a generation ago. We know how to address a much broader range of training needs than we did a generation ago.

NEW PERSPECTIVES ON ESTABLISHED PRACTICE

In preparing this handbook, the emphasis on recent advances in ID has not led us to ''throw the baby out with the bathwater.'' Much of established practice is still valid, although our understanding of even these methods has evolved in response to the advances described above. Thus, in this handbook, you will find:

- A discussion by Anderson (Chapter Four) of the basic methods of data collection and analysis for performance, needs, and task analysis and formative evaluation. This skill was once a part of every ID's repertoire, but, with the loss of the book by Zemke & Kramlinger (1988), this information is presented here for the first time in many years.
- Chapters by Burner (Chapter Five) and by Bell, Andrews, and Wulfeck (Chapter Six) on the ''classical'' methods of front-end analysis and task analysis, both updated to explain how these methods now reflect current cognitive theory and design.
- Chapters by Ruth Clark and Richard Mayer (Chapters Nine and Ten) on the basics of instructional strategies for receptive and directive learning environments in all media. The state of the art for learning environments has evolved substantially, based on current research and theory, and experienced designers are likely to discover that old, familiar design practices are now in need of an update.

- A chapter on Level Two Certification by Shrock and Coscarelli (Chapter Fifteen), which reframes classic competency-based test design in terms of the new paradigms of current theory of learning, instruction, and design.
- A chapter on the role of evaluation in ID by Ross and Morrison (Chapter Sixteen), which updates the classic concept of formative evaluation to reflect the current thinking on design.
- Two chapters on various aspects of the management of ID: project management (Dobson, Dobson, Leemann, and Forsberg, Chapter Seventeen) and management of ID in the training organization (Maitre and Smith's Chapter Nineteen), which are among the very few chapters in a volume on ID to document the management of ID, usually learned only through apprenticeship.
- A chapter with a discussion by Ranshaw (Chapter Eighteen) of the ''human'' side of ID, how to work with a subject-matter expert (SME). This essential ID skill often is learned and passed on only by apprenticeship. It has rarely been documented in a volume of this sort.

CONCLUSION

The thoughtful reader cannot help but conclude that this is an exciting time in the evolution of the instructional design field. Taken together with our focus on measured learning outcomes and performance improvement, HPT-aware ID professionals who practice the state of the art described in this handbook will bring to their clients an approach to high-value, targeted, and cost-effective training that historically has eluded much of the training field.

References

Capella University. (2006). *Capella University's scholar-practitioner educational philosophy, volume 4.0*. Retrieved from www.capella.edu, 11/24/08.

Ferond, C. (2006). The origins and evaluations of human performance technology. In J. Pershing (Ed.), *Handbook of human performance technology* (3rd ed.). San Francisco: Pfeiffer.

Januszewski, A., & Molenda, M. (Eds.). (2008). *Educational technology: A definition with commentary*. Mahwah, NJ: Lawrence Erlbaum Associates.

Merrill, M. D., Drake, L., Lacy, M. J., Pratt, J., & the ID2 Research Group (1996). Reclaiming instructional design. *Educational Technology*, *36*(5), 5–7.

Pershing, J. (Ed.). (2006). *Handbook of human performance technology* (3rd ed.). San Francisco: Pfeiffer.

Schön, D. (1984). *The reflective practitioner*. New York: Basic Books.

Schön, D. (1987). *Educating the reflective practitioner*. San Francisco: Jossey-Bass.

Silber, K. H. (2007). A principle-based model of instructional design: A new way of thinking about and teaching ID. *Educational Technology, XLVII (5)*, 5–19.

Zemke, R., & Kramlinger, T. (1988). *Figuring things out*. Reading, MA: Addison-Wesley.

PART ONE

FOUNDATIONS

This is the part of the book you are most likely to want to skip, either because you think you already know the material or because you don't see any reason to know it.

We suggest that, when you read this part, you are in for some great surprises. Did you know that:

- Cognitivism, not constructivism or behaviorism, is the learning theory that underlies most current ID.
- Inductive learning and learning styles, while popular, are not supported by most of the research as being effective.
- ID is not a process but a set of principles.
- ADDIE is dead.
- The seeds of ID were sewn in the early 1960s.
- Southern California was a hotbed of ID activity and ideas in the 1960s and 1970s.

This part of the book immediately immerses you in the major themes of the volume, with its review of learning theory, its presentation of a new model of ID, and the most comprehensive history of ID yet written.

Chapter One: Some Principles Underlying the Cognitive Approach to Instructional Design. The Foshay chapter on learning theory, based on a book by Foshay, Silber, and Stelnicki, highlights new cognitive research on how those who have done research on learning now believe people learn. It contains new, surprising, and useful findings on short- and long-term memory, cognitive load, encoding, and retrieval. It explains the difference between declarative and procedural knowledge and focuses on the importance of problem solving as the key learning challenge of today. It contrasts those with the older notions of behaviorism and explains why we now use these concepts as the basis for ID. The research findings discussed in this chapter will form the basis for much of what you will read in Chapters Eight, Nine, Ten, and Eleven, which assumes you are familiar with the basics of learning.

Chapter Two: A Principle-Based Model of Instructional Design. Picking up on much of the research about problem solving and combining it with research from other fields (especially architecture) about how people who design actually do their work, Silber (citing the work of many others who have advanced this notion) suggests that the ADDIE process model of ID is dead. In fact, that it was never how experts did ID at all. He goes on to explain how experts use principles, rather than procedures, to form the mental models that were discussed in Chapter One, and how those principles are put together by experts to solve problems. This insight, while relatively new to the ID field, fits with the way most other design fields now view their work. He ends by presenting a principle-based model of ID that previews many of the principles that will be discussed in the rest of this volume.

Chapter Three: Origins and Evolution of Instructional Systems Design. How did ID begin? Who are its founders? How far have we come in the more than forty-five years (yes, it's been that long!) that the field of ID has been around? What were the predecessors of ID (including systems and programmed instruction), and did they morph into the ID that was discussed in Chapter Two? Who were the leaders of early ID who helped take it from its infancy to the robust field it is today? What was the first ID model (hint, it was not ADDIE), and how did it differ from what we do today? In his chapter, Molenda presents the most comprehensive, extensive, exhaustive history of the field of ID ever written, incorporating and adding to all existing histories of the field. His approach provides a new perspective on where we came from and who our founders were.

CHAPTER ONE

Some Principles Underlying the Cognitive Approach to Instructional Design*

Wellesley R. Foshay

In the generation since the birth of the instructional design field, our understanding of the basic psychological mechanisms of memory, perception, learning, and problem solving has seen a great deal of development. Corresponding progress in our understanding of the psychology of instruction (or, if you prefer, design of learning environments) has led to important new definitions of principles of instructional design. For those familiar with the behavioral approach, this chapter will review what you already know and show how the cognitive approach differs. For those who have never had a formal study of the assumptions underlying the behavioral approach, this chapter will provide you with a theoretical understanding of the approach you probably have been using to date. Important additional principles are included in the chapters in Part Three. However, a full discussion of the psychology of learning and instruction is beyond the scope of this chapter and of this handbook. If you are interested in pursuing the subject matter further, references to sources from which this chapter is drawn are provided.

We do not mean to imply a disjunctive contrast between the behavioral and the cognitive approach, nor do we mean to imply that behavioral principles are obsolete—only that the cognitive approach often adds prescriptive utility to

* Portions of this chapter are adapted from Wellesley R. Foshay, Kenneth Silber, and Michael Stelnicki (2003), *Writing Training That Works: How to Train* Anyone *to Do Anything: A Practical Guide for Trainers Based on Current Cognitive Psychology and Instructional Design Theory and Research*. San Francisco: Pfeiffer.

our practice over a wide range of training needs. Few instructional designers follow a purely behavioral or cognitive approach to design. Furthermore, in many cases the behavioral approach and the cognitive approach lead to similar design solutions. Therefore you may find that you are already using some principles of the cognitive approach in designing your instruction.

HOW THE BEHAVIORAL APPROACH IS DIFFERENT FROM THE COGNITIVE APPROACH

Generally speaking, behaviorism is a set of principles concerning both human and non-human behavior. One major behaviorist goal is to explain and predict observable behavior. Behaviorists define learning as the acquisition of new behavior as evidenced by changes in overt behavior. Behaviorism draws conclusions about behavior from research on external events: stimuli, effects, responses, learning history, and reinforcement. These behaviors are studied and observed in the environment and are explained with little or no reference to internal mental processing.

In dramatic contrast to behaviorism, a major tenet of cognitive psychology is that internal thought processes cause behavior. It is their understanding that can best explain human behavior. Cognitive information processing psychologists consider learning to be mental operations that include internally attending to (perceiving), encoding and structuring, and storing incoming information. Cognitive psychologists interpret external stimuli in terms of the way they are processed. They use observable behavior to make inferences about the mind. Furthermore, exciting new work in cognitive neuroscience is relating the structure of the brain to its function, and in the process, validating and elaborating on the accounts of processing and memory induced experimentally by the cognitive psychologists.

The difference in focus between the behaviorist and cognitive theories has important implications for instructional designers who seek design principles based on theory. The biggest differences are in these theoretical areas:

- What learning is
- Factors influencing learning
- The role of memory and prior knowledge
- How transfer occurs
- The goal of instruction
- The structure of instruction
- Specific instructional strategies

Different types of learning are best explained by each approach, and each approach provides basic principles that guide instructional design in different circumstances.

What the implications are for each of the above areas and how they differ in each of the two approaches are shown in Table 1.1. It is important to note that some of the differences are merely semantic (for example, ''fluency'' and ''automaticity'' both describe degrees of learning proficiency), while some are more substantive. For example, ''emphasis on knowledge structures'' reflects the cognitive theory's recognition of the need to think about the parts of knowledge in any given subject and how they fit together.

Table 1.1 Differences Between Behavioral and Cognitive Approaches

Instructional Design Area	*Behavioral Approach*	*Cognitive Approach*
What learning is	''changes in form or frequency of observable performance''; what learners ***do***	internal coding and structuring of new information by the learner; discrete changes in knowledge structures; what learners ***know*** and ***how*** they come to know it
Factors that influence learning	''arrangement of stimuli and consequences in the environment''; reinforcement history; fluency in responding	how learners attend to, organize, code, store, retrieve information as influenced by the context in which information is presented when it is learned and when it is used; thoughts, beliefs, attitudes and values; automatic responding
The role of memory	not addressed in detail; function of the person's reinforcement history; forgetting results from lack of use	''learning occurs when information is stored in memory in a meaningful manner so it can be retrieved when needed''; ''forgetting is the inability to retrieve information from memory because of interference, memory loss,

(*Continued*)

Table 1.1 (*Continued*)

Instructional Design Area	*Behavioral Approach*	*Cognitive Approach*
		or inadequate cues to access the information'' given the way it is organized in memory; therefore, meaningfulness of learning directly affects forgetting
How transfer occurs	focus on design of the environment; stimulus and response generalization to new situations	stress on efficient processing strategies to optimize cognitive load; function of how information is indexed and stored in memory based on expected use of the knowledge; applying knowledge in different contexts by reasoning analogically from previous experiences; construction/ manipulation of mental models made up of networks of concepts and principles; learners believe knowledge is or will be useful in new situation
What types of learning are best explained by the approach	discriminations (recalling facts); generalizations (defining and illustrating concepts); associations (applying explanations); chaining (automatically performing a specified procedure)	``complex forms of learning (reasoning, problem solving, especially in ill-structured situations)''; generalization of complex forms of learning to new situations
What basic principles of the approach are relevant to ID	produce observable, measurable outcomes => task analysis, behavioral objectives, criterion-referenced testing; existing	All of the behavioral principles, and: student's existing mental structures => learner analysis; guide and support for accurate

	response repertoire and appropriate reinforcers => learner analysis; mastery of early steps before progressing to complex performance => simple to complex sequencing; practice; mastery learning; reinforcement => practice. followed by immediate feedback and rewards; use of cues and shaping => prompting, fading, sequencing	mental connections => feedback; learner involvement in the learning process => learner control; metacognitive training; collaborative learning; identify relationships among concepts/principles to be learned, and between them and learners' existing mental models => learner analysis; cognitive task analysis; emphasis on structuring, organizing and sequencing information for optimal processing => advance organizers, outlining, summaries; connections with existing knowledge structures through reflective processing => analogies, relevant examples, metaphors
Goal of instruction	elicit desired response from learner presented with target stimulus	make knowledge meaningful and help learners organize and relate new information to existing knowledge in memory
How should instruction be structured	determine which cues can elicit the desired responses; arrange practice situations in which prompts are paired with target stimuli that will elicit responses on the job; arrange environmental conditions so students can make correct responses in the presence of target stimuli and receive reinforcement	determine how learners' existing knowledge is organized; determine how to structure new information to mesh with learners' current knowledge structure(s); connect new information with existing in meaningful way through analogies, framing, outlines, mnemonics, advance organizers; arrange practice with structurally meaningful

(*Continued*)

Table 1.1 (*Continued*)

Instructional Design Area	*Behavioral Approach*	*Cognitive Approach*
		feedback so new information is added to learners' existing knowledge
Specific instructional strategies	teach fact lesson first, then concepts, then principles, then problem solving; focus on algorithmic procedures for problem solving, including troubleshooting; teach each concept, procedural chain, troubleshooting approach separately; when mastered go on to next; focus on deductive learning; present principles and attributes; build generalization with extended realistic practice, often after initial acquisition	teach problem solving in authentic (job) context; teach principles, concepts, and facts in context as appropriate within the problem-solving lesson; focus on heuristic problem solving and generalization, even in troubleshooting; teach overall mental model, then use coordinate concept, principle, procedure/problem solving teaching to teach all related knowledge at or near the same time; focus on inductive learning; present examples; build generalization through practice in additional problems and contexts which require similar but not identical problem-solving procedures

Based on J. R. Anderson, 1995a; J. R. Anderson, 1995b; Ertmer & Newby, 1993; Fleming & Bednar, 1993; Foshay, 1991; Hannafin & Hooper, 1993; Silber, 1998; West, Farmer, & Wolff, 1991

WHY THE COGNITIVE APPROACH TO INSTRUCTIONAL DESIGN IS IMPORTANT

The cognitive approach to ID has become prominent in the past two decades for two reasons, one based in the theory of learning and instructional design, the other based in business. From the perspective of theory, the cognitive approach seeks to overcome a number of limitations of the behavioral approach. For example, with the behavioral approach to ID:

- Learners sometimes have trouble transferring what they have learned from training to the job;
- Learners can have trouble attaining expert-level performance in troubleshooting and problem solving on the job;
- Learners often have trouble generalizing their training from one situation to another, leading to skill gaps every time the job, content, or technology changes, and creating the need for retraining;
- Learners may have difficulty with divergent reasoning (many right answers or many ways to get to the answer), as opposed to convergent reasoning (one right answer and one way to get it); and
- Designers do not have adequate prescriptions for designing the kinds of training we are now being asked to design—problem solving, troubleshooting (especially in settings where content volatility is high), design, heuristic-based thinking (using guidelines versus algorithmic thinking, which uses formulas with 100 percent predictable outcomes), strategic thinking, and the like.

From the perspective of business, the current behavioral approach to ID sometimes leads to excessive development and delivery costs because it requires:

- Longer training sessions, to cover all the specific algorithms or other content variations;
- More retraining time, to address lack of transfer to new situations; and
- More development time, because there are no guidelines for creating training for higher-order thinking, developers must either guess, or treat problem solving as a large number of low-level procedures and concepts.

The cognitive approach to ID offers remedies to these problems. It provides designers with another way to design training that works well in situations in which higher-order thinking, problem-solving, and transfer to new situations are training goals.

HOW LEARNING OCCURS ACCORDING TO THE COGNITIVE POINT OF VIEW

There are many theoretical models in cognitive psychology. Although there are important differences among them, they broadly agree on how learning occurs. According to these models, there are several components of the mind, and each is involved in the learning process in certain ways. How each component of the mind works has implications for how we design instruction. The components are

- Perception and memory stores
- Short-term or working memory
- Long-term memory

Perception and Memory Stores

Perception Is Selective. There is more stimulation in the environment than we are capable of attending to, and then encoding (internally translating) for storage in memory. Therefore, we only attend to certain things. We attend to and see/hear what we expect to see in a given situation. We attend to those things that interest us because they are either (a) related to what we already know or (b) so novel they force us to attend to them.

Limits of the Sensory Stores. Our sensory stores, also called sensory memories (analogous to a computer's ''buffers''), are capable of storing almost complete records of what we attend to. The catch is they hold those records *very briefly*. During that very brief time before the record decays, we do one of two things: (1) we note the relationships among the elements in the record and encode it into a more permanent memory or (2) we lose the record forever.

ID Implications. The implications of the selectiveness of perception and limitations of sensory stores for instructional designers are that it is crucial to:

- Get the learner to *attend to* the parts of the environment that are crucial (hence the emphasis in the cognitive approach to ID on attention-getting and on motivational statements); and
- Help the learner note *relationships* among the information quickly (hence the importance of organizing the information you are presenting and of clearly relating the new information to existing familiar or important contexts and knowledge).

Short-Term or Working Memory

Controversy. There is disagreement among cognitive psychologists about whether there is a short-term memory that is ''separate and different'' from long-term memory. The disagreement is about whether the two types of memory are physically different, or whether they are just conceptually different constructs. There is also discussion about how they encode information, how they store information, and so on. Regardless of the theoretical differences, some ideas that most psychologists would agree about can affect the design of training.

Rehearsal. When information is passed from the sensory stores to memory, we mentally rehearse it. Examples include repeating phone numbers several times or creating associations to names (for example, **Ted** with the **red** hair) to help memorize them when you first hear them at a party. The former, simply repeating the information over and over, is called *passive rehearsal*. It does not seem to improve memory as well as rehearsing the information in a *deep and meaningful* way, like the latter way of creating associations.

Limited Capacity. There seems to be a limit on the amount of information we can rehearse at one time. A classic paper presented by a Bell Labs psychologist in 1956 showed that we can remember 7 +/− 2 bits of information at most, and that to remember more we have to ‘‘chunk’’ (or group) information in manageable sizes; that’s why your phone number has seven digits, and when area codes became prominent, people were taught to remember phone numbers in three chunks (aaa-bbb-cccc). The findings of this study still seem to apply, with some modifications of how you define a ‘‘bit’’ (element) or a ‘‘chunk’’ (and, as you will see later in Part Four, the ‘‘7 +/− 2’’ estimate is probably too high in many circumstances).

Format. At this point in the learning process, the information being rehearsed is not yet organized and encoded as it will be when it is finally stored in memory. Also, there is some evidence that there are separate spaces for storing and rehearsing verbal information and visual/spatial information, and possibly separate spaces for other types of memories as well.

ID Implications. The implications are that instructional designers need to:

- Help learners use meaningful ways of rehearsing the information, as opposed to simply repeating it (through the use of analogies, by relating new information to existing knowledge or problem situations, etc.);
- Present the information in meaningful ‘‘chunks’’ of appropriate size for the learner population (knowing what your learners already know about the subject they are learning is critical to determining what ‘‘appropriate size’’ for those learners is);
- Present the information in multiple formats (verbal, auditory, visual), which can help learners rehearse, and therefore remember, better; and
- Present the information in a way that allows the learner to move quickly from rehearsing the information to encoding it and integrating (indexing) it with other information into long-term memory.

Long-Term Memory

In general, theorists believe that long-term memory is organized based on context and experience. That means we encode, store, and retrieve information in the way we have used knowledge in the past and expect to use it again in the future. There are several phenomena psychologists agree on about what strengthens the memory process.

Memory Strength. Information in memory has a characteristic called *strength*, which increases with practice. There is a *power law of learning* that governs the relationship between amount of practice and response time or error rates (Strength = Practice to power x). In simple terms, this means that practice increases the strength of learning exponentially (for example, double the practice at least squares the strength of the learned information in memory; triple the practice increases the strength by a factor of nine). Note that other factors, such as meaningfulness, also affect memory strength.

Elaboration. Elaboration means adding information to the information we are trying to learn. The more we elaborate on what we learn through processing, the better we remember it. This is because, as we tie the new information to existing information or as we create other information related to the new information, we create more pathways to get to the new information as we try to remember it.

Chunking. Memories are stored not as individual bits or as long strings of information, but in "chunks," with each chunk containing about seven elements. As explained in the section above on short-term memory, how big an "element" and a "chunk" are differs based on the learner's existing knowledge.

Verbal and Visual Information. It seems we encode verbal and visual information differently in memory. We use a linear code for verbal information, and a spatial code for visual information. We remember visual information very well, especially if we can place a meaningful interpretation on the visuals. In addition, the Gestalt psychology finding that we remember incomplete and strange images better than complete, standard ones still appears to hold true. With verbal information, we remember the meaning of the information, not the exact words.

Associations and Hierarchy. Information is organized in memory, grouped in a set of relationships or structures (for example, hierarchically). Using such a structure makes it easier for us to remember, because there are more related pieces of information activated when we search for information. While you may not remember one specific piece of information in the structure, you may remember the overall structure and some pieces in it, and from that you can remember or infer the missing piece of information. For example, you may not remember all the numbers in the 12 × 12 multiplication tables, but if you remember some key ones (1, 2, 3, and 5 × a number) you can construct the rest.

By comparison with computers, humans can remember far fewer separate pieces of data, but are much better equipped for pattern recognition skills such as analogical reasoning, inference, and comprehension of visual and verbal languages.

ID Implications. The implications are that instructional designers need to:

- Build a lot of meaningful practice into training to increase the probability of retention (for example, the PQ4R method: Preview, Questions, Read, Reflect, Review, Recite);
- Provide learners with information (allowing them to create information that elaborates on the information to be learned);
- Present the information in meaningful ''chunks'' of appropriate size for the learner population (knowing your learners is critical);
- Present the information so it uses the abilities to remember both verbal and visual information, which can increase memory;
- Hierarchically organize the information being presented (to approximate the way information is stored in memory) to increase retention;
- Provide many associations to the information being learned to increase the chances that the information will be retrieved when called for;
- Help learners to organize/index their memories so they have many associations, many retrieval paths, and appropriate structures; and
- Use authentic (real-world) contexts for explanations, examples, and practice, which will help the learners relate what they learn to situations in which they will need to use the knowledge.

CATEGORIES OF KNOWLEDGE: DECLARATIVE AND PROCEDURAL KNOWLEDGE, AND THEIR SUB-TYPES

When they discuss learning, many cognitive psychologists draw distinctions among different categories of knowledge. When you design training, you will probably find it helpful to use these distinctions to help you decide what kind of knowledge you are teaching and how you can best teach that knowledge.

The biggest distinction is between *declarative* and *procedural* knowledge:

- *Declarative* knowledge is knowing *that*.
- *Procedural* knowledge is knowing *how*.

These are examples of *declarative knowledge*:

- Remembering your telephone number;
- Being able to tell the difference between a table and a tray; and
- Stating that for a car engine to run, it must have air, fuel, and electrical current for the ignition.

These are examples of *procedural knowledge*:

- Following a recipe to bake a cake;
- Building a spreadsheet ''from scratch'' using a software package for spreadsheets;
- Fixing the copier so it will stop jamming; and
- Designing a copier that can't jam.

The basic difference between the two types of knowledge is that declarative knowledge tells you *how the world is*, while procedural knowledge tells you *how to do things in the world*.

Trainers who don't understand this distinction often confuse *knowing* and *doing*, and thus make the following kinds of mistakes in designing training:

- They try to teach (and test) procedural knowledge using strategies suited for declarative knowledge;
- They teach declarative knowledge and stop, assuming that the procedural knowledge will naturally follow on its own; and
- They try to teach the procedural knowledge without teaching the associated declarative knowledge.

There are different types of declarative and procedural knowledge. It's important to understand them so that when you plan your instruction, you can use instructional strategies that are appropriate to each type. If you're good at making these distinctions, you may be able to save considerable time and expense in developing and delivering your training while improving its effectiveness. The different types are discussed below, and chapters in Part Three further explain how to teach each of the types of knowledge.

Types of Declarative Knowledge

One common practice is to distinguish three types of declarative knowledge:

- Facts
- Concepts
- Principles and mental models

The discussion below is a synthesis of much that is already familiar and commonly accepted on the topic; it is included here for completeness. The reader will note that these types of declarative knowledge are very similar to the types of learning proposed by Gagné and taught in most basic ID texts (such as Dick & Carey, 2001). To the traditional categories and explanations, however, we have added the notion of mental models, and described their characteristics in slightly different terms to align better with cognitive theory.

Facts. A fact is a simple association among a set of verbal and/or visual propositions. Some examples of facts are

- On a traffic light, red means stop, green means go, and yellow means prepare to stop;
- In 1492 Christopher Columbus sailed from Spain and landed in the Caribbean; he was not the first to do so, nor did he discover America;
- Miller's (Miller, 1956) studies for Bell Labs said the largest number of digits a person could remember easily was seven; and
- The five steps to create a table in MS Word 2003 for Windows are (1) select tables, (2) select number of rows, (3) select number of columns, (4) select line appearance, (5) click OK.

When you know a fact, you have placed it in a structure so you can recall it from memory. Learning facts as part of a structure that will help you recall them in the way you need them is much more efficient than trying to memorize each fact by itself. Simply knowing a fact does *not* mean you can use it in new situations, explain what it means, identify its relationship to other facts, or recall it to do anything with it.

Concepts. A concept is a category of objects, actions, or abstract ideas that you group together with a single name because they share characteristics in common. Some examples of concepts are

- Cars (versus trucks or campers or utility vehicles);
- Jogging (versus running, walking);
- Beautiful sunrises (versus beautiful sunsets, ugly sunrises);
- Justice (versus injustice); or
- Performance improvement (versus training).

When you know a concept, you can classify new objects, actions, or ideas as either in the category or not. People typically learn concepts by remembering the best example of the category they've seen (or imagined). They may or may not be able to articulate a verbal definition. Concepts do not exist in isolation; all concepts have related concepts (parts or kinds, more general, more specific). Items in a given category that do not belong to one concept in the category do belong to another concept in the category.

Principles and Mental Models. *Principles*. A principle is a cause-effect relationship. When you understand a principle, you know how something works. Principles are frequently stated as ''if . . . , then. . . . '' statements. You can demonstrate your understanding of a principle by explaining why something happens or predicting what will happen. For example, you know that:

If you see lightening nearby, you will hear thunder;

If you turn the ignition key in a car, the engine will probably start;

If you rob a bank, you may go to jail;

If you write test items to match instructional objectives, the test will have certain types of validity; and

$E = IR$ (Ohm's law: electrical voltage is the product of current times resistance).

Mental Models. It's also important to know that the three types of declarative knowledge we've talked about so far fit together into structures. These structures are called *mental models*. They are networks of principles along with their supporting concepts and facts stored in a meaningful structure based on (a) the context for which it was created and (b) the past learning and experiences of the learner. For cognitive psychologists, mental models are the key to learning and using knowledge because:

- They tie together all the declarative knowledge in memory;
- They are the structures into which you organize information, put it into memory, retrieve it from memory when needed, and learn by expanding and restructuring existing structures;
- They provide the most meaningful application of declarative knowledge (as adults we rarely spout networks of facts or run around finding new instances of concepts, but we do frequently try to explain how or why things happen or work);
- They form a bridge between declarative knowledge (knowledge about) and procedural knowledge (knowing how) to do procedures (other than rote ones), you have to ''know how the system works,'' that is, have a mental model of the system; and
- The wrong ones (often called *misconceptions*) will actually interfere with performance and further learning.

Therefore, most would argue that for training of adults, the ID must not only teach isolated facts, concepts, and principles, but must also help the learner create the appropriate mental models for optimum structuring of the information learner for storage, retrieval, and application, while guarding against formation (or perpetuation) of misconceptions.

Types of Procedural Knowledge

Procedural knowledge is the ability to string together a series of mental and physical actions to achieve a goal. Procedural knowledge is used to solve problems.

The way ‘‘problem’’ is used in this book may be a new concept for many readers. In the behavioral approach, instructional designers are used to thinking about ‘‘procedures’’ and ‘‘problem solving’’ as two different things—two different levels in a hierarchy such as Gagné’s. In the cognitive approach, and in this book, the tendency is to use ‘‘procedural knowledge’’ and ‘‘problem solving’’ interchangeably, which many might find confusing initially. Because procedural knowledge is used to solve problems, the type of problem the knowledge is used to solve is what leads to the name of the procedural knowledge.

Problem solving always has a starting or initial state (car not running), an end or goal state (running car), a sequence of actions (open door, get in, apply brake, insert key in ignition switch, turn key), and constraints (works only if you have the right key). These are summarized in Table 1.2.

Types of procedural knowledge and problem solving are placed on a continuum:

- At the most precise (procedural) end are well-structured problems;
- At the least precise end are ill-structured problems; and
- In the middle are moderately structured problems.

Table 1.3 summarizes the differences in problem types.

Well-Structured Problem Solving

A term you may sometimes hear for well-defined procedural knowledge is *rote procedure*. We consider performing rote procedures to be well-structured problem solving. All elements of the problem situation are known. The initial state, goal state, and constraints are clearly defined. The operations are also clearly defined, although they may include a choice of alternatives (branches). The learner knows when to start the procedure and when to stop it. Examples of well-structured problem solving include:

- Ringing up a sale in a department store;
- Calculating heating and air-conditioning requirements for a building;

Table 1.2 Problem Characteristics

Definition	*Example*
a. There is an **initial state**, or the elements of the problem the learner is presented with.	You want to record five different TV programs broadcast on five different nights each at a different time.
b. There is a **goal state**, or a description of the situation that would be a solution to the problem.	You need to program the VCR correctly to record the programs.
c. There is a **set of operations** or things the learner can do to get from the initial state to the goal state.	You need to follow the step-by-step programming procedure furnished by your VCR and TV set.
d. There is a **set of constraints** or conditions that must not be violated by the learner in solving the problem (Gagné, 1985). Anderson (1995) uses ‘‘Search’’ instead, as the mechanism of chaining together operations to get from initial to end state.	You must input the correct day, time, and channel for each program in the correct sequence. You must make sure there are no fund drives, presidential press conferences, ‘‘special’’ programs, or any other scheduling changes that would throw off the original times. You also have to make sure you’re correctly specifying a.m. and p.m., correctly associated network name and channel number, and so on.

- Implementing a design for a database; and
- Printing marketing pieces.

Well-structured problems are usually performed simply by recalling procedures and performing them exactly as taught. It’s not even necessary to understand

Table 1.3 How Problem Elements Differ in These Classes of Problems

	Well Structured	*Moderately Structured*	*Ill Structured*
Initial state	Clearly defined	Perhaps known	Not clear or spelled out
Goal state	Clearly defined	Clearly defined	Not clear or spelled out
Operations	Clearly defined	Must be created	Not clear or spelled out
Constraints	Clearly defined	May be known	Not clear or spelled out
Example	Access a computer file	Car won’t start	No marketing plan for a new product

why the procedure works. Thus, in many situations it is optional to understand underlying *principles* that explain the *why* of a well-structured procedure.

Moderately Structured Problem Solving

In moderately structured problems, which include troubleshooting, the goal state is clear and the learners might know the initial state and constraints. However, the learners probably have to recall and assemble in a novel way the operations that will get them from the initial state to the goal state, given the constraints. Examples of moderately structured problems include:

- Fixing the cause of a ''mis-ring'' on a sale item in a department store;
- Developing a floor plan for a building;
- Planning how to implement a redesigned work process; and
- Planning a marketing focus group.

Other examples are deciding on the most advantageous retirement package, deciding whether or not to fire an employee, determining whether to repair your old laptop computer or buy a new one, determining whether or not to recommend that an employee seek company-provided counseling.

For moderately structured problems, it is important to understand the principles that underlie them. For example, a manager who wants to figure out how to motivate an employee needs to understand a few basic principles of motivation, if only at the common-sense level.

Troubleshooting (discussed further in Chapter Eleven) is a special ''compound'' case, in which an expert treats unfamiliar and/or complex problems as moderately structured and generates the operations. Some examples include determining the cause/source of a food poisoning outbreak, finding the source of a scraping noise when your car starts, determining why a metal stamping machine damages its stampings on a random basis, and figuring out why your refrigerator defrosts continually or your coffee maker doesn't work.

Ill-Structured Problem Solving

In ill-structured problems, which include most of the complex problems our learners encounter, three or all four of the elements of a problem are either missing completely or are present but not clear. The range of ill-structured problems is discussed further in Chapter Eleven. Examples of ill-structured problems include:

- Deciding on the sale price for an item in a department store;
- Designing a new building;

- Redesigning a work process; and
- Introducing a new product.

Other examples are holding a press conference on a highly controversial issue, conducting a workshop with learners who are highly resistant to learning the content, and designing an artificial pancreas or an acceptable human blood substitute or an automobile that never wears out.

You've probably heard the old saw that ''defining a problem is most of solving it.'' That refers especially to these ill-structured problems.

The view of this book. For purposes of instructional design, in most circumstances there is little difference between moderately and ill-structured problems. Therefore, we will consider only two classes of problems: well-structured and ill-structured. Jonassen's treatment in Chapter Eleven is consistent with this view.

CONCLUSION: TWELVE INTERPRETIVE PRINCIPLES TO APPLY TO INSTRUCTIONAL DESIGN

Figuring out the implications of learning theory for instruction is neither direct nor simple. That's what instructional designers do, and it's what much of this book (especially Part Three) is about. To get you started, however, we will state twelve principles we hope will help you make the inferential leap from learning to instruction. Many of these are elaborated upon in Part Three.

Principle 1: Any Job Task Includes Both Declarative and Procedural Knowledge. It is naïve to think that entry-level jobs require rote procedures, and only higher-level jobs require workers to know both the ''why'' and the ''how.'' If that was ever true, is certainly isn't in today's knowledge economy! Your front-end analysis, and your training, should always include both.

Principle 2: All Knowledge Is Learned in Structures, Which Are Related to the Logic of the Knowledge, What You Already Know, and How You Use the Knowledge. Stated differently, understanding the forest is as important as knowing the trees. You must think structurally about what is to be learned, and so must the learner.

Principle 3: There Are Different Types of Declarative Knowledge, and You Learn Each a Different Way. This insight is one of the cornerstones of instructional design. Your strategies, tools, and templates must vary by knowledge type, which means you must constantly be aware of what knowledge type(s) need to be learned to master a given task you are training.

Principle 4: Concepts and Principles Are Best Learned from a Combination of Examples and Definitions. A common design error we see is to treat examples as an afterthought, to fail to include them, or to include ones that don't work for the learning task and the learner. You and your learner need both examples and definitions, and both must be carefully constructed.

Principle 5: Teach the Knowledge Structure, Not Just the Parts in Isolation. You must help the learner build (or modify) appropriate knowledge structures. This is a separate learning task from understanding the pieces of knowledge separately, and you must provide opportunities for the learner to create and integrate these structures.

Principle 6: Procedural Knowledge Is How to Do. Anything that involves a series of steps is a procedure. Real-world job tasks usually include a combination of embedded procedures.

Principle 7: Procedural Knowledge Varies According to Its Structure. A common design error is to look only for the well-structured procedures, or to treat all procedures as if they were well-structured. You must get the degree of structure right in your analysis, and you must be aware of the degree of structure in your instruction.

Principle 8: Procedural Knowledge Uses Declarative Knowledge. Another common design error is to forget that there are always declarative knowledge components embedded in procedural knowledge—and that you must use appropriate instructional strategies to teach both.

Principle 9: People Solve Problems Inductively, But Only If It's an Unfamiliar Problem. Early work on problem solving sought to identify general (inductive) principles that could solve any problem. These turn out not to be used much by experts, and they usually are not the place to start when teaching novices.

Principle 10: An Expert Problem Solver Knows More and Different Things Than a Novice Does. An expert is not simply a novice after lots of practice. Experts know more, but they also know differently (with different knowledge structures). Therefore you need to understand the level of expertise of your learners and adapt your instruction accordingly.

Principle 11: Experts Know More Domain-Specific Strategies Than Novices Do. Experts know a great many patterns and insights (knowledge structures) specific to their problem domain. You should be looking for these in your cognitive task analysis (see Chapter Seven), and you should help the learner to see, learn, and use them.

Principle 12: Manage Cognitive Load in Training and in Performance. Cognitive load is all about how to prevent the cognitive ''buffers'' from ''overflowing.'' Yet management of cognitive load is still rare in instruction and assessment. See Chapters Nine and Ten for further discussion of how to manage cognitive load.

These principles are meant to help you see some of the most important implications for instructional design of current cognitive learning theory. As you study the chapters in this handbook, we hope these principles will help you build or revise *your* mental model of what learning is, what instruction is for, and how it works to help the learner do the work of learning.

References

Anderson, J. R. (1995). *Cognitive psychology and its implications* (5th ed.). New York: W.H. Freeman.

Anderson, J. R. (1995). *Learning and memory: An integrated approach*. Hoboken, NJ: John Wiley & Sons.

Dick, W., & Carey, L. (2001). *The systematic design of instruction* (5th ed.). New York: HarperCollins.

Ertmer, P. A., & Newby, T. J. (1993). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly*, *6*(4), 50–72.

Fleming, M., & Bednar, A. (1993). Concept-learning principles. In M. L. Fleming & W. H. Levie (Eds.), *Instructional message design: Principles from the behavioral and cognitive sciences* (2nd ed., pp. 233–252). Englewood Cliffs, NJ: Educational Technology Publications.

Foshay, R. (1991, May). Sharpen up your schemata. *Data Training*, pp. 18–25.

Gagné, R. (1985). *The conditions of learning* (4th ed.). New York: Holt, Rinehard and Winston.

Hannafin, M., & Hooper, S. (1993). Learning principles. In M. Fleming & H. Levie (Eds.), *Instructional message design* (2nd ed.). Englewood Cliffs, NJ: Educational Technology Publications.

Miller, G. A. (1956). The magical number seven, plus or minus two—some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Silber, K. H. (1998). The cognitive approach to training development: A practitioner's assessment. *Educational Technology Research and Development*, *46*(4), 58–72.

West, C., Farmer, J., & Wolff, P. (1991). *Instructional design: Implications from cognitive science*. Boston: Allyn & Bacon.

CHAPTER TWO

A Principle-Based Model of Instructional Design*

Kenneth H. Silber

INTRODUCTION

It is time for a new way of looking at instructional design (ID). ID has been written about as a sequential process since at least 1965; each new model just tweaks its basic steps differently. But everyone agrees that experts DON'T really follow the steps of ID models, and students have difficulty learning them. Time to change—to think about ID as a set of *principles* underlying designing instruction. A *principle-based ID model* is easier to learn, faster to use, and transfers better.

THE ARGUMENT

ID has been described as a systematic process or well-structured procedure. It has taken the form of the generic ADDIE (Analysis, Design, Development, Implementation, Evaluation) model in its many variations (Andrews & Goodson, 1980; Gustafson & Branch, 2002; Reigeluth, 1983, 1999). Constructivism has not killed off this model and its variants, though Constructivists claim not to use the model. Even modern attempts to present ID as a non-linear process/procedure, including different shaped models and concepts like rapid

* Excerpted from Silber, K. H. (2007). A Principle-Based Model of Instructional Design: A New Way of Thinking About and Teaching ID. *Educational Technology, XLVII*(5) September–October 2007. 519. Reprinted with Permission.

prototyping (Cennamo & Kalk, 2005), seem to be struggling to explicate the process in a non-linear way.

Clients, ID critics, and students who work in the ''real world'' all balk at ''how long the ID process takes,'' and how IDers stick ''slavishly'' to a linear procedure in developing instruction. Some IDers, like Braden (1996), argue that the basic strength of ID lies in this very linearity. Tessmer and Wedman (1990) argue for a ''layers of necessity'' approach. Dorsey, Goodrum, and Schwen (1997) argue that a rapid prototyping (cyclical) approach to ID is a more appropriate representation of what IDers do. Gustafson (2000) suggests that the combination of linearity and cyclical prototyping better represents how ID is actually practiced (though not how it is *taught*). Cennamo and Kalk (2005) conclude: ''Although classic models such as ADDIE are presented in a linear sequence, instructional design is rarely practiced that way'' (p. 4).

This chapter will argue that all these discussions are about the *wrong* question. They all assume ID is a *process*, and that what matters is how you theorize about it, present it, practice it, and learn it.

ARGUMENTS OF THE CHAPTER

This chapter argues that the real questions change completely if one considers that ID is NOT a systematic procedure at all. *What if ID is moderately structured problem solving instead* (Jonassen, 2004; Newell & Simon, 1972)? And, as such, according to Foshay, Silber, and Stelnicki (2003) and Jonassen (2005), it is based NOT on a procedure/process at all, but rather on the understanding of (1) a set of principles and (2) a way of thinking about design problems. Rowland's (1992, 1993) notion that expert designers use a mental ''template,'' and think a set of questions, is an early statement of this idea. Thiagaragan (2004) suggests that ID is really a set of principles. Lawson, a design guru in the field of architecture, has written extensively about design as a *way of thinking* (Lawson, 2004, 2006). Nelson and Stolterman (2003) look at design even more broadly, defining it as utilizing a ''process of composition, which pulls elements into relationship with one another, forming a functional assembly that can serve the purposes, and intentions, of diverse human populations'' (p. 22). If ID is problem solving, then the *real questions* are about how designers think, and the principles they use—the subject of this chapter, which presents the three arguments that:

- ID, as experts do it, is a problem-solving process, not a procedure, made up of a thinking process and a set of underlying principles.
- The thinking process is similar to the one designers in other fields use.
- ID is a well-known and agreed-on set of principles and heuristics that form the mental model for expert designers.

BOUNDARIES OF THE CHAPTER

There are many authors who have written about ''a new approach to thinking about instructional design'' that goes beyond the scope of the focus of this chapter. This presentation offers a unique look at this issue, without going beyond the scope of the three arguments presented above. Therefore, this chapter will NOT focus on the following issues that are addressed by others also interested in new ways of thinking about ID:

Osguthorpe and Osguthorpe (2007) describe a ''conscience of craft'' of ID. They see it as made up of theoretical and practical knowledge; this present chapter does not make that distinction, using instead the distinction between declarative and procedural knowledge. They include foundational assumptions, which this chapter includes, but also personal beliefs, which this chapter does not. They talk about ID as a professional discipline that results in jobs and includes ID, program evaluation, and educational research; this chapter is focused only on ID, and ID only as a process or a set of principles and heuristics. They talk about four foundational disciplines (historical, psychological, sociological, and philosophical), while this chapter, sympathetic to the inclusion of the sociological, focuses only on the psychological.

Nelson and Stolterman's (2003) *The Design Way* is interested in the ''true'' and the ''ideal'' as well as the ''real'' when thinking about design; we focus only on the real. They are interested in a complex theory that covers all design; we are interested only in the thinking process that designers use. They are interested in seven designerly qualities (competence, confidence, capability, capacity, courage, connection, and character [p. 293]); this chapter is interested, in their terms, in competence (knowing), capability (making), and capacity (learning).

Rowland (2004) is interested in a ''design epistemology'' and addresses questions like ''What is knowledge?'' (p. 41); this chapter has the narrower focus of what a design mental model might look like in the area of actual design. Rowland is also interested in designer core competencies (judgment, creativity, composition, mindfulness, tolerance for ambiguity, positive attitude toward error, bias toward service and responsibility, and systems thinking [p. 46]); this chapter is interested in more specific observable declarative (mental model) and procedural (problem-solving) knowledge of designers.

Hardré, Ge, and Thomas (2005, 2006) have theorized about the differences between novice and expert IDers and have done a landmark study demonstrating that the theorized differences between novices and experts in general do actually manifest themselves in the learning and practice of ID. This chapter will cite some of their theory and research findings as they relate to the teaching of ID, but we are more focused on the characteristics of how expert IDers *solve problems*.

The field of ''learning sciences,'' sometimes discussed among IDers, does not believe in instructional design, finding it to be too linear, rigid, and designer versus learner focused. This chapter is not interested in even addressing their arguments or their questioning of the validity of ID as a field.

Constructivists like Winn, Jonassen, Grabinger, and others have argued that traditional ID does not work in a constructivist paradigm, and have proposed alternative ways of looking at ID. The author is a cognitivist who, like Merrill (2002) and Foshay, Silber, and Stelnicki (2003), bases his ideas on research rather than on a philosophical stance, which only a fraction of IDers follow. Therefore, this chapter will not address the issues and concerns of social negotiation of meaning, ''discovery'' learning, and other terms.

Finally, there are others who have written about detailed strategies for teaching ID (for example, Shambaugh & Magliaro, 2001). While they and others go into detailed instructional strategies for teaching people to design (including designing, modeling, and reflection strategies), this chapter is concerned with instructional strategies only at a macro level—since if ID is ill-structured problem solving, the strategies have already been identified (Foshay, Silber, & Stelnicki, 2003; Merrill, 2002).

THE THREE ARGUMENTS

1. ID, as experts do it, is a problem-solving process, not a procedure, made up of a thinking process and a set of underlying principles. Newell and Simon (1972) suggest that problems, and types of problem solving, range on a continuum from well-structured to ill-structured. Table 2.1 shows the basic differences among well-structured, moderately structured, and ill-structured problems.

Table 2.1 Differences in Problem Types

	Well-Structured Procedures	*Moderately Structured Troubleshooting*	*Ill-Structured Problem Solving*
Beginning State	well defined	perhaps known; perhaps just symptoms known	Not clear or spelled out
Actions	well defined	many possible actions; must be created or selected	Not clear or spelled out
End State	well defined	well defined or not clear	Not clear or spelled out
Constraints	well defined	may be clear	Not clear or spelled out

From Newell and Simon, 1972, as summarized in Foshay, Silber, and Stelnicki, 2003

Jonassen (1997, 2005) has elaborated on the differences in the characteristics between well- and ill-structured problem solving, but his delineation does not contradict the original distinctions. Anyone who has practiced ID would need no lengthy dissertation, I believe, to accept the premise that the problems IDers solve regularly fall into the moderately and ill-structured categories almost exclusively. It seems clear that the call of every ID model for needs analysis to clarify the problem (for example, Rossett, 1987) demonstrates the belief that the problem is *never* presented in a well-structured manner that allows the IDer to merely ''follow the ID process'' and ''solve it.''

Nelson and Stolterman (2003), who use the terms ''wicked problems'' and ''tame problems,'' point out the danger of confusing the two:

> ''By treating a wicked problem as a tame problem, energy and resources are misdirected, resulting in solutions that are not only ineffective, but can actually create more difficulty; because the approach used is an intervention that is, by necessity, inappropriately conceptualized.'' (p. 17)

Lawson (2006), writing about the analogous process of architectural design, says that ''design problems are often not apparent but must be found . . . neither the goal nor the obstacle to achieving that goal is clearly expressed in fact, and the initial expression of design problems may often be quite misleading'' (p. 56). Further, he clearly puts design in the ill-structured category, when he says:

> ''First, it is not clear that in the case of design problems the improved state can be undeniably and accurately identified. There may be an infinite number of states that offer some form of improvement over the current state, and it may not be possible to entirely agree on their relative benefits.'' (p. 19)

It may be true that ID experts tend to see problems presented to them in a range between well- and ill-structured, while ID novices see all problems as well-structured, that is, how IDers see the problems presented to them is one of the discriminators between novices and experts.

Foshay, Silber, and Stelnicki (2003) cite theory and research to support the contention that novices and experts look at problems differently, with novices focusing on surface features of the problem and having more difficulty creating ''problem spaces'' organized in ways that facilitate problem solving.

Hardré, Ge, and Thomas (2005, 2006) present a model that proposes differences between novices and experts in design thinking, practice, and product, and they present research results that demonstrate that novice IDers do indeed approach problems differently from experts. Lawson (2004), citing Kees Dorst, suggest five levels of expertise in design problem solving:

1. ''The novice tends to follow strict rules as laid down probably from instruction.''

2. ''The beginner has moved on slightly and is more sensitive to the situation context and more aware of exceptions to the rules.''
3. ''The competent problem solver . . . works in a substantially different way, being much more selective as to which problem features to attend to and having much more clearly articulated plans of working.''
4. ''The proficient problem solver has . . . acquired enough . . . experience and reflection to accurately recognize important features and make appropriate plans on a frequent basis.''
5. ''The expert recognizes the nature of the situation intuitively and performs actions without the need for conscious mental effort.'' (p. 107)

Lawson's work supports the notion that there are differences in how novices and experts approach problems.

The importance of this distinction is that the characteristics that competent through expert performers exhibit are only necessary if the kinds of problems they solve are ill-structured. And the description of the problem-solving processes they employ clearly does NOT fall into following well-structured procedures. Since the problems that IDers face are NOT well-structured, it seems to follow logically that well-structured problem-solving procedures are NOT the optimal solution strategies for these problems. Solving moderately and well-structured problems, like ID, involves the use NOT of algorithms, but rather the use of (a) mental models that arrange the relevant declarative knowledge in a way that allows for its use in solving the problem, and (b) heuristics (or guidelines) for actions to take (Foshay, Silber, & Stelnicki, 2003). Lawson (2006) explains:

> ''Most designers adopt strategies which are heuristic in nature. The essence of this approach is that it is simultaneously educational and solution seeking. Heuristics strategies do not so much rely upon theoretical first principles as on experience and rules of thumb.'' (pp. 184–185)

In other words, ID involves the use of a set of related principles about (a) what good instruction is, (b) good design practices, and (c) a set of heuristics that guides one through the extensive set of principles to apply them to a particular problem.

2. The thinking process is similar to one designers in other fields use. While instructional design may be unique in the kinds of problems it tries to solve, it is not necessarily the case that the principles and heuristics—the way designers think about problems—are unique. In fact, there is, in my mind (thanks to Dr. Elizabeth Boling of Indiana University for first putting me on this track), a great deal of similarity between the way IDers think about problems and the way designers in general do.

The author who has spent the greatest amount of time in creating theory and conducting research on design is Bryan Lawson, a British professor of architecture. *How Designers Think* (2006) is now in its fourth edition—the first edition was written in 1980; and its companion is *What Designers Know* (2004). Although it is on his work we draw most heavily, this work seems to be ignored in the references of other authors in the ID field. In addition, Nelson and Stolterman's (2003) *The Design Way* provides an even more macro conceptual view of design. Rowland (1992, 1993, 2004) has written a series of articles addressing various aspects of thinking about design in a new way. The goal of this chapter is not to summarize these books and articles, but rather to highlight some key notions from these works that describe how designers in general, including IDers, think.

Though much of the work by Rowland, as mentioned earlier, is beyond the scope of this chapter, his definition of design is a good place to begin to look at design as a *generic set of skills,* rather than ''just an ID model'':

> ''Design is a disciplined inquiry engaged in for the purpose of creating some new thing of practical utility. It involves exploring an ill-defined problem, finding—as well as solving—a problem(s), and specifying ways to effect change. Design is carried out in numerous fields and will vary depending on the designer and on the type of thing that is designed. Designing requires a balance of reason and intuition, an impetus to act, and an ability to reflect on actions taken.'' (Rowland, 1993, p. 80)

Beginning with the notion that design problems are ill-structured, Lawson argues that they ''cannot be comprehensively formulated and that solutions cannot be logically derived from them'' (p. 182). He cites empirical evidence that designers ''use solution rather than problem-focused strategies. That is to say, their emphasis is more on reaching a solution rather than on understanding the problem'' (p. 182). These ideas seem consistent with what expert IDers do. They do just enough analysis to lead to a hypothesis about a solution, which they propose, then do a prototype, see results, and then modify. These ideas seem inconsistent with a linear ADDIE-style model in which novice IDers are frequently accused, by clients, of being trapped in ''analysis paralysis.''

Lawson does an extensive analysis of design problems, design solutions, and the design process itself. His summary principles, and explanations of these principles, are presented in Table 2.2. His analysis of design problems expands on the analysis cited in the previous section, clearly focusing on their ill-structured nature. His discussion of design solutions and process also expands on the analysis cited above, clearly delineating both problems and the process as ill-structured, defining characteristics of the solutions and process, and even leading to the conclusion that the solutions and process are as much a learning

Table 2.2 Summary of Brian Lawson on Design Problems, Solutions, Process

Category	*Principle*	*Explanation*
Design Problems	''Design problems cannot be comprehensively stated'' (p. 120)	Many parts of the problem do not become visible or clear until some prototype solution has been generated
	''Design problems required subjective interpretation'' (p. 120)	The way we understand new ID problems is a function of the mental models we have, based on our prior training and experience
	''Design problems tend to be organized hierarchically'' (p. 121)	Design problems are frequently the symptoms of higher-order problems; there is no ''right level'' at which to attack such problems, but it makes sense to start at the highest level possible
Design Solutions	''There are an inexhaustible number of different solutions'' (p. 121)	There can never be a complete list of all possible solutions to design problems; because they are ill-defined, there is a whole range of acceptable solutions to problems based on optimization
	''There are no optimal solutions to design problems'' (p. 121)	Design almost always involves compromise, and the designer cannot optimize all the requirements at the same time; one optimizes one or two, and sub-optimizes the others
	''Design solutions are often holistic responses'' (p. 122)	The solution does not individually address each aspect of the problem; rather there is one holistic solution that addresses many problem aspects at once
	''Design solutions are a contribution to knowledge''	Once a solution has been formulated and a design completed, they can be studied by others; ''They are to design what hypotheses and theories are to science'' (p. 122)
	''Design solutions are parts of other design problems'' (p. 122)	Most design solutions have some unintended good and bad effects, and, as such, create other design problems to be solved

Category	*Principle*	*Explanation*
Design Process	''The process is endless'' (p. 123)	''There is no way of decision beyond doubt when a design problem has been solved. Designers simply stop designing when they run out of time or when . . . it is not worth pursuing the matter further . . . . one of the skills is knowing when to stop'' (p. 55)
	''There is no infallibly correct process'' (p. 123)	*''In design the solution is not just the logical outcome of the problem, and there is therefore no sequence of operations which will guarantee a result''* (p. 124, emphasis added)
	''The process involves finding . . . problems'' (p. 124)	The designer spends as much time in identifying the problem as in solving it
	''Design inevitably involves subjective value judgment'' (p. 125)	The way the designer addresses issues about the most important problem to solve, and the ways to optimize a solution, are based on the designer's values
	''Design is a prescriptive activity'' (p. 125)	Problems of science are not the same as problems of design, and therefore the process of science (to understand and predict) is not a good model design, whose job is to prescribe and create
	''Designers work in the context of need for action'' (p. 125)	The purpose of design is not design; rather it is to create something that changes the situation or environment; the end, not the process, is the focus

From Lawson, 2006, pp. 121–125

process as a designing one—an idea that fits well with Argyris and Schön's (1996) Type II learning, and with the notion of metacognition.

3. ID is a well-known and agreed-on set of principles and heuristics that form the mental model for expert designers. If, as argued above, ID is not a well-structured procedure, but rather a set of principles and heuristics that expert IDers use to solve ill-structured problems, then the next issue to be addressed is the nature and content of this set of principles and heuristics. A model of such principles and heuristics has been developed (Silber, 2006). The model derives principles and heuristics from standard basic texts in the ID field. It is inclusive rather than exclusive, including principles that come from different philosophies of ID, which may even contradict each other. Wherever possible, the principles are derived from research; where not, they are derived from commonly accepted practices in the ID field. By definition, creating such a model is an ill-structured problem. And, therefore, according to the discussion in the previous section about problem solutions:

- This set of principles is merely one possible solution to the problem.
- This set of principles optimizes some components at the expense of others (e.g., ID management).
- This set of principles is *never complete*.
- This set of principles provides a starting place for others to change and add.
- It will have some unintended positive and negative effects.

Most importantly, to avoid the model's misuse, it is crucial to understand that this set of principles is *not a linear set of principles to be followed completely and slavishly by every IDer in addressing each ID problem*. That would make it as unrepresentative of how IDers behave as is ADDIE. This is a set of principles that IDers keep in mind as they define and solve ID problems. As they work through an ID problem, IDers access these principles, stored in their ID mental models, as needed. They may access one principle or several principles in combination, then act, then access another combination of principles. These principles may lead to questions they ask, to hypotheses they form, to problem-space definitions, to problem solution strategies, or to problem solutions.

But, again, two things IDers should never do are (a) memorize and recite this list of principles and (b) apply all the principles in this sequence every time.

The model is an attempt to represent a mental model of what an expert IDer knows and how an expert IDer approaches an ill-structured ID problem. Since all mental models are unique in both content and structure, based on individual experience and problem context (Foshay, Silber, & Stelnicki, 2003), it is impossible to ''draw'' what each IDer's mental model looks like. Therefore, for convenience, the set of principles and heuristics is presented in an outline form and is organized, after two introductory sections, around the generic ADDIE phase. This organization was chosen based on the cognitive psychology principle that suggests it is easier to learn new information if it is clearly related to existing knowledge (Foshay, Silber, & Stelnicki, 2003), and not because the author believes that such phases or sequence are actually the way ID experts organize these principles in their own mental models. An early version of the model was presented at an ISPI Conference (Silber, 2006). It has been modified and expanded since then. The most current version of the model is shown in Exhibit 2.1.

Exhibit 2.1 The Silber ID Principle Model

I. Underlying Philosophy (Dick & Carey; Morrison, Ross, & Kemp; Seels & Glasgow; Smith & Ragan)

A. Theories

1. Learning

- Behaviorism

  Practice/responding is key to learning

  Reinforcement is key to continuing responses, and different reinforcement schedules produce different rates of responding

  Environment must help learners discriminate stimuli that lead to different responses

  Environment must help learners shape responses

- Cognitivism

  Learning occurs situated in real settings and specific contexts with real problems

  Information is stored in long-term memory in structures called mental models that represent all the relationships among information

  Selective perception and <1 sec time operate in getting information into sensory registers

  Learning is the retrieval, adding to and reshaping, and restoring of mental models

  Limited capacity ($7+/-2$), short time (10 seconds), and relation to existing knowledge determine what information is kept and lost in working memory

  Retrieval of knowledge for use depends on context stored and context presented by problem

- Constructivism

  Learning is collaborative with meaning negotiated

(*Continued*)

Exhibit 2.1 (*Continued*)

Learning/meaning results from personal interpretation of knowledge and experience

Learning is an active process of constructing knowledge

2. Communication
   - Communication takes place between a sender and receiver, each of whom creates a message to send, and interprets the message received
   - The design of the message influences the amount of noise that is included in it, and the way the receiver interprets it
3. Systems
   - Development of instruction follows a systematic process
   - View instruction as a system: all components related and must function together to produce learning
   - Instruction is a system that is part of a larger system of performance improvement which is part of a larger system of education or training or organizational performance

B. Disadvantages

1. No incentive in school settings
2. Question about whether it works with goal-free/problem-based learning environments
3. Time-consuming

C. Advantages

1. Systematic approach for dealing with problems
2. Facilitates design in teams
3. Instruction that is
   - Replicable
   - Efficient
   - Effective
   - Aligned (all elements)
4. Needed for mediated and e-learning
5. Can accommodate multiple learning theories

D. Principles

1. Designer must have clear idea of what learners will learn.

2. Learners can learn from different media; live teacher not required.
3. There are general principles of instruction that work across subject matters and learner characteristics.
4. The best instruction is efficient, effective, and motivating.
5. There should be congruence among objectives, learning activities, and assessment.
   - Use behavioral objectives to be clear what learner is supposed to do
   - Use criterion measurement assessment that matches objectives and compares learners only to objectives
6. The instruction itself should be evaluated to ensure it works.
7. Learners are evaluated against attainment of objectives, not against each other.

II. General ID Guidelines (Dick, Carey, & Carey; Morrison, Ross, & Kemp; Seels & Glasgow; Smith & Ragan)

A. ID is a collaborative process, involving client, SME, ID, production

1. Sometimes there will be conflicts among interested parties; ID's job is to resolve them.
2. Clients frequently need persuading to do portions of the ID process.

B. Despite client preferences, we always validate the ''givens'' of a project

1. We always validate the given problem (see III. Analysis), but without overanalyzing.
2. We always validate the appropriateness and feasibility of the requested delivery medium during analysis and design.

C. Clients are interested in results, NOT ID language or process

D. ID does not impose a cookie cutter ID process or ID solution on a problem

E. IDs make instructional decisions based on theory and research, not intuition or fad

F. ID requires project management of time, resource, and quality, as well as of clients

III. Analysis (Carlyle; Jonassen, Tessmer, & Hannum; Mager & Pipe; Rossett)

A. Rarely is the given problem one solved by instruction, and if it is, it is not the root cause problem

(*Continued*)

Exhibit 2.1 (*Continued*)

1. Identify the real reasons people do not perform as desired to determine what to put in the text, IG, online, job aids, videos, etc.
    - Not told what desired performance is
    - No feedback on performance
    - No incentives to perform
    - Internal motives do not match doing performance
    - Person not capable of performing
    - No resources to aid performance
    - Doing other things interferes with desired performance
    - Used to be able to do it; forgot
    - Culture/environment goes counter to performance
    - Performance not doable as specified
    - Never had, and now don't have, skill/knowledge

B. We have to know what they are doing now and what we want them to be doing

1. Data-based decisions about current and desired performance/ knowledge are better than gut feelings.
2. SMEs do not understand what the problem is; they are experts, and think differently from novices.
3. Collect data about performance, learners, and environments from at least three sources:
    - Extant data/documents
    - Learners
    - Experts
    - Managers
    - Benchmark data
    - Societal norms
    - Actual live performance
    - Test scores
4. Collect data about performance, learners, and environments in at least three ways:
    - Interview
    - Observation

- Reading
- Tests
- Survey
- Groups

5. Balance the amount of analysis you do with other time and resource requirements of the project, but do NOT skip it completely.

C. Collect data about the environments in which people will perform and will learn as much as will help guide the design:

1. What job aids and equipment do they have when they perform the tasks?
2. What facilities and equipment are available in the classroom setting?
3. What computer equipment is available in the instructional setting and at learners' homes?

D. Collect data about the learners, as much as will help guide the design.

1. General demographic characteristics:
   - Age
   - Gender
   - School level
   - Reading level
   - Visual literacy
   - Ethnicity
   - Native language
   - Information processing ability (high/low amounts at a time)
   - Interests in free time outside of food industry (to be used for analogies)
   - Emotional maturity
   - Sociability
   - Motivation to learn the content
2. Prerequisite knowledge
   - Prior experience eating food (what foods they've eaten/not eaten/not heard of)
   - Prior experience handling food (e.g., cooking at home, shopping)

(*Continued*)

Exhibit 2.1 (*Continued*)

- Prior experience eating in establishments (restaurants, hospital cafeterias, etc.)
- Prior work in food service industry (e.g., fry cook at McDonald's)

IV. Lesson Design (Foshay, Silber, & Stelnicki; Gagné; Merrill)

A. Help the learner select the information to attend to

1. Gain the learners' attention
   - Draw on learners' personal interests by using visuals that make the relevance of the instruction obvious.
   - Avoid Las Vegas-type approaches to learning by minimizing visuals that stimulate emotional interest.
   - Present a realistic problem/case situation.
2. Motivation: Explain what's in it for them (can be part of A1) to increase interest
   - What good can happen if they learn it?
   - What bad can happen if they learn it?
3. Give the learners confidence (You can do it!) that they can learn it (especially important in difficult content like temperature and microbes) (Can be part of A1)
   - Explain how other learners like them learned it.
   - Explain how it is something like what they've already learned.
   - Use visuals to make lessons more concrete, familiar, and coherent.

B. Activate relevant prior experience

1. Encourage learners to recall a knowledge structure to organize new knowledge (something they already know from either learner analysis or prior section of book).
2. Provide learners with real-life experience/example that is foundation for new knowledge.
3. Recall and relate their prior knowledge to the new knowledge.

C. Organize the knowledge

1. Present a structure of content (advance organizer, visual if possible, that shows not just list of content but relationships among content) at the beginning.
2. Present lesson objectives or questions to think about.

3. Use graphic design to visually represent the content to aid in organization.

D. Demonstrate what is to be learned (vs. telling)

1. Use of examples
   - Presenting examples is more important than facts for learning and recall.
   - Where possible, present facts and examples together; if must be sequenced, put *examples* first.
   - Use ''worked examples'' (i.e., showing the concept, procedure, etc., being done); for novice learners, it requires less processing load than does practice exercises.
   - The most efficient way to present material is to present an example, then a similar problem to solve immediately following.
   - The example should be placed contiguous to the fact, concept, principle, procedure being exemplified so the learner can see the information and example together.
   - If the example is complex, provide visual and/or verbal cues on where the learner should focus his or her attention.
   - Examples should be realistic; they should demonstrate the content in a job realistic context.
2. Facts = structure; mnemonics; job-aids
   - Present facts in a structure, using

     Organized visuals and text to show how facts relate to one another in a structure

     Illustrative visuals and text to show concrete facts in job context
   - Use mnemonic visuals when physical memory aids are not available and facts must be recalled
   - In the work or training environment, display discrete factual data where it can easily be seen when needed
   - Use tables and charts and graphs to illustrate or support discovery of relationships or trends in numeric data
3. Concepts = examples and non-examples
   - Show the structure of the concepts using organization visuals
   - Present:

     The attributes of the concept

(*Continued*)

Exhibit 2.1 (*Continued*)

Examples of the concept

Non-examples of the concept (which may be examples of a related concept)

The *examples first, or simultaneously*

- Wherever possible, present related concepts simultaneously.

  Highlight similarities and differences.

  Use examples of one as non-examples of the other.
- Use illustrative visuals to present the examples/non-examples where possible.
- Use visual analogies especially for more abstract or unfamiliar concepts.

4. Principles
   - Present the context for the principles by showing the principles in action in a real-world setting, using illustrative visuals (example first).
   - Present the structure of the principles using an organizational visual.
   - Present the principles as if . . . then . . . statements (*not* then . . . if . . . format).
   - Use illustrative visuals as appropriate to clarify principles.
   - Present examples of principles working together in realistic setting.
5. Procedures = demonstrations
   - Present the context for the procedure by showing it being performed in a real-world setting, using illustrative visuals (example first).
   - Present the structure of the procedure, using a flowchart or table; use visuals if possible.
   - Manage load when presenting procedure steps by chunking, teaching procedure chunk by chunk, step by step, from beginning, and using attention-focusing strategies.
   - Use visuals to show both the steps and the results of the steps.
   - Present warnings *before* teaching the procedure, *not* at the end.

6. Processes = visualizations
   - Present the context for the process by showing it happening in a real-world setting, using illustrative visuals (example first).
   - Present the structure of the process, using a flowchart or table; use visuals if possible.
   - Use words and visuals such as flow diagrams and animations that show state changes in the process.
   - Manage load when presenting process visuals by teaching system components first, and using attention-focusing strategies.
   - Use visuals such as schematics and visual analogies to represent abstract processes such as how illness develops.
7. Provide guidance for learners
   - Direct them to relevant information.
   - Use multiple examples/demos.
   - Explicitly compare multiple demonstrations.
8. Use media for relevant instructional role, not to compete.
9. Use graphic design principles extensively here as described above.

E. Require learners to use their knowledge
1. Use practice exercises that require learners to process information in a job-realistic context.
2. Use practice activities that require the learner to respond in similar ways to the ways they will on the job, and match the objectives, NOT just multiple-choice questions.
   - Fact objective = use the fact to complete a task, provided with a job aid with the facts on it
   - Concept objective = identify a novel instance of the concept
   - Process objective = solve a problem or make a prediction about the process
   - Principle objective = solve a problem or make a prediction using the principle
   - Procedure objective = perform a task by following the step
3. Intersperse practice exercises (distributed) throughout the lesson, rather than bunched (massed) at the end of the lesson.

(*Continued*)

Exhibit 2.1 (*Continued*)

4. Following the power law of practice, use more practice to produce greater and quicker learning.
5. Use more than one practice exercise per objective.
   - The number of practice exercises depends on the complexity of the content and the level of the learners.
   - Require learners to solve problems varied in type and difficulty.
6. Initially guide learners with feedback and coaching; then withdraw gradually.
7. Present instructions for practice in writing, NOT in live/recorded voice.
8. Present feedback using some or all of the following:
   - Actual answer judging followed by ''you are correct'' or ''that is not correct, and here's why. . . ; try again''
   - Providing the correct answer and saying ''If your answer was similar to this, you were correct.''
   - Providing common errors and saying ''If you did make any of the common mistakes, then your answer is not correct'' or ''be sure to avoid doing any of the following''
9. Present feedback in written format so learners can compare their answers to the correct one.

F. Engage learners in real-world problem solving.

1. Have learners solve progression of problems; compared to one another.
2. Engage learners at problem or task, NOT action, level (practicing whole more important than practicing part).
3. Show learners task/problem at beginning.

G. Have learners integrate new knowledge/skill into everyday life, and reflect on it.

1. Provide a summary; repeat structure of content or use bulleted list.
2. Give job aids to assist in later recall and use.
3. Present activities learners can do to assist in recall and application.

V. Creating Instructional Solutions (Development) (Clark & Estes; Clark & Lyons; Clark & Mayer; Hartley; Horn; Jonassen)

A. Reduce cognitive load.

1. Use cuing to draw attention to salient features.
2. Chunk—no more than five to seven items per list or per visual.
3. Present and build mental models.
4. Avoid including distracting and decorative visuals, effects, emotions, sounds that distract the learner from important instructional content.
5. Use similarity/repetition of design approach, elements, colors, layout throughout.
6. Use proximity of images and words to facilitate dual encoding.
7. Create a metaphor for the design (desktop, chart, image).
8. Use standard structures to show hierarchy and relationships.
9. Check all symbols/metaphors for cross-cultural appropriateness.
10. Use visuals in place of text when content can be communicated visually.
11. Plan visuals that are consistent in style and low in complexity.
12. Use words or visuals alone when information is self-explanatory. Avoid presenting redundant information when one representation is sufficient.
13. Present individual components of a complex visual first when learners need to build a deep understanding of the content.
14. Use visual mnemonics.
15. Use cuing visuals to draw attention to important instructional content when the display is complex for the learner.
16. Readability should be:
    - Flesh-Kincaid reading level = 7 or 8 (*New York Times* = 6.2)
    - Reading Ease Score (0 = hard; 100 = easy) = 60 to 70 = 8th grade (*Time* magazine = 52; *New York Times* = 70.4)
    - Fog Level = 7 or 8 (*New York Times* = 8.9)
17. B and C below provide specific guidelines for reducing cognitive load.

B. Media selection

1. The medium is NOT the key factor in learning.
2. It is the ‘‘active ingredients’’ (the instructional design strategies built into medium) that affect learning.

C. Use appropriate text design

1. Use *Information Mapping* as a technique to organize and display information clearly, especially on paper

(*Continued*)

Exhibit 2.1 (*Continued*)

2. Divide content into types:
   - Facts
   - Concepts
   - Principles
   - Structures
   - Processes
   - Procedures
3. Put each type of information in a separate map
4. Each map begins on a new page
5. Simple information in blocks should be presented in *chunked bullet lists*
6. Complex information in blocks should be presented *in tables of two or three co*lumns with
   - Concept/step number
   - Definition/description
   - Example/demonstration
7. To present facts and principles, use blocks that contain bullet lists or tables
8. To present facts, concepts, and principles, use blocks that contain tables with
   - Term (or principle name)
   - Definition (or description of principle)
   - Example
9. To present structures, use blocks
   - Visual showing the structure
   - Tables that name and describe the parts of the structure
10. To present processes and procedures, use blocks
    - Procedure/process tables (properly chunked), with words and visuals adjacent
    - Step/stage number and name
      Who does it (for process only)
      Description

D. Use appropriate text font, size, spacing for print
   1. Serif for body text in print

2. Sans serif for headlines in print and for all in slides and websites
3. Use font that matches tone of message
4. Use two or three fonts maximum
5. Differentiate titles/headings/chunks/etc. by:
   - Font change
   - Bold
   - Italic
   - Bold italic
   - Color

E. Use appropriate visuals in text and PowerPoint slides
   1. Use *illustrative* visuals to provide examples of the concept/process/procedure.
   2. Use *organizer* visuals in the advance organizer to organize the content.
   3. Avoid *decorative* (gratuitous) visuals.
   4. Avoid cartoons and cutesy visuals that are just there to be there.
   5. Avoid starting a lesson with a dramatic but extraneous visual that will activate inappropriate prior knowledge.

F. Use general graphic design principles in text and PowerPoint slides
   1. *Warnings:* Use visuals to draw attention to and illustrate warnings.
   2. *Point of view:* Visuals must show activity from the performer's point of view in the job environment.
   3. *Repetition:* Repeat design elements throughout each image, and throughout presentation (related to Gestalt).
   4. *Color:* Limit colors to two or three.
   5. *Proximity:* Put text and related visuals directly adjacent.
   6. *Figure/Ground:* Make sure the background color/visual is NOT so busy or dark as to hide or confuse the foreground text/information.
      - Make the most important information stand out to avoid information processing overload.

        Avoid image where figure and ground compete

        Avoid image where figure should be ground and ground should be figure (figure/ground reversal)

        Avoid image where figure and ground create optical illusion

(*Continued*)

Exhibit 2.1 (*Continued*)

- To highlight the figure:

  Use space, shape, typography, color to highlight figure

  Use appropriate lower levels of intensity for ground

  Watch intensity and width of lines in tables

  Highlight figure in color or with arrows/circles, or with enlargement/zooming/etc.

7. *Contrast:* Emphasize the key points by contrasting with rest (related to figure/ground)
8. *Hierarchy:* Present information in structures
   - Chunk information (no more than five to seven bits per chunk)
   - Show structure of chunks
   - Within each chunk, show hierarchy of information in that chunk
   - Use horizontal, vertical, or diagonal alignment to help show hierarchical organization of information
   - Use tables and charts to visualize and compare information
   - Use columns, not rows, when comparing information
   - Use shading, spacing, color, figure-ground, and chunking to make tables more readable
9. *Gestalt:* Show relation between whole and parts
   - Show big picture at beginning, and keep image with pointer throughout to show location
   - Leave some incompleteness so the mind can complete the image
   - Use contiguity so that eye follows leading images (arrows, hands, etc.)
10. *Point of view:* Shoot all visuals in procedures and process from learner's point of view

G. Use appropriate slide design principles

1. Slide content
   - Use illustrative (NOT decorative) visuals as the main content and focus of slide
   - Use organizer graphic (small size) on all slides to show learner where the current information fits in the structure of knowledge being presented

- Present words as audio narration rather than onscreen text

  This splits information across two separate cognitive channels (phonetic processing and visual processing)

  Limit on-screen text to three or four key words per idea (not script to read)

  Limit number of ideas per slide to four or five

2. Slide layout/design
   - Use screen ''real estate'' appropriately: prime, secondary, and tertiary areas
   - Follow ''golden rectangle'' rule: for each 1-inch high, 1.618-inch long (slide copy = 5×8 or 6×9)
   - Follow ''rule of thirds'': focus is one-third from left of golden rectangle, not center of slide
   - Follow ''backwards S'' rule: eye enters from top left, goes across to top right, then down one level on right, then back left, then down one level on left, then across right
   - Follow all guidelines from III A, C, D, E, F, G above

H. Use text, pictures, and audio together slide presentations

1. Use words and visuals rather than words alone; allows for dual encoding
2. Words = printed or spoken; Illustrative/Organizer visuals printed or projected
3. Redundancy principle
4. Do NOT present words in both text and audio; this decreases learning (redundancy principle) because it leads to conflicts between phonetic processing and visual processing
5. Explain complex visuals with words presented in audio.
6. Use conversational rather than formal style in visuals and audio
7. Use on-screen coaches
   - To serve a valid instructional purpose, NOT for entertainment
   - May look lifelike or be an image related to subject matter
   - Should be someone/something learners can relate to
   - Words of coach presented in audio, NOT on screen

VI. Implementation (Gelinas & James; Hale; Mourier & Smith)

A. Eighteen reasons for failure

(*Continued*)

Exhibit 2.1 (*Continued*)

1. No compelling reason
2. No sense of urgency
3. Lack of undivided attention
4. Leaders not accountable
5. Failure to learn
6. Insufficient communication
7. Staff takes the lead
8. Lack of a critical mass
9. Insufficient change in the culture
10. See change as an organizational process
11. Get stuck in the present
12. Misalignment of rewards
13. Insufficient infrastructure
14. Not enough quick wins
15. A disconnect between the vision and actions
16. Not understanding the time and resources needed for people to learn
17. No outside perspective
18. The environment shifts

B. Cornerstones of success
  1. Clarity of purpose
  2. Ongoing sponsorship
  3. Customer collaboration
  4. Measuring and reporting
  5. Rewarding adoption

C. Agree on the goal
  1. Define and agree on the need, results, business measures, target audience, critical mass

D. Set the baseline
  1. Document the current state.
  2. Identify what you will track at the organizational and individuals levels to measure change

E. Assess feasibility

F. Develop a strategy for sustaining adoption
  1. Project, communication, succession plans

G. Identify leading indicators
    1. Oversight including sponsor, program manager, steering committee, active sub-team assigned specific responsibilities
    2. Sustain attention: Keep the initiative on senior management's agenda
    3. Measure rate of adoption and report
    4. Shift ownership for the target audience to embrace the program or initiative; take responsibility for its success; take responsibility for improving it; and take responsibility for enrolling others
    5. Reward adoption

VII. Measuring Success (Evaluation) (Dick, Carey, & Carey; Morrison, Ross & Kemp; Merrill; Smith & Ragan)

A. ID evaluates/assesses learners
    1. Use criterion-referenced assessment: compare learners against objective, not each other
    2. Learner assessments should be as authentic as possible
    3. Learner assessments should be reliable and valid
    4. Learner assessments must match objectives exactly in conditions, performance, and criteria (FSS)
        - Parts of = locate/name
        - Fact = recall or recognize
        - Concept = classify novel examples/non-examples
        - Procedure = do it in novel situation
        - Principle = predict/explain consequences in a given situation

B. ID evaluates the effectiveness of the instruction and ID process
    1. Reviews
        - Content experts
        - ID experts
        - Client
    2. Formative evaluation
        - One-on-one
        - Small group
        - Large group
    3. Summative evaluation
        - Long term (aka confirmative) to measure retention

(*Continued*)

Exhibit 2.1 (*Continued*)

- Short term to measure transfer of behavior to real world (aka Level 3)
- Immediately at end of class to measure learning (aka Level 2)

4. Results
   - Improvement in organization performance (aka Level 4)
   - Return on investment (aka Level 5)

Abbreviated from Silber, 2006. © Kenneth H. Silber. All rights reserved.

CONCLUSION

This chapter has attempted to make several, perhaps controversial, arguments, to present evidence for the positions taken by the author, and to provide the reader with mental models and heuristics for thinking about ID itself, in a new way. It should be considered an initial foray into this area. It is hoped that others will expand upon or challenge the ideas contained herein. This chapter, then, has tried to make a case for the following basic ideas:

- ID is NOT a procedure; rather, it is ill-structured problem solving.
- ID has a mental model of principles and a set of heuristics for identifying and solving ID problems.

References

Andrews, D. H., & Goodson, L. A. (1980). A comparative analysis of models of instructional design. *Journal of Instructional Development*, *3*(4), 2–16.

Argyris, C., & Schön, D. A. (1996). *Organizational learning II: Theory, method, and practice*. Reading, MA: Addison-Wesley.

Braden, R. A. (1996). The case for linear instructional design and development: A commentary on models, challenges, and myths. *Educational Technology*, *36*(2), 5–23.

Carlyle, K. (1986). *Analyzing jobs and tasks*. Englewood Cliffs, NJ: Educational Technology Publications.

Cennamo, D., & Kalk, D. (2005). *Real world instructional design*. Belmont, CA: Thompson-Wadsworth.

Clark, R. C., & Lyons, C. (2005). *Graphics for learning*. San Francisco: Pfeiffer.

Clark, R. C., & Mayer, R. (2003). *e-Learning and the science of instruction*. San Francisco: Pfeiffer.

Clark, R. E., & Estes, F. (2002). *Turning research into results*. Atlanta, GA: CEP Press.

Dick, W., Carey, L., & Carey, J. (2005). *The systematic design of instruction* (6th ed.). New York: Pearson/Allyn & Bacon.

Dorsey, L. T., Goodrum, D. A., & Schwen, T. M. (1997). Rapid collaborative prototyping in an instructional development paradigm. In C. R. Dills & A. J. Romiszowski (Eds.), *Instructional development paradigms*. Englewood Cliffs, NJ: Educational Technology Publications.

Foshay, W. R., Silber, K. S., & Stelnicki, M. B. (2003). *Writing training that works*. San Francisco: Pfeiffer.

Gelinas, M., & James, R. (1998). *Collaborative change*. San Francisco: Pfeiffer.

Gagné, R. M., Briggs, L. J., & Wager, W. W. (2005). *Principles of instructional design* (5th ed.). Belmont, CA: Wadsworth/Thomson Learning.

Gustafson, K. L. (2000). Designing technology-based performance support. *Educational Technology*, *40*(1), 38–44.

Gustafson, K. L., & Branch, R. (2002). *Survey of instructional design models* (4th ed.). Syracuse, NY: ERIC.

Hale, J. (2003). *Performance-based management*. San Francisco: Pfeiffer.

Hardré, P. L., Ge, X., & Thomas, M. K. (2005). Toward a model for the development of instructional design expertise. *Educational Technology*, *45*(1), 53–57.

Hardré, P. L., Ge, X., & Thomas, M. K. (2006). An investigation of development toward instructional design expertise. *Performance Improvement Quarterly*, *19*(4), 63–90.

Hartley, J. (2003). Designing instructional and informational text. In D. Jonassen (Ed.), *Handbook of research on educational communications and technology* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Horn, R. (1982). Information mapping. In D. Jonassen (Ed.), *Technology of text* (Vol. 1). Englewood Cliffs, NJ: Educational Technology Publications.

Jonassen, D. (1982). *The technology of text* (Vol. 1). Englewood Cliffs, NJ: Educational Technology Publications.

Jonassen, D. (1985). *The technology of text* (Vol. 2). Englewood Cliffs, NJ: Educational Technology Publications.

Jonassen, D. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology Research and Development*, *45*(1), 65–94.

Jonassen, D. (2004). *Learning to solve problems*. San Francisco: Pfeiffer.

Jonassen, D., Tessmer, M., & Hannum, W. (1998). *Task analysis methods for instructional design*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Lawson, B. (2004). *What designers know*. Oxford, UK: Architectural Press.

Lawson, B. (2006). *How designers think* (4th ed.). Oxford, UK: Architectural Press.

Mager, R. F., & Pipe, P. (1997). *Analyzing performance problems* (3rd ed.). Atlanta, GA: Center for Effective Performance.

Merrill, M. D. (2002). First principles of instructional design. *Educational Technology Research and Development*, *50*(3), 43–59.

Morrison, G. R., Ross, S. M., & Kemp, J. E. (2001). *Designing effective instruction* (3rd ed.). Hoboken, NJ: John Wiley & Sons.

Mourier, P., & Smith, M. (2001). *Conquering organizational change*. Atlanta, GA: CEP Press.

Nelson, H. G., & Stolterman, E. (2003). *The design way*. Englewood Cliffs, NJ: Educational Technology Publications.

Newell, A., & Simon, S. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Osguthorpe, R. R., & Osguthorpe, R. D. (2007, July/August). Instructional design as a living practice: Toward a conscience of craft. *Educational Technology*, *47*(4), 13–23.

Reigeluth, C. M. (Ed.). (1983). *Instructional design theories and models*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Reigeluth, C. M. (Ed.). (1999). *Instructional design theories and models. Vol. 2. A new paradigm of instructional theory*. Mahwah, NJ: Lawrence Erlbaum Associates.

Rossett, A. (1987). *Training needs assessment*. Englewood Cliffs, NJ: Educational Technology Publications.

Rowland, G. (1992). What do instructional designers actually do? An initial investigation of expert practice. *Performance Improvement Quarterly*, *5*(2), 65–86.

Rowland, G. (1993). Designing and instructional design. *Educational Technology Research and Development*, *41*(1), 79–91.

Rowland, G. (2004). Shall we dance? A design epistemology for organizational learning and performance. *Educational Technology Research and Development*, *52*(1), 33–48.

Seels, B., & Glasgow, Z. (1998). *Making instructional design decisions* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Shambaugh, N., & Magliaro, S. (2001). A reflexive model for teaching instructional design. *Educational Technology Research and Development*, *49*(2), 69–92.

Silber, K. H. (2006, April). A radical, principle-based approach to looking at instructional design. Paper presented at the International Conference of the International Society for Performance Improvement, Dallas.

Smith, P. L., & Ragan, T. L. (2004). *Instructional design* (3rd ed). Hoboken, NJ: John Wiley & Sons.

Tessmer, M., & Wedman, J. F. (1990). A layers of necessity instructional development model. *Educational Technology Research and Development*, *38*(2), 77–85.

Thiagaragan, S. (2004, October). Faster, better, cheaper: Training design principles. Paper presented at the Instructional System Design Conference of the International Society for Performance Improvement, Chicago.

CHAPTER THREE

Origins and Evolution of Instructional Systems Design

Michael Molenda

INTRODUCTION

This historical survey strives to synthesize existing accounts about the origins and evolution of instructional design—such as Dick's (1987), Saettler's (1990), Reiser's (1987, 2001, 2007), Shrock's (1991), and Molenda's (1997), to add several first-person reports (for example: Briggs, 1980; Diamond, 1980; Hannum, 2005; Popham, 1980; Schuller, 1986), and to place all the accounts into a larger and more coherent framework.

The History of What?

This chapter surveys the history of the concept of instructional design as it is known in the educational technology field. Obviously, there are as many ways of preparing instruction as there are teachers and trainers. Often the approach is more intuitive than deliberate. If deliberate, the approach may follow any of a number of different paradigms. For example, in the early days of audiovisual education, most media producers followed a planning process borrowed from commercial film-making—the treatment and script. Nowadays, creators of instructional products may look to design traditions such as those in architecture, graphics, fine arts, software engineering, and product design. Within educational technology, however, the dominant paradigm is the *systems approach* to the design of instruction, the essence of which is ''to subdivide the instructional planning process into steps, to arrange those steps in logical order, then to use the output of each step as the input of the next'' (Molenda & Boling, 2008, p. 104). Those steps are typically given as analysis, design, development, implementation, and evaluation.

To be sure, within the field of educational technology, many other approaches have been proposed. Dills and Romiszowski (1997) describe approaches such as cybernetic, behavioral analysis, situated cognition, semiotic, direct instruction, constructivist, existentialist, structural communications, rapid collaborative prototyping, simulations, and intelligent tutoring, among others. Some of these are genuinely divergent ways of thinking about the creation of learning environments, but many are not intended as guides to the whole process of planning instruction, dealing, rather, with strategies and tactics within the ''design'' step of the systems approach model.

All of these approaches have different histories, so it would be impossible to encompass all of them in one grand narrative. Consequently, this chapter focuses on the story of the systems approach version of instructional design, usually referred to as *instructional systems design* or *instructional systems development*. To avoid a hair-splitting argument about which of these two terms is more legitimate (since ''design'' and ''development'' both have persuasive claims to being the broader term), I will use the acronym ''ISD'' to refer to the broad concept of instructional systems design/development. A definition by Leslie Briggs of Florida State University (1977) just as the concept was taking hold proposes several key elements:

A systematic approach to the planning and development. . . .

- All components of the system . . . are considered in relation to each other;
- The resulting delivery system is tried out and improved before widespread use. (p. xxi)

As this definition indicates, the term *systems* is meant to connote an approach that is both systematic and systemic.

THE HISTORICAL PRE-CONDITIONS FOR ISD

The Post-War Environment for Education

With the end of World War II in 1945 came the end of rationing restrictions and the return of the millions of men and women from the military services. Among them were many who had experienced first-hand the gargantuan job of ''rapid mass training'' that had been accomplished through the use of motion pictures and other audiovisual media. The pool of highly skilled audiovisual developers and users who returned to civilian life provided a jolt of energy that accelerated the pace of change in education. With their experience of using media to multiply the effect of good teaching, these trainers and educators were open to ideas for increasing the efficiency of instruction as well improving its quality.

By the early 1950s the products of the Baby Boom were entering elementary school, triggering a corresponding school construction boom. With new, modern schools came new, modern technology: classrooms outfitted with electrical outlets at the front and back, permanently mounted projection screens, and shades or blinds for room darkening to accommodate the burgeoning audiovisual media. The expense of constructing and equipping new school buildings was challenging school districts across the United States, but there was another even more daunting challenge facing them—providing the teachers for all these new classrooms. Reports from agencies such as the Fund for the Advancement of Education (1955) were projecting a serious teacher shortage in the coming decade. Educators were fearful there would simply not be enough teachers to go around, at least not enough fully qualified teachers.

Some forward-looking leaders in the audiovisual education field were thinking about the possibilities of using technology to automate some aspects of the educational process, to leverage the human resources that were available. The most visible of these leaders was James D. Finn, a professor at the University of Southern California (USC), who would become president of the Department of Audio-Visual Instruction (DAVI), the predecessor of AECT, in 1960. In a series of articles published between 1957 and 1960 (Finn, 1957a, 1957b, 1960), he proposed the application of the methods of the ‘‘second industrial revolution’’ to formal education under the rubric of ‘‘automatizing the classroom.’’ However, the political will to undertake such a sweeping change was slow to appear. That would change very shortly.

Sputnik and a Crisis in Education

The U.S. public was shocked to attention on October 4, 1957, when the Soviet Union successfully launched Sputnik I, the world’s first artificial satellite. Both the U.S. and USSR had been working on earth-orbiting satellites, but everyone was caught off-guard by the USSR’s launching first. The stunning Soviet technological achievements prompted an urgent examination of the U.S. scientific establishment, including educational preparation in areas of science and technology. The U.S. Congress did not take long to come to a consensus that American schools and colleges were not producing the quantity and quality of scientific and technical specialists necessary to keep pace with the Soviet Union.

This perceived crisis propelled Congress to pass a number of emergency measures in 1958, including the National Defense Education Act (NDEA). Until this time, the federal government had not played a substantive role in public education, leaving it to the individual states. But the urgency of the crisis overrode old trepidations, and federal funds, hundreds of millions of dollars, began to be invested in the teaching of science, technology, foreign languages, and other fields related to the Cold War struggle.

NDEA in Action

The new education legislation supported numerous types of activities under different titles of the NDEA. Those with greatest impact on instructional technology were Titles III, IV, VI, and VII.

Title III. This section authorized grants ($70 million per year) for purchase of equipment to strengthen science, math, and foreign language instruction. A sizable portion of these funds was used by schools to purchase AV equipment and materials.

Title IV. This program funded hundreds of fellowships per year to support three years of graduate study for individuals intending to become college teachers. Many of the future leaders in instructional design and technology were educated at NDEA fellowship programs at Syracuse, Michigan State, and USC. Title VI supported research on methods and materials for language teaching and area studies centers at universities. It also provided stipends to teachers to attend summer institutes on methods and materials for teaching. These summer institutes introduced thousands of teachers to the new educational media, many of whom became technology advocates back home.

Title VII. The first part of Title VII promoted ''research and experimentation in more effective utilization of television, radio, motion pictures, and related media for educational purposes.'' As Saettler reports (1990, p. 413), this part of the act was an afterthought, instigated by lobbyists for the audiovisual trade association, National Audio-Visual Association (NAVA). Nevertheless, it provided, in the first year alone, $1.6 million for forty-five research projects at universities across the United States. A comprehensive evaluation of the impact of Title VII activities by Filep and Schramm (cited in Saettler, p. 414) concluded that this program was successful in bringing new researchers into the educational media field, upgrading the quality of research, and encouraging the growth of academic programs in educational media. It also promoted individualized instruction and teacher acceptance of media.

Overall, NDEA programs helped create the infrastructure—the people, hardware, and ideas—necessary to support the dawning idea of a systems approach to the design and implementation of instruction.

CONCEPTUAL UNDERPINNINGS OF ISD: BEHAVIORIST LEARNING THEORY

Teaching Machines and Programmed Instruction

In the midst of growth and change in American education, some radically new concepts were coming to the fore. Behavioral psychologist B.F. Skinner had presented his first teaching machine, based on operant conditioning principles, in 1954 (Skinner, 1954) and major school demonstration projects were underway between 1957 and 1962 (Saettler, 1990, pp. 297–302). Shortly after, Norman Crowder introduced a variant format for teaching machines that was not based on any particular theory of learning, but on a practical concern for efficiency. It featured a more flexible program structure that allowed learners to skip ahead through material that was easy for them or to branch off to remedial frames if they encountered difficulty (Crowder, 1962). His method was quickly dubbed *branching programming* because a schematic outline of the program resembled a tree trunk with multiple branches. Initially, Crowder's programs were used in the AutoTutor teaching machine, but Crowder soon joined the rush to convert programs to book form, and his TutorText series became one of the best-known series of programmed materials.

From PI to Technology of Teaching

As research and field experience accumulated, it became clearer that the ''magic'' of PI was not in the hardware, and possibly not even in the software—the step-by-step breakdown of information followed by questions, responses, and confirmation or correction of the response. Rather, the success of PI could be attributed more to the planning process by which the software was developed. Referring to this process as a ''technology of teaching'' was first proposed by B.F. Skinner (1965) and elaborated in his later book (1968) to describe his view of programmed instruction as an application of the science of learning to the everyday tasks of teaching. This view coincided with the notion promoted earlier by Finn that instructional technology could be viewed as a *way of thinking* about instruction, not just a conglomeration of devices, echoing the recently popularized notion of economist John Kenneth Galbraith (1967) that technology should be seen as ''the systematic application of scientific or other organized knowledge to practical tasks'' (p. 12).

From Technology of Teaching to Design Methodology

The procedures for creating PI materials followed the prescriptions for operant conditioning experiments: analyzing the task to be learned in order to break it down into a series of small steps, specifying the behavioral indicator of mastery of each step (performance objective), sequencing the behavioral responses in hierarchical order, creating prompts for the desired responses, observing the learner response, and administering appropriate consequences for each response.

Since reinforcement theory called for practicing mostly correct responses, each frame of the program had to be tested for efficacy. In fact, developmental testing was a mandatory specification for materials destined for the military training market. The U.S. Air Force required that ''at least 90 percent of the target population will achieve 90 percent of the objectives'' (Harris, 1964, p. 142). This was known as the 90/90 criterion and was widely accepted as the standard benchmark of effectiveness. This demanded a commitment to evaluation and revision far beyond what had been typical in the past. So the PI development process that evolved was characterized by careful specification of objectives, active responses, immediate feedback, and repeated rounds of testing and revision.

Gradually, PI developers began to realize that it was the painstaking development *process* that made PI successful:

> ''The uniqueness and strength of programmed instruction lie mainly in its production process . . . Programmed instruction is developed through a process which has empirical and analytic qualities.'' (Lange, 1967, p. 57)

The focus on the design process was championed by Susan Meyer Markle, one of Skinner's brilliant associates. She and her partner, Phil Tiemann, proclaimed that ''programming is a process'' (Markle & Tiemann, 1967). That is, it is not the PI *format* that accounts for success, but rather the developmental *process*; she particularly emphasized the importance of developmental testing of prototypes of the lesson (Markle, 1967). At about the same time, Michael Scriven, a mathematician and leading theorist in evaluation, coined the term *formative evaluation* to describe procedures for testing and revising prototypes while they were still in development, rather than waiting until the final product was mass-produced and ready for rollout (Scriven, 1967).

Markle and Tiemann's procedural flow chart for PI product development consisted of analyzing learners and learning tasks, specifying performance objectives, requiring active practice and feedback, and subjecting prototypes to testing and revision; it can be seen as a precursor to the analyze, design, develop, implement, evaluate cycle proposed in later ISD models.

In addition to developmental testing or formative evaluation, one of the procedures that was central to both PI writing and ISD was the specification of precise learning objectives. During the heyday of PI, Robert Mager wrote a brief, humorous, branching programmed booklet on how to write—and how *not* to write—objectives. It was so popular that he prepared a more polished version for publication, entitled *Preparing Objectives for Programmed Instruction* (Mager, 1961). As the book's renown spread to broader audiences, including teacher education programs, the publisher reissued it with a more generic title, *Preparing Instructional Objectives* (Mager, 1962). It became a classic, selling over two million copies in the following three decades (Heinich, Molenda, & Russell, 1989, p. 45). For many educators, this would be their closest brush with ISD concepts, hence the importance of this book in promulgating ideas related to ISD. In addition, the book became the anchor of a series of brief, breezy, programmed texts by Mager, known as ''The Mager Library,'' packaged as a boxed set, comprising five titles (Mager, 1984a, 1984b, 1984c, 1984d, 1984e). The series was widely used in corporate train-the-trainer programs and academic programs for teaching about instructional design, and it also constituted the first primer for the nascent field of performance technology.

Individualized Instruction

The dramatic breakthrough of PI was the idea that self-study materials could be structured in such a way that each learner could move through the material at his or her own pace and could even be directed to content that was highly specific to his or her needs. The individualized instruction notion was expanded to include audiovisual materials as PI projects added various types of playback devices under machine control, later under computer control.

At the same time, others were experimenting with self-instructional systems that were not based in behaviorist learning theory. A ''poor man's'' version of multimedia self-instruction was developed by a Purdue University biology instructor, Sam Postlethwait, beginning in 1961, under the label of Audio-Tutorial System. He began very modestly by making audio recordings of his botany lectures for students who missed class (Heinich, Molenda, & Russell, 1989, pp. 318–319). He later enhanced this by placing a tape recorder in each of several learning stations equipped with plants and lab apparatus needed to do experiments. Gradually, Postlethwait amplified his lectures with slides, filmstrips, and 8mm film loops. Students could come to the lab at their convenience and listen to his lectures while looking at supporting visual materials, then do experiments and write up their reports. In the fully developed Audio-Tutorial System, the lab was supplemented with periodic discussion sessions wherein students were responsible for being prepared for questioning by grad assistants. Large-group lectures were scheduled for guest speakers and film showings (Postlethwait, 1968; Postlethwait, Novak, & Murray, 1972). This formula proved so successful that Postlethwait and fellow enthusiasts formed an organization in 1970, the International Audio-Tutorial Congress, which morphed over the years through several identities before arriving at its present name, International Society for Exploring Teaching and Learning (ISETL), while continuing to attract adherents.

Federal funds lure business involvement. Interest in PI and individualized instruction mushroomed in the mid-1960s with the rapid growth of federal government investment in education in connection with President Lyndon B. Johnson's ''War on Poverty.'' The Economic Opportunity Act of 1964 created the Job Corps, which provided general and vocational education, technical training, and work experience at residential centers for young people from poverty backgrounds. Overnight there was a huge market for self-instructional materials and programs for the tens of thousands of learners in dozens of Job Corp centers, and the ''learning industry'' was launched. Companies such as GE, Westinghouse, Litton Industries, and Morton Thiokol established large units to create individualized materials and to manage learning systems. A number of future contributors to the ISD movement, including Robert Morgan, Robert Branson, and Donald Tosti, among others, gained real-world experience working in the learning industry on Job Corps projects (D. Tosti, personal e-mail communication, July 24, 2008).

The financial opportunities prompted a frenetic phase of mergers and acquisitions among hardware and software companies, for example: 3M + *Newsweek*, RCA + Random House, Sylvania + *Reader's Digest*, IBM + Science Research Associates, and General Electric + *Time* (Bern, 1967). Involvement by such big businesses investing such large sums in educational technology directly and indirectly promoted growth in the academic sector of the field. As one small example of an indirect effect, the fees paid by commercial exhibitors at the DAVI (later AECT) conventions provided a majority of the annual budget of the association; this allowed the association to publish journals and hold conferences, allowing scholars to communicate and share ideas.

Federal funding for school technology projects. The Johnson Administration funneled federal funds into formal education as well as into non-formal education programs such as the Job Corps. The Elementary and Secondary Education Act (ESEA) of 1965, among other things, established regional educational laboratories, which in turn supported innovative activities in schools. Two major systems for individualized instruction were tested in schools through such programs. Individually Prescribed Instruction (IPI) was developed by the Learning Research and Development Center (LRDC) at the University of Pittsburgh and implemented at Oakleaf Elementary School near Pittsburgh in 1964. In IPI, students worked through self-instructional units individually, took a post-test, and, if they demonstrated mastery, moved on to the next unit. IPI's independent-study materials and tests became quite widely disseminated through Research for Better Schools, another regional lab, but after about a decade, federal funds were withdrawn, ending the project (Saettler, 1990, p. 305).

In a similar project, American Institutes for Research (AIR) collaborated with the Westinghouse Learning Corporation to develop Project PLAN (Program for Learning in Accordance with Needs), focused on individualizing education and demonstrating how computers could contribute to the process, mainly by keeping track of student progress. Within a decade, around one hundred schools were participating, but PLAN ground to a halt after federal funds dried up. Schools were unable or unwilling to pay the cost of participation, and the corporate vendors were unable to make a profit (Saettler, 1990, p. 306).

Professional Associations Support PI

DAVI. A number of scholars in the audiovisual instruction field were quick to recognize the connections between their concerns and the potentialities of PI. The 1959 convention program of DAVI had a single research paper devoted to this topic: ''Teaching Machines and Self-Instructional Materials: Recent

Developments and Research Issues,'' but by the late 1960s the convention offered about a dozen sessions a year on PI.

The concept of a pre-packaged set of materials to be used independently by learners, although new to many educators, fit into a conceptual niche that already existed for audiovisual professionals. By the late 1950s there had already been a number of demonstration projects in which entire courses of study had been presented successfully via film or television (Heinich, 1970, pp. 120–122). For example, the Rocky Mountain Area Project, 1958–1960, demonstrated that a high school physics course on film could be used successfully to substitute for teachers in schools lacking qualified staff (Scott, 1960). During the same period, the schools in Hagerstown, Maryland, used closed-circuit television to transmit whole lessons in core subjects via television (Washington County Board of Education, 1963).

Hence, many leaders in the audiovisual profession already had a systemic vision of the classroom of the future, in which the task of presentation of information could be performed by pre-recorded material. The PI notion just went one step further, allowing each student to interact individually with the material.

A special interest group was formed in 1959, conducting sessions at the next several DAVI conventions under the name of the Teaching Machine Group. However, researchers quickly realized that the hardware of teaching machines was subordinate to the software, the written materials inside them. DAVI's embrace of these new phenomena was signaled by the publication of a collection of key documents on teaching machines and programmed learning (Lumsdaine & Glaser) in 1960, and then a later compilation of research and commentary (Glaser, 1965b).

NSPI. But DAVI was not the only, or even the primary, professional association interested in PI. When Air Force experiments in 1961 demonstrated the dramatic time and cost advantages of PI, military trainers and university researchers quickly formed an informal interest group, which by 1962 became a national organization, the National Society for Programmed Instruction (NSPI). The organization grew to encompass thousands of members in the United States, Canada, and other countries; its periodical, *NSPI Journal*, later *Performance and Instruction*, during the formative years of ISD chronicled the advance of new ideas and newly developed procedures for the improvement of instruction.

In 1973 the society's name was changed to the National Society for Performance and Instruction, reflecting the shift of focus from the PI format to the larger process of creating materials and systems that changed human performance. Decades later, as the interests of members grew and evolved to include all sorts of technological interventions for improved human performance, the name, too, evolved to its current form, International Society for Performance Improvement (ISPI) in 1995.

CONCEPTUAL UNDERPINNINGS OF ISD: SYSTEMS APPROACH

Systems Approach in the Military

Undoubtedly, the most important influences on the emergence of ISD originated in the military services. The titanic military struggles of World War II had ushered in an era of innovation in warfare. An analytical technique that grew out of submarine hunting was called operations research, in which computers were used to make the calculations required. After the war, this approach to man-machine operations, now referred to as *the systems approach*, was applied to the development of training materials and programs. During the post-war period, each of the U.S. military services had developed its own model for training development, all of which were based on the systems approach, a ''soft science'' version of systems analysis, itself an offshoot of operations research (McCombs, 1986). Alexander Mood (1964), a pioneer in the application of statistical methods to complex problems, speaking at an early conference on the systems approach in education, explained the distinction. In his view, *systems analysis* is the name of a rigorous analytical method involving the construction of a mathematical model of some phenomenon in order to experiment with some of the functions, to determine whether changes lead to desired effects. *Systems approach*, on the other hand:

> ''Is simply the idea of viewing a problem or situation in its entirety with all its ramifications, with all its interior interactions, with all its exterior connections and with full cognizance of its place in its context.'' (p. 1)

The systems approach was viewed in the military as a methodology for combining the human element with machine elements, an antidote to purely mechanistic thinking. They no longer had weapons; they had weapons *systems*.

This concept had a direct impact on training in the 1960s when the U.S. armed forces changed their bidding procedures for new weapons systems, requiring contractors to provide not only the hardware, but the training needed by the operators (Dick, 1987). Defense contractors had to become systems thinkers.

The next step was applying the systems approach to training within the military itself. The systems approach offered the armed forces a way to standardize training procedures and doctrines within very large, complex, and far-flung organizations. Further, PI and other forms of individualized instruction offered a vital lifeline to military training managers. In the late 1960s and early 1970s, they were facing a ''crunch'' stemming from three factors: (1) the shift to an all-volunteer military, meaning a higher turnover of lesser skilled recruits, (2) the new military technologies coming on line, requiring ever more sophisticated training, and (3) Defense Department budgets that were not expanding enough to accommodate the needed training as currently designed and delivered

(Hannum, 2005). The military services *had* to find ways to get ''more bang for the buck'' in training. Over the next decades, the U.S. armed forces would invest billions of dollars in research, development, and implementation of technology-based training solutions. Many of those dollars went to private corporations, some of the same companies offering services to federally funded school innovation projects.

The U.S. Army had been experimenting with a systems approach to training for several years. For example, Project Minerva within the Army Security Agency developed a ten-stage system design model that contained all the elements found in later ISD models (Tracey, Flynn, & Legere, 1967). By 1968 the U.S. Army had officially adopted a training doctrine called Systems Engineering of Training, CON REG 350-100-1 (Quinn, 1970). The U.S. Air Force adopted a similar doctrine in 1972. This was a ground-breaking development, but not yet the most important in terms of large-scale dissemination beyond the armed forces. That development came a bit later.

Military ISD Model at Florida State

In the early 1970s, Robert Morgan and Robert Branson, who had been immersed in the systems approach in Job Corps projects, were at Florida State University, where they participated in launching a new academic program in instructional systems. The Center for Performance Technology there was selected in 1973 by the U.S. Department of Defense to develop procedures to substantially improve Army training. This team, according to Hannum (2005), was asked by the Army to ''(1) uncover the best approaches for developing and delivering training and (2) develop a set of procedures to guide the implementation of such approaches.'' The team conducted a thorough review of documentation of training across all the armed services and made site visits to key military training installations, and they sought the advice of Robert Gagné regarding theoretical bases of training and instruction (Hannum, 2005, p. 5).

The resulting ISD procedures developed for the Army evolved into a model that was adopted by the Army, Navy, Air Force, and Marines, called the Interservice Procedures for Instructional Systems Development (IPISD). As reported later by Branson (1978), the detailed procedures clustered around five major functions—analyze, design, develop, implement, and control. The IPISD model eventually had enormous influence in military and industrial training because its use was mandated not only in all the U.S. armed services but also among defense contractors.

Systems Approach in Business and Industry

As the military services were experimenting with the systems approach and moving toward specific models of the ISD variety, similar movements were taking place in the private sector, often in businesses that were involved in the defense industry, which helps explain the parallel developments. David Curl, reporting on the emergence of the systems approach in a range of businesses (1967), proposed ''a basic plan to follow in preparing an instructional program or system'' (p. 41); it was an eight-step procedure incorporating the major elements found in later ISD models.

Large corporations were dealing with the same ''crunch'' that the armed forces were—large, complex organizations with thousands of employees engaged in using increasingly sophisticated tools, requiring efficient, effective training that could be standardized across sites. The largest corporations in the United States were among the first to report their tentative steps toward systematic design processes, most notably AT&T and its subsidiaries, as reported by Bumstead (1968), Dyer (1969), and Ford (1970). Meanwhile, Douglas Aircraft (Nicely, Nelson, & Kaufman, 1970) and Kodak (A system to create training systems, 1971) were among the companies that had progressed furthest toward developing full-fledged ISD models as their training design lodestones. It is no coincidence that the efforts at these last mentioned companies were led, respectively, by Roger Kaufman and Joe Harless, both of whom figure prominently in the history of ISD and human performance technology.

ISD took root in the corporate world because it delivered results: it helped employees gain needed knowledge, skills, and attitudes faster, better, and cheaper than conventional approaches. For large and dispersed organizations, it provided a common training doctrine—a standard vocabulary and mindset—across geography and over time.

Systems Approach and the Audiovisual Field

The concept of systems approach probably was first introduced to the leaders of the educational technology field at the Lake Okoboji leadership conference in 1956. This annual conference, to which leading members of the state audiovisual associations were invited, often featured a keynote speaker, of whom perhaps the most influential was the first—Charles F. Hoban, who spoke to the 1956 conference on the topic, ''A Systems Approach to Audio-Visual Communication.'' As it happened, the conference's ''systems approach'' theme coincided with a series of articles by Finn published around the same time (for example, Finn, 1956) on a similar theme. These influences helped create interest in the idea of the systems approach, which eventually—about two decades later—became a hallmark of the field.

The vision that drove ''the systems view'' was expressed succinctly by Phillips (1966): ''to fashion a coherent assemblage of learning resources, specifically designed *from their inception* to be used with and make possible the implementation of a new curriculum'' (p. 373). The idea was to look at the education setting as a total system and to design a coherent package of hardware, software, manpower, facilities, and an implementation plan to most efficiently and effectively pursue the stated goals of the system. The link between the audiovisual education world and the systems world was also explained cogently by Gilpin (1962).

Among the earliest and most authoritative voices to reach educators with the ISD message were Robert Gagné (1962) and Robert Glaser (1962, 1965a). They were advocating instructional improvement from the standpoint of emerging psychological principles, but also placing these principles under a ''systems'' umbrella. These highly influential works are considered *precursors* of ISD inasmuch as they did not attempt to lay out specific detailed procedures or models for ISD.

Robert Corrigan and Roger Kaufman, both of whom had worked on Air Force programmed instruction projects and both of whom were affiliated with Chapman College in Southern California in the early 1960s, collaborated in the writing of a brief programmed monograph (1966) on the principles of systems engineering. Although it did not explicitly address educational issues, it made systems engineering concepts accessible to educators. Earlier, both authors had contributed influential papers to the first national conference sponsored by the National Society for Programmed Instruction (NSPI), the predecessor of ISPI, held in 1962 and reported in 1964. Corrigan's paper, ''Programmed Instruction as a Systems Approach to Education'' (1964), demonstrated that a teaching-learning situation employing PI could be viewed as an instructional system for individual learners, just as a classroom situation could be viewed as an instructional system for groups of learners. This perspective allows planners to restructure traditional educational environments, melding the most effective individual and group methods to create a more cost-beneficial hybrid, thus combining behavioral and systems engineering at the level of classroom organization.

Kaufman's paper at the same conference, ''The Systems Approach to Programming'' (1964), proposed that the production of programmed materials, previously viewed as a step in a psychological intervention, could be represented as a series of functions in a flow diagram, thus marrying behavioral and systems engineering at the level of lesson planning.

Exploring the Systems Approach in Higher Education

Southern California hotbed. In the early 1960s, Leonard C. Silvern, a senior scientist at Northrop Norair, was introducing systems engineering concepts at the University of Southern California (USC) as an adjunct instructor in the instructional technology department. There, James D. Finn presided over a program to which he had already introduced widely promulgated ideas about systematizing education (1956, 1957a, 1957b, 1960). Silvern had been working on instructional methods in the Navy since World War II, had done extensive research on fire and safety training, had become expert on the programming of teaching machines, and by the early 1960s was advocating systems engineering as an approach to designing learning environments (Silvern, 1963). In the spring of 1963, as an adjunct professor, he began offering the first course in applying the systems approach to instruction, ''Designing Instructional Systems,'' at USC. He also produced a detailed procedural model (1965) that influenced later model builders.

Robert Heinich, a graduate student in Silvern's first ID course, continued this line of exploration at USC with monographs on systems engineering of education (1965) and a dissertation that was later published (1970) by DAVI as a monograph, becoming one of the foundational works on applying systems thinking to education. Around the time Heinich completed his doctoral studies, there was a surge of interest among textbook publishers in branching into the publication of complete systems of instruction. Heinich became director of the Educational Systems Division of Doubleday and Company in 1967. During his two years there he produced a number of integrated learning systems consisting of films, audiotapes, and filmstrip sets. In 1969 he left to join the Instructional Systems Technology faculty at Indiana University, and later become the long-term editor of *Audio-Visual Communication Review*, which became *Educational Communication and Technology Review* under his guidance.

Back in southern California, at the School of Education at UCLA in 1962, Jim Popham taught the first college course on PI. With colleagues including Arthur Lumsdaine, Evan Keislar, Susan Markle, and John McNeil, Popham played a catalytic role in promoting research and development around PI (Popham, 1980).

Michigan State University: Instructional Systems Development project. During the late 1950s and early 1960s, the major academic programs in educational technology were groping from their roots in audiovisual media toward theoretical grounding in communication theory and learning theory. By the middle of the 1960s, systems theory was emerging as a potential place to stand to look at all the processes entailed in using learning resources in formal and non-formal education and training.

The Instructional Systems Development project, led by John Barson and headquartered at Michigan State University, was a multi-university demonstration and evaluation effort, testing a systems-approach ID procedure by applying the procedure to actual course development efforts during 1966 and 1967. The other collaborating institutions were Syracuse University, University of Colorado, and San Francisco State College. The collaborating researchers carefully documented time expenditures and costs associated with a systems approach to course development, reporting their findings in a final report (Barson, 1967). The heuristic guidelines and procedural model tested in this project (Figure 3.1) were widely disseminated and played a seminal role for later ISD model builders.

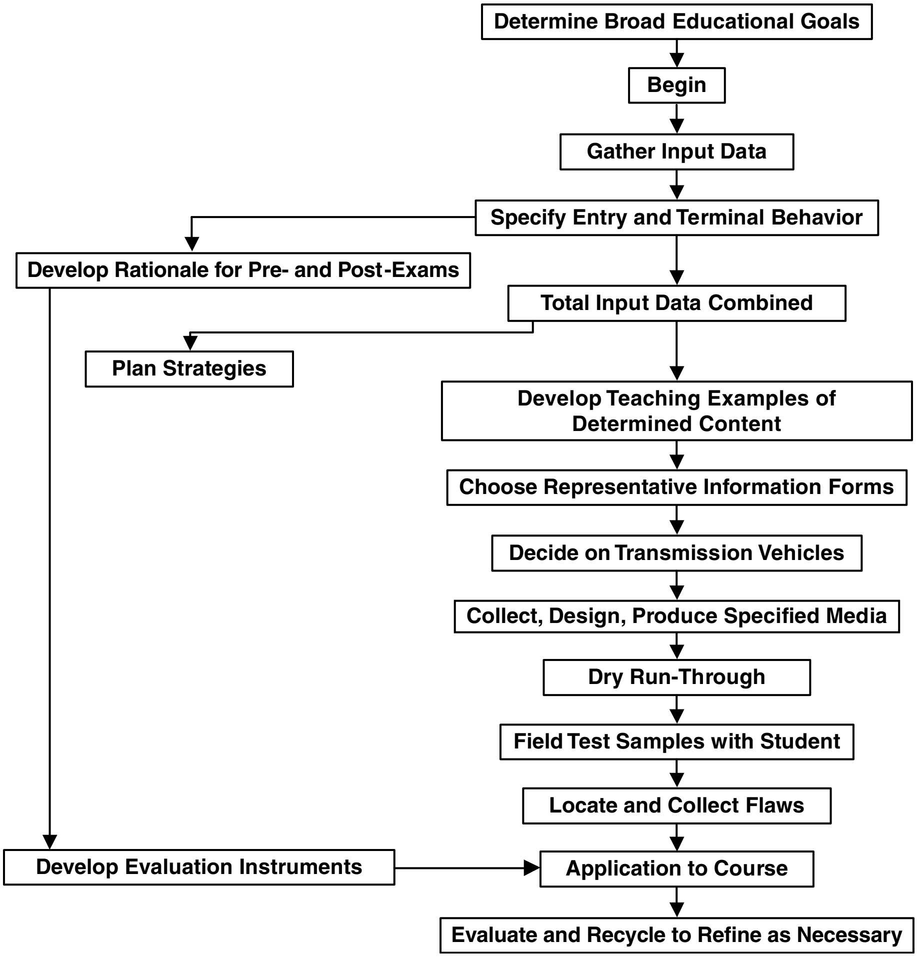


Figure 3.1 Facsimile of Barson Model.

Syracuse University. At Syracuse University, one of the Barson project's participating institutions, Donald P. Ely served as the head of both the academic program and the university audiovisual center from 1959 until the service center split off in 1971, becoming the Center for Instructional Development. This center was led for the next quarter-century by Robert M. Diamond, who became a nationally visible champion for ISD and its application to the improvement of college instruction (see Diamond, 1975, 1980, 1985).

The academic program area (there were no departments in the School of Education), dubbed *instructional communications* in 1963, was one of the national flagship programs, and its faculty sought to keep the program on the cutting edge. This included sponsorship of a conference in April 1964, ''To

Develop New Dimensions for Research in Educational Media Implied by the Systems Approach to Instruction.'' Led by Eugene Oxhandler, the conference proposed a new paradigm to guide inquiry about educational media (Oxhandler, 1965). Although the naturalistic research methodology envisioned in this conference did not become a standard approach in the near term, the spotlight on the systems approach added to the momentum that was gathering in the field behind the systems concept.

Indiana University. Indiana University's Audio-Visual Center was headed from 1942 to 1972 by the visionary L.C. ''Ole'' Larson, who supported explorations in applying the systems approach to college teaching. Larson was early to see the value of building an organization that cohered around a holistic theory. He bought into the systems view and forged an organization that had clearly designated functional units: research and analysis, development, production, evaluation, and implementation (note the congruence with the phases of the ISD process). It was no accident that the academic program took the name Instructional *Systems* Technology in 1969, at the recommendation of a committee chaired by Bob Heinich.

As director of research of the Audio-Visual Center, Henry Bern (1961) was among the first to advocate for the systems approach as a pedagogical methodology. A little later he was predicting a bright future for ''educational engineers'' (1967), echoing a concept proposed a quarter-century earlier by W. W. Charters (1945).

Working with faculty on course development projects in the Audio-Visual Center, Gene Faris and Richard Stowe generated an early ISD model that was tested during 1966 and 1967; this model would later be published by Faris (1968) as one of the first full-fledged ISD models.

Florida State University. As part of an effort to enhance research capabilities in the Florida State University (FSU) College of Education, Professor Russ Kropp established a center for research and development on computer-assisted instruction (CAI) and in 1966 brought in Duncan Hansen and Walter Dick as assistants. The center's contract with IBM included the training of CAI specialists, and Hansen and Dick established a series of courses to support this program. One course, designed and taught by Dick in 1967, focused on a systems approach to the development of CAI materials (W. Dick, personal e-mail communication, December 23, 2008). Dick also developed a visual-verbal model of the systematic design process, which was used at FSU in 1968 and later incorporated in modified form in the Dick and Carey textbook (1978), discussed below.

Over the next several years, Robert Morgan, Robert Gagné, Leslie Briggs, Robert Branson, and Roger Kaufman joined the nascent Instructional Systems program (W. Dick, personal e-mail communication, December 23, 2008), constituting one of the most prestigious academic programs in the nascent field.

The Systems Approach at Regional R&D Laboratories

The Elementary and Secondary Education Act of 1965 established a nation-wide network of twenty regional educational research and development laboratories. The Southwest Regional Laboratory for Educational Research and Development (later Southwest Educational Development Lab) in Austin, Texas, began by carrying out a number of curricular materials development projects, led by Richard Schutz and Robert L. Baker (coming from Arizona State University). Their experiences fed into a set of handbooks on instructional product research and development, one of which, by Baker and Schutz (1971), became used as an ISD textbook. Similar work was also being done at other regional labs, which were encouraged to use systems-approach procedures to produce high-quality instructional materials for use in schools. The Far West Laboratory in San Francisco was especially active in ISD. A project directed by Bela Banathy there produced a twenty-three-volume library of paperback programmed modules for each step in the ISD process, under the title *Training Resources*, published in 1975.

Although the federal support, virtually eliminated during the Reagan Administration, was not sustained for a long enough period to significantly impact school practice, the knowledge gained in these enterprises enriched the literature of ISD and demonstrated the feasibility of ISD as a replicable process.

CONCEPTUAL UNDERPINNINGS OF ISD: COGNITIVE LEARNING THEORY

Although it can be fairly claimed that the outlines of the ISD process derive predominantly from the systems approach and from behaviorist theory, it is equally true that the *instructional strategies* drawn upon by instructional designers derive heavily from cognitive learning theory. Hannum (2005), in telling the story of the IPISD project at Florida State University in the early 1970s, emphasizes that the developers were strongly influenced by the theories of Robert Gagné, who was then at Florida State (p. 11). Gagné's work had reached the educational technology field as early as 1962 with his seminal book, *Psychological Principles in System Development*. It continued with *Conditions of Learning* (1965), in which he introduced his Events of Instruction framework. Subsequent editions of this book showed a deft eclectic touch, harmonizing the findings from behaviorist and cognitivist research into a coherent whole.

During the period of the late 1960s, Jerome Bruner was the most visible representative of the cognitive orientation. His *Toward a Theory of Instruction* (1966) directly challenged the behaviorist paradigm, arguing instead that human learning is driven by active minds that are continuously seeking to make sensible meaning out of their everyday experiences. He led the development of one of the most influential curriculum development projects of those times, a humanities course, *Man: A Course of Study*, which was used widely in schools in the United States and the UK in the 1970s. It incorporated the discovery method, instantiated in pre-packaged sets of materials. The movement led by Bruner had a decisive impact on curriculum development in American schools, and likewise on instructional designers who served the school market. In an early survey of instructional designers, Hoban (1974) found that 59 percent of respondents used concepts from Bruner's theories in their work, second only to the 71 percent who used Skinner's theories (p. 463).

Of course, one limitation of the cognitivist perspective is that it offers solutions primarily for learning tasks in the cognitive domain—intellectual tasks. It offers little guidance to the achievement of objectives lying in the interpersonal, attitudinal, or motor skill areas.

CONCEPTUAL UNDERPINNINGS OF ISD: COMPUTER PROGRAMMING

The systems approach itself evolved out of analytical methods associated with the original general-purpose computers, so it should not be surprising that logic diagrams, process flow charts, and mathematical expressions were prominently visible tools in the early days of ISD. For example, the first ISD-type model to appear in the journal *Educational Technology* (Childs, 1968) consists of an elaborate flow chart, complete with activity blocks, decision blocks, and feedback loops to guide activities referred to as *programming*.

Not only were the analytical methods of computer programming influential in early ISD thinking, but computers themselves were envisioned as a delivery system for instruction virtually from the beginning. Although computer hardware in the 1960s was limited to centralized mainframe units, some educators and some computer specialists were convinced that computer-assisted instruction (CAI) could offer a cost-effective alternative to labor-intensive face-to-face instruction. By 1968 a number of experimental CAI programs were under way: ULTRA at New York Institute of Technology, TRAC at the Oregon College of Education, PLATO at the University of Illinois, and TICCIT in Reston, Virginia. Almost invariably, the initial instructional strategies used in these programs resembled those of PI, blending two of the major conceptual threads of ISD. However, the costs involved in delivering instruction via mainframe computers proved to be prohibitive, so programs such as these languished with little impact until the era of the microcomputer changed the cost equation.

THE EMERGENCE OF ISD AS A FULL-FLEDGED CONSTRUCT

Sparks in the Stubble

Glimmers of what would become a vision of a generic, systemic, and systematic process of creating instructional materials and environments can be glimpsed in the educational technology literature throughout the 1960s. Early versions of the systems approach appeared even in the popular education literature; for example, articles by Mauch (1962) and Bern (1967) in *Phi Delta Kappan* advocated the utility of viewing teaching-learning situations as systems, and thus amenable to deconstruction (system analysis) and restructuring (system synthesis) into new, more productive forms.

The systems approach gained visibility in the audiovisual instruction world when the third general session of the 1966 DAVI convention in San Diego was devoted to a pair of presentations by John Barson and Bob Heinich. Barson's presentation, ''The Systems Approach in Higher Education,'' summarized the work of his Instructional Systems Development project, described earlier. Heinich's presentation, ''The Systems Approach in Elementary and Secondary Education,'' recapitulated his dissertation findings, alluded to earlier. This event was important for raising awareness in a key sector—a large organization whose membership was still primarily focused on producing and delivering audiovisual materials to teachers and professors. The following year, the DAVI convention devoted a half-day and a half-dozen sessions to ''The Systematic Design of Instructional Materials,'' indicating a growing interest in this topic.

Reaching Critical Mass

By 1967 and 1968, the earlier conceptual sparks had been nourished by the fuel of federal support and big-business investment and were beginning to glow brightly enough to be noticed in educational technology and related fields. The various conceptual elements—PI as an application of behavioral psychology, systems engineering, and computer programming—were converging into a new compound, under the label of *instructional development* (sometimes *instructional design*). Several authors now were ready to propose systematic procedural models that laid out specific steps of lesson-development fully and in some detail, published in a venue that reached a wide swath of educational technology scholars.

The Barson project final report (1967) contained such a model, but was not published in a widely disseminated venue until later (see Figure 3.1).

Eraut's (1967) article was ''an attempt to summarize and to advocate a methodology for course development'' (p. 92), but his box-and-arrow charts described the overall strategy without giving a succinct procedural guide.

Bela Banathy (1968) provided a book-length treatment of the application of systems thinking to education and included a flowchart for ''The Design of

Instructional Systems.'' It is very close to the mark, but the elements in the flowchart are left at a rather general level and in the terminology of systems engineering rather than that of lesson planning; it does not explicitly take into account the psychological processes entailed in a learning episode.

Five publications from 1968 appear to meet the criteria for the earliest full-blown ISD model, including being widely promulgated and using the label *instructional design* or *instructional development*. John Childs (1968) provides a complex box-and-arrow flowchart, which he summarizes at the end as ''the procedural steps . . . in the process of instructional design'' (p. 14); and his twelve steps overlap well with conventional ISD models.

Gene Faris (1968) advocates for the job title of *instructional developer* and epitomizes the job by showing the course development model used at Indiana University—a box-and-arrow chart that contains the basic elements of ISD models. As it happened, this model was not developed further nor emulated to any extent, so could not be cited as particularly influential.

Haney, Lange, and Barson (1968) presented the Barson 1967 model, but now in a peer-reviewed, widely circulated research journal. It is a full-fledged ISD model and, further, the authors advocate a *heuristic* approach to its use—a perspective that was elaborated quite fully and effectively later by Romiszowski (1981).

The fourth of the five publications is a spinoff of the Barson model, developed by a team at Teaching Research in Oregon led by Dale Hamreus (Hamreus, 1968). The Oregon model is more detailed than Barson's—a box-and-arrow flowchart with twenty-two steps. It suffers a bit because of its complexity; the elements of the model were later reconfigured into a much more mnemonic arrangement in the form of the IDI model, discussed later. The Oregon model was quite widely known, although published only as a local report, then made available through the ERIC microfiche system. According to Gustafson and Branch (1997), it was used primarily ''by teams developing large-scale curriculum projects, a common activity of the period'' (p. 73).

The fifth candidate for the earliest complete and widely promulgated ISD model is the article by Roger Kaufman (1968) in the same issue of *AV Communication Review* as that of Haney, Lange, and Barson. His box-and-arrow charts depict a full systemic problem-solving process, and one of the charts is a credible procedural model for lesson development (Figure 4, p. 422). But it suffers a bit by comparison because it does not adopt the *instructional design* label, instead being captioned as ''A Possible Mission Profile for Preparing Instructional Materials Using a Systems Approach.''

These ''first'' ISD models helped disseminate ideas about ISD but had little practical impact outside the academic realm. The soon-to-be-developed IDI model, discussed in the next section, eventually was taught to teams of educators at hundreds of school districts around the United States and, in the process, disseminated widely in academia. The later IPISD model, implemented throughout the Department of Defense, would have a tremendous impact on later models developed in the business sector and in academia.

Instructional Development Institute (IDI). The earlier path-setting activities at USC, Syracuse, and Michigan State provided the foundation for a consortium formed around 1970 under the leadership of James D. Finn, composed of those three institutions plus U.S. International University. (Indiana University joined in the mid-1970s; Florida State, Arizona State, and Georgia became members later, after the era of the IDI project.) This consortium, initially known as National Special Media Institute (NSMI) and later as University Consortium for Instructional Design and Technology (UCIDT), worked together to develop, test, and disseminate a packaged training program on instructional design, funded under Title XIII of NDEA. The IDI was a fully programmed five-day workshop on ISD intended for teachers at the K-12 level. Between 1971 and 1974 it was offered to 300 to 400 groups of educators in the United States and later in several other countries. In 1976 it was expanded to seven days, adding units on evaluation and diffusion strategies developed at Indiana University.

IDI workshops were usually conducted by faculty and graduate students from participating universities, who later used the IDI materials in their own college courses. Thus the IDI became an influential vehicle for disseminating the IDI model (see Figure 3.2) and other workshop materials and methods

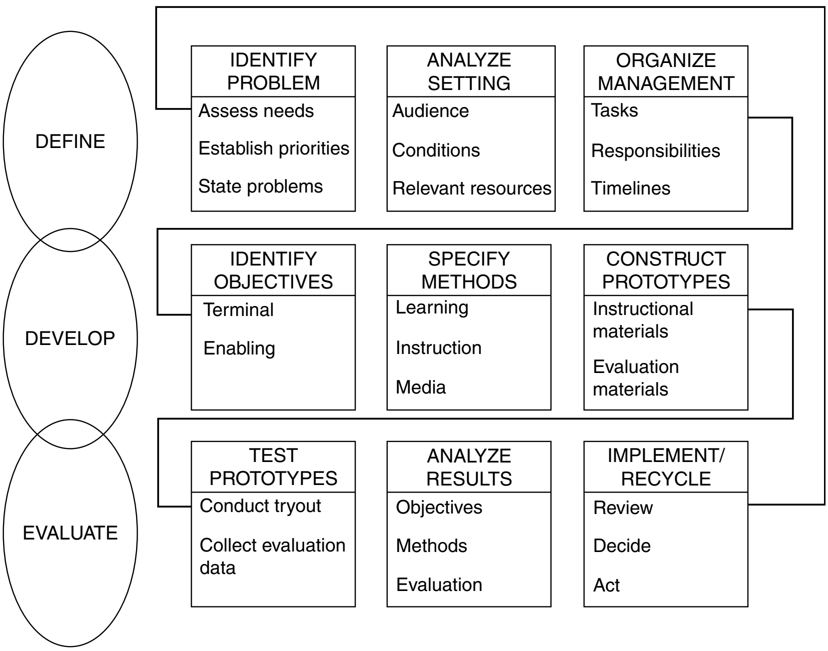


Figure 3.2 The IDI Model.

among educational technology faculty and students across the United States (Schuller, 1986).

THE INFLUENCE OF SCHOLARLY INSTITUTIONS

New theories and scientific constructs tend to be taken seriously to the extent that they are recognized by credible authorities and institutions. ISD, which emerged partly out of research and development within academia, was embraced quite quickly and fully by educational technology academics. Progressively oriented professors and staff members at leading audiovisual centers saw the potential of ISD as a new paradigm. For theorists, it offered a scientific foundation for thinking about the improvement of instruction, which was long the concern of audio-visualists and instructional psychologists. For practitioners, it provided an accessible, succinct methodology to guide the entire process of creating instructional materials or whole learning environments; it could be used as a road map for an instructional video production and also for the redesign of a whole curriculum. So there was a latent followership, ready to participate when leaders stepped forward to found institutional infrastructure for ISD.

AECT and Its Division of Instructional Development

The primary scholarly organization for what would become the educational technology field was until 1970 known as the Division of Audio-Visual Instruction (DAVI) of the National Education Association (NEA). At that time the NEA decided to restructure, requiring DAVI to reorganize as a free-standing professional association, the Association for Educational Communications and Technology (AECT).

ISD was not yet prominent enough to merit serious consideration as the name of the field, but it had a following that grew at an accelerating rate in the late 1960s. As soon as the new AECT organizational structure allowed the formation of special-interest groups, the first group to petition for ''division'' status was the Division of Instructional Development (DID) in 1971. Richard Stowe, a member of the staff of the Audio-Visual Center at Indiana University, became the first president, and two other Indiana University faculty members led a symposium at the 1971 AECT convention to explore ''a definition of instructional development'' (Davies & Schwen, 1971). In the subsequent three decades the DID continued to prosper, with a series of recognized leaders in the field serving as directors of the division.

The DID was from the beginning one of the largest divisions of AECT—*the* largest during many years. As such, it garnered a good share of the platform space at AECT conventions, sponsoring dozens of symposia, research reports, and other papers every year. This sort of institutional base is critical to attract scholars to invest time and energy in research and theory development. Without an outlet to report their efforts, interest would have surely waned.

Having a critical mass of dues-paying members also enabled the establishment of outlets for professional writing, another necessary component of academic sustainability. The most important of these outlets was the *Journal of Instructional Development*.

Journal of Instructional Development

Scholarly communication in the area of instructional design and development expanded greatly after 1977, thanks to the launching of a specialty journal, the *Journal of Instructional Development* (JID), in that year. The notion of a special-interest journal devoted to instructional design and development topics was championed by Kenneth Silber, then a professor at Governors State University, who had been deeply involved in formulating AECT's new definition of educational technology (AECT, 1977). He argued that the association's own definition of *a profession* required a high level of scholarly communication, a criterion that was not being met by the association's current lineup of journals (K. Silber, personal e-mail communication, September 17, 2008). His proposal for a new journal was accepted by AECT, and by the end of 1977, volume 1, numbers 1 and 2, of JID were published, with Silber as editor and John B. Johnson, also of Governors State University, as managing editor.

The journal attracted submissions from many of the leading scholars in the nascent ISD field and provided an outlet for others who were interested in doing research in this area but had been unsure whether there would be a place to publish their findings. Its referees included leading scholars in the field, and it maintained high standards for acceptance, typically accepting only about one-quarter of the manuscripts submitted (K. Silber, personal e-mail communication, September 17, 2008).

JID continued to be published quarterly through volume 11 in 1988. Unfortunately, the period of 1986 to 1988 found AECT struggling with unsustainable deficits, leading to the hiring of a new association manager, who undertook major cost-cutting measures. In 1988 the board of directors decided to merge JID with the other leading research journal under a new name and structure. The new *Educational Technology Research and Development* (ETR&D) would have two sections, each with its own editor, with Norman Higgins of Arizona State University as the first editor of the Development section.

During its heyday, JID served as an important forum for new ideas in instructional design and development, although its subscriber base did not extend much beyond the membership of the ID interest group in AECT. Nevertheless, key JID articles were widely cited in the educational technology literature, proof of its wider readership and scholarly impact.

Other Influential Associations and Journals

As discussed earlier, the National Society for Programmed Instruction had become the National Society for *Performance* and Instruction in 1973 in recognition that its members' interests now extended beyond programmed instruction to encompass a growing palette of strategies, formats, and processes to achieve results-oriented improvements in human performance. Its monthly periodical, *Performance and Instruction*, carried success stories, particularly from the corporate realm, along with new-and-improved ISD models. During this period it took the lead in exploring the ''front end'' of the design process—analytical methods to determine the source of performance deficiencies, be they susceptible to training solutions or not.

The American Society for Training and Development (ASTD) represented a larger and more general population than AECT or NSPI—corporate trainers, designers, and training managers. Its monthly journal, *Training & Development,* spoke mainly to training managers but did document the emergence and spread of the ISD approach to training design. Its annual survey of ''the state of the industry,'' along with the annual survey conducted by *Training* magazine, provide the best documentation of the adoption of new instructional media and methods over the years.

Although not published by a professional association, the monthly magazine *Educational Technology* exhibited consistent thought leadership in educational technology, not to mention admirable resiliency. Founded by Larry Lipsitz in the early 1960s to promote the study and dissemination of programmed instruction, it evolved into the most widely circulated and widely read periodical in the field of educational technology, still going strong into the 21st century. Lipsitz viewed the magazine as a platform for debate about emerging ideas regarding media and methods, especially instructional design, and it garnered more than its share of ground-breaking papers by major authors.

Codification of ISD in Textbooks

By the early 1970s the elements of a generic ISD process had jelled and were being codified in a form that could be communicated to many potential users through textbooks and handbooks. The 1970s saw the birth of a spate of textbooks that would help disseminate the ISD approach. The first in the market were Kemp (1971), Baker and Schutz (1971), and Gerlach and Ely (1971), although the latter textbook was primarily devoted to instructional media—putting media utilization into the context of a systems approach.

These textbooks were preceded by a number of monographs and paperback workbooks that presented many of the ISD elements, but in a bit more rudimentary or less widely marketed form—for example, Leslie J. Briggs' monograph, *Handbook for the Design of Instruction* (1970), developed at Florida State

University in 1968–1970 for an ISD course he was teaching there; it built upon Briggs' earlier work at American Institutes for Research (Briggs, Campeau, Gagné, & May, 1967). Another precursor to the fully developed ISD textbook was the series of programmed workbooks by W. James Popham and Eva Baker: *Establishing Instructional Goals* (1970a), *Planning an Instructional Sequence* (1970b), and *Systematic Instruction* (1970c).

The second round of textbooks arrived in the mid-1970s; examples include Gagné and Briggs (1974), Davis, Alexander, and Yelon (1974), Diamond (1975), Briggs (1977), and Dick and Carey (1978). In the following decades, new editions, particularly of the Dick and Carey textbook (Dick & Carey, 1985, 1990, 1996; Dick, Carey, & Carey, 2001, 2005) and Jerrold Kemp's—who was later joined by Gary Morrison and Steve Ross (Kemp, 1971, 1977, 1985; Kemp, Morrison, & Ross, 1994, 1998; Morrison, Ross & Kemp, 2001, 2003, 2007)—continued to offer updated ISD procedures to new generations of students of educational technology.

Codification of ISD Through Certification Standards

Although there is no national program of professional certification for instructional designers, since the 1970s the major professional associations have supported efforts in this direction. In 1977 AECT and NSPI formed a joint task force on certification. This task force evolved into a separate organization in 1983, the International Board of Standards for Training, Performance, and Instruction (IBSTPI), which issued a list of competency-based standards for instructional design in 1986 (*Instructional Design Competencies: The Standards*). The competencies described in the IBSTPI standards are very closely aligned with the generic ISD models of the time. Thus the publication and dissemination of these standards, which were accepted and promoted by two major professional associations, lent legitimacy to the ISD approach.

WIDENING DISSEMINATION OF ISD

ISD Models

By the late 1970s, the standard way of expressing prescriptions about the components of the ISD process and their sequencing was through an ISD model. Authors by the dozens proposed different variations on the basic systems approach model (Andrews & Goodson, 1980). These models tended to agree on the most fundamental components and their sequencing: analysis of the problem, followed by making design decisions, leading to the development of prototype solutions, which could be implemented on a pilot basis, then be evaluated before full implementation. This common core procedure of *a*nalysis, *d*esign, *d*evelopment, *i*mplementation, and *e*valuation gradually came to be referred to colloquially as the *ADDIE* process. This term was not used as a formal title and was not the name of any specific model or other procedural guide (Molenda, 2003); it was just a shorthand nickname used mainly in oral discussion. From the late 1980s onward it became the most commonly used label to refer colloquially to the ISD family of models.

By the beginning of the 1980s, there were enough ISD models on the market to justify scholarly analysis. The ERIC Clearinghouse on Information Resources commissioned a study by Professor Kent Gustafson at the University of Georgia. The resulting ''information analysis document'' (Gustafson, 1981) yielded a taxonomy of four categories: classroom ID models, product development models, systems development models, and organization development models, with multiple examples of each category. This sort of official recognition, like the textbooks and standards discussed earlier, gave further legitimacy to the notion of ISD models.

Differing Rates of Adoption of ISD

During this period of expansion, advocates for ISD attempted to promote its use in K-12 and higher education. These efforts were largely unsuccessful, possibly for reasons related to the social and economic dynamics of these institutions (see Martin & Clemente, 1990), exploration of which would go beyond the scope of this chapter. Within the educational technology academic community, by the end of the 1980s, skill in instructional design was viewed as *the* core competency of the professional working in higher education or being educated in one of the growing number of academic programs.

Meanwhile, ISD flourished in corporate and military training as a way to standardize design practices for more efficient and effective training. Large corporations, such as AT&T, IBM, NCR, and Motorola, adopted ISD as their corporate training doctrine, creating their own ISD models to guide training designers' work. The ISD concept was also disseminated by giant consulting firms such as Ernst & Young and Arthur Andersen & Co. (later Andersen Consulting, then Accenture in 2000) that offered training services to their clients. Client companies learned about consulting firms' ISD procedures and often decided to adopt similar practices within their own training programs. By the late 1980s, ISD had become the ''gold standard'' for corporate training design.

Arrival as the Reigning Paradigm

When a new definition of instructional technology was devised by AECT in 1994 (Seels & Richey, 1994), it was obvious that ISD had come to occupy the center of the stage. The core terms of the new definition were taken right from ISD: ''Instructional technology is the theory and practice of *design, development, utilization, management, and evaluation* [emphasis added] of processes and resources for learning'' (p. 1). It would be difficult to dispute that by 1994 the ISD construct had become the reigning paradigm in instructional technology.

QUESTIONING THE ISD PARADIGM

Research on ISD Process

Although instructional design was a popular topic for research during the period of growing popularity of ISD in the 1970s and 1980s, only a small fraction of instructional design research was devoted to the ISD process itself (Molenda, 1987). Most studies dealt with the variables associated with the ''design'' stage of the process—particularly, the efficacy of various instructional strategies and tactics. Few studies tackled the larger questions such as ''Does ISD work?'' or ''Who uses ISD?'' or ''What is needed for successful implementation of ISD?''

Two research syntheses published in 1986 (Ellson, 1986; McCombs, 1986) can serve to summarize the findings of those studies that did examine the whole ISD process. Ellson, looking for instructional treatments associated with major improvement in learning productivity, identified ''performance-based instructional design'' as one of the few educational methods that achieved a level of productivity that was at least double that of conventional instruction (p. 119). On the other hand, McCombs, in her synthesis of the early research on the ISD process, most of it done in the military services, emphasized the faults rather than the successes of the method. She found that what was done in the name of ISD often yielded unsatisfactory results, and she identified factors that were crucial to successful implementation of ISD. For instance, users noted that ISD models tended to be deficient in providing specific guidance on how to do each step. McCombs thus inferred that organizations intending to use ISD must ensure that their designers have the requisite skills to fill the gaps in the methodology (p. 72).

Corporate Pressures

By the late 1990s, however, an accumulation of pressures in the business world—including rapidly evolving digital technologies (see Liu, Gibby, Quiros, & Demps, 2002), intense cost competition with the accompanying need to reduce human resources costs, and the increasing pace of organizational change—led to a period of critical questioning of ISD orthodoxy. This dissatisfaction culminated in a lead article in *Training* magazine entitled ''The Attack on ISD'' (Gordon & Zemke, 2000). Experts quoted in the article charged that the ISD approach was too slow and clumsy for the fast-changing digital environment, failed to focus on what is most important, and tended to produce uninspired solutions.

Other critics in the corporate sphere argued that ISD should be viewed as a subordinate process within the larger process of performance improvement, on the grounds that training alone was never a sufficient solution to any training problem within an organization (Molenda & Pershing, 2004). It is this viewpoint that inspired this very handbook.

A Challenge to Underlying Instructional Theory

At the same time as business pressures were mounting, theoreticians and researchers were debating the merits of bold new (or recently rediscovered) claims regarding the fundamental nature of human learning . . . and how different methods of instruction did or did not fit with these new understandings. The debate took off when several influential scholars proposed a new paradigm for the design of instruction, which they called ''constructivism.'' The most frequently cited beginning of this movement is Bednar, Cunningham, Duffy, and Perry's ''Theory into practice: How do we link?,'' initially an occasional paper, later published in an anthology in 1991. These authors plus David Jonassen combined to write a number of manifestoes promoting this new paradigm (for example, Duffy & Cunningham, 1996; Duffy & Jonassen, 1992; Jonassen, 1991). They made strong claims about the invalidity of the psychological and philosophical bases of prior theories. These claims were difficult to evaluate because, first, the concept of ''constructivism'' was not clearly defined and, second, the proponents' examples of ''constructivist'' instructional prescriptions—situated cognition, anchored instruction, cognitive flexibility, problem based learning, cognitive apprenticeship, and everyday cognition—had previously been proposed by psychologists guided by cognitivist theories of learning, not constructivist philosophy. These definitional and labeling issues are discussed in depth by Robinson, Molenda, and Rezabek (2008).

Dave Merrill considered the shift of focus from a behaviorist to a cognitivist view of the learner to constitute a paradigm shift to what he termed ''second generation instructional design'' or ID_2 (Merrill, Li, & Jones, 1990). Other contributors to the debate took a less revolutionary posture, and simply proposed that the findings of cognitive psychology research could provide a great deal of guidance to instructional designers when it came to the *design* stage in the ISD process, the stage at which instructional strategies and tactics were being selected (Dick, 1997).

In many of the latter cases, scholars have actually been proposing new *frameworks* around which to organize lessons or instructional units—not new models of the total instructional development process (Molenda & Russell, 2006). Such frameworks specify the sequence of learning activities that should be incorporated into effective lessons. A familiar example is the Events of Instruction framework (Gagné & Medsker, 1996). Another even more detailed set of prescriptions is offered by Foshay, Silber, and Stelnicki (2003) as ''a cognitive training model'' (p. 23). The authors offer seventeen specific tactics organized around the various psychological stages of a lesson: gaining attention, linking to prior knowledge, structuring the content, presenting the new knowledge, and strengthening the new knowledge through practice and feedback. Many other prescriptive guides are discussed in detail in Reigeluth's comprehensive three-volume series on instructional-design theories (Reigeluth, 1983, 1999; Reigeluth & Carr-Chellman, 2009). The guidelines offered in these volumes revolve around which teaching-learning tactics to use, when to use them, and how to sequence them within the lesson.

Selection and sequencing guides and templates such as these should not be mistaken for procedural guides for conducting the entire planning process. Authors contribute to the semantic confusion when they label selection-and-sequencing guides or lesson frameworks as *models*. This label usually is and ought to be reserved for guides to the overall planning process.

In any event, in response to this constructivist challenge, theorists and practitioners have been busy exploring ways to design learning environments that place learners in realistic settings, that engage them in problem solving, and that give them greater ownership of the whole learning process. The ''constructivist'' movement coincided with the flowering of digital media that made it more feasible to create the sorts of interactive, exploratory, immersive environments recommended by this theory.

A Changing Digital Environment

In the late 1980s, as computing power multiplied geometrically and became more ubiquitous through networking, and as computer systems became more capable of offering multimedia presentations, they began to be seen as a new delivery platform: ''digital media.'' Just as the shift from audiovisual material production to television production entailed changes in the design process in the 1950s, so did the shift from traditional media to digital interactive media in the 1980s and 1990s (Jonassen & Mandl, 1990). For example, the increased complexity of interactive materials fostered concerns that such materials might be difficult for learners to use, to understand, or to accept; thus user-centered design and usability methods (Corry, Frick, & Hansen, 1997; Frick & Boling, 2002), borrowed from software design, became subjects of debate and study.

The rapid growth of the Internet and the World Wide Web in the 1990s presented instructional designers with another fundamentally different media environment in which to work. Web-based instruction, by its very nature, requires more learner-controlled activities, such as reading, writing, discussion, and reflection, as opposed to the teacher-controlled activities of the face-to-face classroom—lectures, demonstrations, and question-answer exchanges between teachers and learners. Thus, instructional designers had to think afresh about the sorts of instructional solutions to be created.

As the proportion of instruction delivered over the web increased, designers, particularly those in military and business environments, considered borrowing another concept from software engineering, the *object*. Proponents suggested that the use of reusable *learning objects*—‘‘small (relative to the size of an entire course) instructional components that can be reused a number of times in different learning contexts’’ (Wiley, 2002, p. 4)—was the key to gaining greater efficiency in churning out the thousands of hours of course material needed in the hundreds of web-based distance learning programs. It is not yet clear whether this particular approach will fulfill the dreams of its proponents, but the search continues for ways to automate the ISD process to the extent possible.

TO WHAT HAS ISD EVOLVED?

After the challenges of several paradigm battles and of adaptation to emerging media platforms, ISD continues to be a robust construct. After launching ‘‘the attack on ISD’’ (Gordon & Zemke, 2000), Ron Zemke later teamed with Allison Rossett to reconsider the criticisms raised in the original article. Their analysis (Zemke & Rossett, 2002) concluded that the flaws attributed to ISD lay more in how the process was executed rather than flaws in ISD as a theory.

That is, what is implemented in the name of ISD is not always in conformance with the canonical definition of ISD (and it could be argued that there is not a canonical definition). An example of this gap is the widespread failure to actually conduct formative evaluation, as specified in ISD theory. For years, ASTD has carried out an annual review of trends in corporate training, often asking survey respondents if they conduct formative or summative evaluation of learning gains from newly created instructional products. The response typically shows that about 40 percent of organizations do so (see, for example, Sugrue, 2003, p. 19). Thus, one of the key components of the ISD approach appears to be omitted much of the time. Other research studies document similar shortcomings in execution. These findings are consistent with those of McCombs back in 1986: that to be implemented successfully, ISD needs to be carried out rigorously, and that it needs to be conducted by people who are able to bring skill and creativity to the process.

There are others who feel that ISD, even if implemented adequately, still has blind spots that limit its suitability as the reigning paradigm. They suggest that design traditions in other disciplines—such as art, architectures, and software engineering—offer alternatives worthy of consideration (Bichelmeyer, Boling, & Gibbons, 2006; Molenda & Boling, 2008, pp. 119–122). Of particular current concern is the extent to which clients or users are involved in the design process. Carr-Chellman and Savoy (2004) discuss a range of design approaches from user-based, to user-centered, to truly user-controlled or emancipatory design, which they claim can be transformational for learners and the institutions in which they operate.

The psychological underpinnings of ISD have evolved over time as well. After two decades of debate about which is the ''one correct'' theory to inspire instructional design, there seems to be a new consensus, voiced well by Willis (1998), that an eclectic posture is warranted. As he points out, ''strategies developed within one paradigm are used by those who support another'' (p. 15), indicating that practitioners continue to adapt on a pragmatic basis. By observing how designers work, it appears that they intuitively adapt the process to the environment in which they work and the audience of learners they serve. For example, those who work with adult learners would more readily find value in a user-centered or participatory design approach.

CONCLUSION

The concept of ISD was created over forty years ago and has been evolving ever since. It is probably safe to say that ISD in practice will continue to evolve in response to changing social and economic forces, advances in understanding how humans learn, and new telecommunications technologies. Wallace Hannum's career retrospective (2005) summarizes aptly the confidence of ISD's proponents: ''Still the processes and procedures specified in the ISD model seem our best bet for developing and delivering high-quality training, regardless of how it is delivered'' (p. 19).

References

Andrews, D. H., & Goodson, L. A. (1980). A comparative analysis of models of instructional design. *Journal of Instructional Development*, *3*(4) 2–16.

Association for Educational Communications and Technology (AECT). (1977). *The definition of educational technology*. Washington, DC: Author.

Baker, R. L., & Schutz, R. E. (Eds.). (1971). *Instructional product development*. New York: Van Nostrand Reinhold.

Banathy, B. (1968). *Instructional systems*. Palo Alto, CA: Fearon.

Barson, J. (1967). *Instructional systems development. A demonstration and evaluation project: Final report*. U.S. Office of Education, Title II-B project OE 3-16-025. East Lansing, MI: Michigan State University. 125 p. (EDRS: ED 020 673).

Bednar, A. K., Cunningham, D., Duffy, T. M., & Perry, J. D. (1991). Theory into practice: How do we link? In G. Anglin (Ed.), *Instructional technology: Past, present and future*. Denver, CO: Libraries Unlimited.

Bern, H. A. (1961, July/August). Audiovisual ''engineers''? *Audio-Visual Communication Review*, *9*(4), 186–194.

Bern, H. A. (1967, January). Wanted: Educational engineers, *Phi Delta Kappan 48*, 230–236.

Bichelmeyer, B., Boling, E., & Gibbons, A. S. (2006). Instructional design and technology models: Their impact on research and teaching in instructional design and technology. In M. Orey, V. J. McClendon, & R. M. Branch (Eds.), *Educational media and technology yearbook 2006* (pp. 33–49). Westport, CT: Libraries Unlimited.

Branson, R. K. (1978, March). The interservice procedures for instructional systems development. *Educational Technology*, *18*(3), 11–14.

Branson, R. K., Rayner, G. T., Cox, J. L., Furman, J. P., King, F. J., & Hannum, W. H. (1975). *Interservice procedures for instructional systems development* (five volumes). Fort Benning, GA: U.S. Army Combat Arms Training Board. (NTIS Nos. ADA 019 486, ADA 019 487, ADA 019 488, ADA 019 489, ADA 019 490).

Briggs, L. J. (1970). *Handbook of procedures for the design of instruction*. Pittsburgh: American Institutes for Research.

Briggs, L. J. (Ed.). (1977). *Instructional design: Principles and applications*. Englewood Cliffs, NJ: Educational Technology Publications.

Briggs, L. J. (1980). Thirty years of instructional design: One man's experience. *Educational Technology*, *20*(2), 45–50.

Briggs, L. J., Campeau, P. L., Gagné, R. M., & May, M. A. (1967). *Instructional media: A procedure for the design of multi-media instruction, a critical review of research, and suggestions for future research*. Pittsburgh: American Institutes for Research.

Bruner, J. S. (1966). *Toward a theory of instruction*. Cambridge, MA: Belknap Press.

Bumstead, R. A. (1968, May). AT&T systems approach for love and money. *Training in Business and Industry*, *5*(5), 43–46, 62, 64.

Carr-Chellman, A., & Savoy, M. (2004). User-design research. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (2nd ed.) (pp. 710–716). Mahwah, NJ: Lawrence Erlbaum Associates.

Charters, W. W. (1945, February). Is there a field of educational engineering? *Educational Research Bulletin*, *24*(2), 29–37, 56.

Childs, J. W. (1968, August 30). A set of procedures for the planning of instruction. *Educational Technology*, *8*(16), 7–14.

Corrigan, R. E. (1964). Programmed instruction as a systems approach to education. In G. D. Ofiesh & W. C. Meierhenry (Eds.), *Trends in programmed instruction: Papers from the first annual convention of the National Society for Programmed Instruction* (pp. 36–45). Washington, DC: Department of Audiovisual Instruction, National Education Association.

Corrigan, R. E., & Kaufman, R. A. (1966). *Why system engineering*? Palo Alto, CA: Fearon Publishers.

Corry, M. D., Frick, T. W., & Hansen, L. (1997). User-centered design and usability testing of a web site: An illustrative case study. *Educational Technology Research and Development*, *45*(4), 65–76.

Crowder, N. A. (1962). Intrinsic and extrinsic programming. In J. E. Coulson (Ed.), *Programmed learning and computer-based instruction* (pp. 58–66). Hoboken, NJ: John Wiley & Sons.

Curl, D. H. (1967, March). Essentials of a training system. *Training in Business and Industry*, *4*(3), 37–41.

Davis, R. H., Alexander, L. T., & Yelon, S. L. (1974). *Learning system design: An approach to the improvement of instruction*. New York: McGraw-Hill.

Davies, I. K., & Schwen, T. M. (Eds.). (1971). *Toward a definition of instructional development*. Paper presented at the annual convention of the Association for Educational Communications and Technology, Philadelphia.

Diamond, R. M. (1975). *Instructional development for individualized learning in higher education*. Englewood Cliffs, NJ: Educational Technology Publications.

Diamond, R. M. (1980). Instructional development: One biased view (problems, issues, and the future). *Educational Technology*, *20*(2), 51–54.

Diamond, R. M. (1985). Instructional design: Systems approach. In T. Husén & T. N. Postlethwaite (Eds.), *International encyclopedia of education* (pp. 2558–2563). Oxford, UK: Pergamon.

Dick, W. (1987). A history of instructional design and its impact on educational psychology. In J. A. Glover & R. R. Ronning (Eds.), *Historical foundations of educational psychology* (pp. 183–202). New York: Plenum Press.

Dick, W. (1997). Better instructional design theory: Process improvement or reengineering? *Educational Technology*, *37*(5), 4–7, 50.

Dick, W., & Carey, L. (1978). *The systematic design of instruction*. Glenview, IL: Scott, Foresman.

Dick, W., & Carey, L. (1985). *The systematic design of instruction* (2nd ed.). Glenview, IL: Scott, Foresman.

Dick, W., & Carey, L. (1990). *The systematic design of instruction* (3rd ed.). Glenview, IL: Scott, Foresman.

Dick, W., & Carey, L. (1996). *The systematic design of instruction* (4th ed.). New York: HarperCollins.

Dick, W., Carey, L., & Carey, J. O. (2001). *The systematic design of instruction* (5th ed.). New York: Longman.

Dick, W., Carey, L., & Carey, J. O. (2005). *The systematic design of instruction* (6th ed.). Boston: Pearson/Allyn and Bacon.

Dills, C. R., & Romiszowski, A. J. (Eds.). (1997). *Instructional development paradigms*. Englewood Cliffs, NJ: Educational Technology Publications.

Duffy, T. M., & Cunningham, D. J. (1996). Constructivism: Implications for the design and delivery of instruction. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 170–198). New York: Macmillan Library Reference USA

Duffy, T. M., & Jonassen, D. H. (Eds.). (1992). *Constructivism and the technology of instruction: A conversation*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Dyer, R. H. (1969, March). AT&T Long Lines training centers are heavy on audiovisual and programmed instruction. *Training in Business and Industry*, *6*(3), 33–37.

Ellson, D. G. (1986). Improving productivity in teaching. *Phi Delta Kappan*, *68*(2), 111–124.

Eraut, M. R. (1967). An instructional systems approach to course development. *Audio-Visual Communication Review*, *15*(1), 92–101.

Faris, G. (1968, November). Would you believe . . . an instructional developer? *Audiovisual Instruction*, *13*(9), 971–973.

Finn, J. D. (1956). AV development and the concept of systems. *Teaching Tools*, *3*(4), 163–164.

Finn, J. D. (1957a). Automation and education: I. General aspects. *Audio-Visual Communication Review*, *5*(1), 343–360.

Finn, J. D. (1957b). Automation and education: II. Automatizing the classroom—Background of the effort. *Audio-Visual Communication Review*, *5*(2), 451–467.

Finn, J. D. (1960). Automation and education: III. Technology and the instructional process. *Audio-Visual Communication Review*, *8*(1), 5–20.

Finn, J. D., Perrin, D., & Campion, L. E. (1963). *Studies in the growth of instructional technology, 1: Audiovisual instrumentation for instruction in the public schools, 1930–1960, a basis for take-off*. Occasional Paper No. 6. Washington, DC: National Education Association.

Ford, G. A. (1970, July). Four steps are no longer enough: A systems approach adaptation of JIT for the 1970s. *Training and Development Journal*, *24*(7), 29–34.

Foshay, W. R., Silber, K. H., & Stelnicki, M. B. (2003). *Writing training materials that work: How to train anyone to do anything: A practical guide for trainers based on current cognitive psychology and ID theory and research*. San Francisco: Pfeiffer.

Frick, T., & Boling, E. (2002). Effective web instruction: Handbook for an inquiry-based process. Unpublished manuscript. Bloomington, IN: Indiana University.

Fund for the Advancement of Education. (1955). *Teachers for tomorrow*. New York: Author.

Gagné, R. M. (1962). Introduction. In R.M. Gagné (Ed.), *Psychological principles in system development*. New York: Holt, Rinehart & Winston.

Gagné, R. M. (1965). *The conditions of learning*. New York: Holt, Rinehart and Winston.

Gagné, R. M., & Briggs, L. J. (1974). *Principles of instructional design*. New York: Holt, Rinehart and Winston.

Gagné, R. M., & Medsker, K. L. (1996). *The conditions of learning: Training applications*. Fort Worth, TX: Harcourt Brace.

Galbraith, J. K. (1967). *The new industrial state*. Boston: Houghton Mifflin.

Gerlach, V. S., & Ely, D. P. (1971). *Teaching and media; a systematic approach.* Englewood Cliffs, NJ: Prentice-Hall.

Gilpin, J. (1962, March/April). Design and evaluation of instructional systems. *Audio-Visual Communication Review*, *10*(2), 75–84.

Glaser, R. (1962). Psychology and instructional technology. In R. Glaser (Ed.), *Training research and education*. Pittsburgh: University of Pittsburgh Press.

Glaser, R. (1965a). Toward a behavioral science base for instructional design. In R. Glaser (Ed.), *Teaching machines and programmed learning II: Data and directions*. Washington, DC: Department of Audiovisual Instruction, National Education Association.

Glaser, R. (Ed.). (1965b). *Teaching machines and programmed learning II: Data and directions*. Washington, DC: Department of Audiovisual Instruction, National Education Association.

Glaser, R. (1968). Educational technology as instructional design. *Educational Technology 8*(1), 5–6.

Gordon, J., & Zemke, R. (2000, April). The attack on ISD. *Training*, *37*, 43–53.

Gustafson, K. L. (1981). *Survey of instructional development models*. Syracuse, NY: ERIC Clearinghouse on Information Resources.

Gustafson, K. L. (1994). Instructional design models. In T. Husen & T. N. Postlethwaite (Eds.), *The international encyclopedia of education* (2nd ed.). (Vol. 2, pp. 2856–2862). Oxford, UK: Pergamon.

Gustafson, K. L., & Branch, R. M. (1997). Revisioning models of instructional development. *Educational Technology Research and Development*, *45*(3), 73–89.

Hamreus, D. (1968). The systems approach to instructional development. In *The contribution of behavioral science to instructional technology*. Monmouth, OR: Oregon State System of Higher Education, Teaching Research Division. (EDRS: ED 041 448. Microfiche only.).

Haney, J. B., Lange, P. C., & Barson, J. (1968, Winter). The heuristic dimension of instructional development. *Audio-Visual Communication Review 16*(4), 358–371.

Hannum, W. (2005). Instructional systems development: A 30-year retrospective. *Educational Technology*, *45*(4), 5–21.

Harris, R. F. (1964). Programmed instruction at Chanute AFB, Illinois. In G. D. Ofiesh & W. C. Meierhenry (Eds.), *Trends in programmed instruction*. Washington, DC: Department of Audiovisual Instruction, National Education Association.

Heinich, R. (1965). *The systems engineering of education II: Application of systems thinking to instruction*. Los Angeles: School of Education, University of Southern California.

Heinich, R. (1970). *Technology and the management of instruction*. Monograph 4. Washington, DC: National Education Association.

Heinich, R., Molenda, M., & Russell, J. D. (1989). *Instructional media and the new technologies of instruction* (3rd ed.). New York: Macmillan.

Hoban, D. (1974). The instructional developer. *Audio-Visual Communication Review, 22*(4), 453–466.

Instructional design competencies: The standards. (1986). Batavia, IL: International Board of Standards for Training, Performance and Instruction (IBSTPI).

Jonassen, D. H. (1991). Objectivism versus constructivism: Do we need a new philosophical paradigm? *Educational Technology Research and Development, 39*(3), 5–14.

Jonassen, D. H., & Mandl, H. (Eds.). (1990). *Designing hypermedia for learning*. Berlin: Springer-Verlag.

Kaufman, R. A. (1964). The systems approach to programming. In G. D. Ofiesh & W. C. Meierhenry (Eds.), *Trends in programmed instruction: Papers from the first annual convention of the National Society for Programmed Instruction* (pp. 33–35). Washington DC: Department of Audiovisual Instruction, National Education Association.

Kaufman, R. A. (1968, Winter). A system approach to education: Derivation and definition. *Audio-Visual Communication Review, 16*(4), 415–425.

Kemp, J. E. (1971). *Instructional design: A plan for unit and course development*. Belmont, CA: Fearon.

Kemp, J. E. (1977). *Instructional design: A plan for unit and course development* (2nd ed.). Belmont CA: Fearon-Pitman.

Kemp, J. E. (1985). *The instructional design process*. New York: Harper & Row.

Kemp, J. E., Morrison, G. R., & Ross, S. M. (1994). *Designing effective instruction*. New York: Merrill.

Kemp, J., Morrison, G. R., & Ross, S. M. (1998). *Designing effective instruction* (2nd ed.). Upper Saddle River, NJ: Merrill.

Lange, P. C. (1967). Introduction to section II, program development. In P. C. Lange (Ed.), *Programmed instruction: The sixty-sixth yearbook of the National Society for the Study of Education. Part II* (pp. 57–60). Chicago: University of Chicago Press.

Liu, M., Gibby, S., Quiros, O., & Demps, E. (2002). Challenges of being an instructional designer for new media development: A view from the practitioners. *Journal of Educational Multimedia and Hypermedia, 11*(3), 195–219.

Lumsdaine, A. A., & Glaser, R. (Eds.). (1960). *Teaching machines and programmed learning: A source book*. Washington, DC: Department of Audio-Visual Instruction, National Education Association.

Mauch, J. (1962). A systems analysis approach to education. *Phi Delta Kappan, 43*, 158–161.

Mager, R. F. (1961). *Preparing objectives for programmed instruction*. San Francisco, CA: Fearon.

Mager, R. F. (1962). *Preparing instructional objectives*. Belmont, CA: Fearon.

Mager, R. F. (1984a). *Analyzing performance problems, or, You really oughta wanna* (2nd ed.). Belmont, CA: David S. Lake Publishing.

Mager, R. F. (1984b). *Developing attitude toward learning, or, SMATS "n" SMUTS* (2nd ed.). Belmont, CA: David S. Lake Publishing.

Mager, R. F. (1984c). *Goal analysis* (2nd ed.). Belmont, CA: David S. Lake Publishing.

Mager, R. F. (1984d). *Measuring instructional results, or, Got a match?* (2nd ed.). Belmont, CA: David S. Lake Publishing.

Mager, R. F. (1984e). *Preparing instructional objectives* (rev. 2nd ed.). Belmont, CA: David S. Lake Publishing.

Markle, S. M. (1967). Empirical testing of programs. In P. C. Lange (Ed.), *Programmed instruction: The sixty-sixth yearbook of the National Society for the Study of Education* (pp. 104–140). Chicago: University of Chicago Press.

Markle, S. M., & Tiemann, P. W. (1967). *Programming is a process*. Sound filmstrip. Chicago: University of Illinois at Chicago.

Martin, B. L., & Clemente, R. (1990). Instructional systems design and public schools. *Educational Technology Research and Development*, *38*(2), 61–75.

McCombs, B. L. (1986). The instructional systems development (ISD) model: A review of those factors critical to its successful implementation. *Educational Communications and Technology Journal*, *34*(2), 67–81.

Merrill, M. D., Li, Z., & Jones, M. K. (1990). Second generation instructional design (ID2). *Educational Technology*, *30*(1), 7–14.

Molenda, M. (1987, March 1). *An agenda for research on instructional development*. Presented at annual convention of Association for Educational Communications and Technology, Atlanta. Available online at http://www.indiana.edu/~molpage/Agenda%20Res%20on%20ID.pdf. Retrieved November 15, 2008.

Molenda, M. (1997). Historical and philosophical foundations of instructional design: A North American view. In R. D. Tennyson, F. Schott, N. M. Seel, & S. Dijkstra (Eds.), *Instructional design: International perspectives, vol. I: Theory, research, and models*. Mahwah, NJ: Lawrence Erlbaum Associates.

Molenda, M. (2003). In search of the elusive ADDIE model. *Performance Improvement*, *42*(4), 34–36.

Molenda, M., & Boling, E. (2008). Creating. In A. Januszewski & M. Molenda (Eds.), *Educational technology: A definition with commentary*. Mahwah, NJ: Lawrence Erlbaum Associates.

Molenda, M., & Pershing, J. A. (2004). The strategic impact model: An integrative approach to performance improvement and instructional systems design. *TechTrends*, *48*(2), 26–32.

Molenda, M., & Russell, J. D. (2006). Instruction as an intervention. In J. A. Pershing (Ed.), *Handbook of human performance technology* (3rd ed.). San Francisco: Pfeiffer.

Mood, A. (1964, April). *Some problems inherent in the development of a systems approach to instruction*. Paper presented at Conference on New Dimensions for Research in Educational Media Implied by the Systems Approach to Education, Syracuse University, Syracuse, New York.

Morrison, G. R., Ross, S. M., & Kemp, J. E. (2001). *Designing effective instruction* (3rd ed.). Hoboken, NJ: John Wiley & Sons.

Morrison, G. R., Ross, S. M., & Kemp, J. E. (2003). *Designing effective instruction* (4th ed.). Hoboken, NJ: John Wiley & Sons.

Morrison, G. R., Ross, S. M., & Kemp, J. E. (2007). *Designing effective instruction* (5th ed.). Hoboken, NJ: John Wiley & Sons.

Nicely, C. S., Nelson, M. J., & Kaufman, R. A. (1970, May). A system for learning systems. *Training in Business and Industry*, *7*(5), 40–45.

Oxhandler, E. K. (1965, May). Afterthoughts on a systems conference. *Audiovisual Instruction*, *10*(5), 395–397.

Phillips, M. G. (1966). Learning materials and their implementation. *Review of Educational Research*, *36*(3), 373–379.

Popham, W. J. (1980). Two decades of educational technology: Personal observations. *Educational Technology*, *20*(1), 19–21.

Popham, W. J., & Baker, E. L. (1970a). *Establishing instructional goals*. Englewood Cliffs, NJ: Prentice-Hall.

Popham, W. J., & Baker, E. L. (1970b). *Planning an instructional sequence*. Englewood Cliffs, NJ: Prentice-Hall.

Popham, W. J., & Baker, E. L. (1970c). *Systematic instruction*. Englewood Cliffs NJ: Prentice-Hall.

Postlethwait, S. N. (1968). Audio-tutorial system. *Journal of Animal Science*, *27*, 938–940.

Postlethwait, S. N., Novak, J. D., & Murray, H. (1972). *The audio-tutorial approach to learning*. Minneapolis, MN: Burgess.

Quinn, A. K. (1970, February). In training, the system's the thing. *Training and Development Journal*, *24*(2), 25–29.

Reigeluth, C. M. (Ed.). (1983). *Instructional-design theories and models*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Reigeluth, C. M. (Ed.). (1999). *Instructional-design theories and models, Volume II: A new paradigm of instructional theory*. Mahwah, NJ: Lawrence Erlbaum Associates.

Reigeluth, C. M., & Carr-Chellman, A. (Eds.). (2009). *Instructional-design theories and models, Volume III: Building a common knowledge base*. New York: Routledge.

Reiser, R. A. (1987). Instructional technology: A history. In R. M. Gagné (Ed.), *Instructional technology: Foundations*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Reiser, R. A. (2001). A history of instructional design and technology: Part II: A history of instructional design. *Educational Technology Research and Development*, *49*(2), 57–67.

Reiser, R. A. (2007). A history of instructional design and technology. Chapter 3 in R. A. Reiser & J. V. Dempsey (Eds.), *Trends and issues in instructional design and technology* (2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Robinson, R., Molenda, M., & Rezabek, L. (2008). Facilitating learning. In A. Januszewski & M. Molenda (Eds.), *Educational technology: A definition with commentary*. Mahwah, NJ: Lawrence Erlbaum Associates.

Romiszowski, A. J. (1981). *Designing instructional systems: Decision making in course planning and curriculum design*. London: Kogan Page.

Saettler, P. (1990). *The evolution of American educational technology*. Englewood, CO: Libraries Unlimited.

Scott, D. T. (1960). Teaching high school physics through the use of films. *Audio-Visual Communication Review*, *8*(4), 220–221.

Schuller, C. F. (1986). Some historical perspectives on the instructional technology field. *Journal of Instructional Development*, *8*(3), 3–6.

Scriven, M. (1967). The methodology of evaluation. In R. Tyler, R. M. Gagné, & M. Scriven (Eds.), *AERA monograph series No. 1, Perspectives of curriculum evaluation* (pp. 38–83). Chicago: Rand McNally.

Seels, B., & Richey, R. (1994). *Instructional technology: The definition and domains of the field*. Washington, DC: Association for Educational Communications and Technology.

Shrock, S. A. (1991). A brief history of instructional development. In G. J. Anglin (Ed.), *Instructional technology: Past, present, and future*. Englewood, CO: Libraries Unlimited.

Silvern, L. C. (1963). *Systems engineering in the educational environment*. Hawthorne, CA: Northrop Norair Division of the Northrop Corp.

Silvern, L. C. (1965). *Basic analysis*. Los Angeles: Education and Training Consultants.

Skinner, B. F. (1954). The science of learning and the art of teaching. *Harvard Educational Review*, *24*, 86–97.

Skinner, B. F. (1965). The technology of teaching. *Proceedings of the Royal Society*, Series B, 1965: 162, 427–443.

Skinner, B. F. (1968). *The technology of teaching*. New York: Appleton-Century-Crofts.

Sugrue, B. (2003). *State of the industry*. Alexandria, VA: American Society for Training and Development.

A system to create training systems. (1971, September). *Training in Business and Industry*, *8*(9), 40–41.

Tracey, W. R., Flynn, E. B., Jr., Legere, C.L.J. (1967). Systems approach gets results. *Training in Business and Industry*, *4*(6), 17–21, 32–38.

Washington County Board of Education. (1963). *Washington County closed-circuit television report*. Hagerstown, MD: Author.

Wiley, D. A., II. (2002). Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. In D. A. Wiley II (Ed.), *The instructional use of learning objects*. Bloomington, IN: Agency for Instructional Technology and Association for Educational Communications and Technology.

Willis, J. (1998). Alternative instructional design paradigms: What's worth discussing and what isn't. *Educational Technology*, *38*(3), 5–16.

Zemke, R., & Rossett, A. (2002, February). A hard look at ISD. *Training*, *39*, 27–34.

PART TWO

ANALYSIS

One of the key tenets of HPT, and of the praxeology of ID, is the simple notion that before we design anything for a client or try to solve a problem, we do some analysis to clarify the nature and real causes of the problem.

As Molenda (see Chapter Three) and Ferond (Chapter Seven in the *Handbook of Human Performance Technology*) both point out clearly, one of the first and major contributions of the fields of HPT and ID is this notion of doing analysis at the beginning of a project.

Whether it is called front-end analysis, root cause analysis, performance analysis, needs analysis, needs assessment (or any of the multitude of terms authors have used to name the process(es) involved), authors have always argued that:

- Clients frequently misdefine and misdiagnose the causes of problems they call us in to solve.
- Clients frequently ask for solutions from us that will not solve the real problems they have.
- The time and money spent in analysis are recouped many times over when we are able to design the correct interventions based on our analyses.
- Analysis produces value-added results for clients.

There is not, however, universal agreement about the methods for, objects of, or timing of this analysis. Different authors favor different data collection methods. Some favor doing complete analyses before starting design, while others believe in rapid prototyping models that do brief analysis combined with brief design, repeated several cyclical times throughout the process. Most importantly, some authors believe in looking at external behaviors, while others believe the only true analysis is one that probes the mental models of experts. The chapters in this section explore all these issues.

Chapter Four: Collecting Analysis Data. This chapter provides a basic, but extensive and intensive grounding, in the different methods for collecting data for analyses. It is the most comprehensive presentation of how to collect data found in any recent publication. It presents and discusses the pro's and con's of collecting data from different sources. It explains the advantages, disadvantages, and procedures for collecting data using a great variety of possible methods. It explains the why's and how's of summarizing and reporting the data collected.

Chapter Five: From Performance Analysis to Training Needs Assessment. The author walks us through the why's, what's, and how's of these analyses. It begins by untangling the plethora of nomenclature associated with identifying and clarifying the performance problem. Then it explains the why's and how's of conducting both a performance analysis and a training needs analysis, providing both step-by-step guidelines and job aids.

Chapter Six: Behavioral Task Analysis. This chapter presents an up-to-date and important look at the ''old'' process of behavioral task analysis. This process, which was used exclusively in the early years of ID, before cognitive task analysis (CTA) came along, is still (as we explained in the Introduction to this volume) an important tool for identifying how experts perform on the job and may be a sufficient tool for analysis of certain types of declarative and procedural knowledge. We include this chapter to underline the notion that CTA does ***not*** replace behavioral analysis in the ID toolkit, but rather is complementary to it.

Chapter Seven: Cognitive Task Analysis. Here, the authors lift the mysterious veil that obscures the purpose and procedure for doing CTA. The authors present a clear and understandable description of cognitive task analysis and how to conduct one. They summarize several of the major classical methodologies for doing CTA and then present a method they have synthesized and implemented. When you finish this chapter, you will not only be able to converse intelligently about CTA with your colleagues, but actually be able to conduct one yourself.

CHAPTER FOUR

Collecting Analysis Data

Jeanne Hites Anderson

INTRODUCTION

''Help! Telecom product and service sales are down. I need a two-day sales training program!''

Sound familiar? Most performance technologists have heard this cry for help in some form. The manager sees symptoms of a performance problem, assumes that the cause of the problem is lack of skills and knowledge, and decides that training is the solution. We know we need some diagnostics to determine the actual cause of the problem before designing a solution. We also need to partner with our client to determine ways to achieve their goals through data-driven solutions. In the scenario above, this is the point when a performance analysis begins.

There are several points at which an analyst will collect data throughout the instructional design process:

- *Performance analysis* (a preliminary analysis of a performance problem and recommendation for the next step);
- *Needs analysis* (an analysis aimed at determining the gap between what the current performance is and what it should be. It identifies performance problems, their causes, possible solutions, and whether training is one of the needed solutions);

- *Learner analysis* (analysis of learning characteristics, including specific entry competencies, and general characteristics such as learning styles, academic, personal, and social learner characteristics);
- *Setting or context analysis* (characteristics of the setting in which the learner works, including technical, managerial, and environmental constraints to the application of new job skills);
- *Task analysis* (an analysis of how work tasks are performed, including breaking tasks (both observable and cognitive) into steps at an appropriate level of detail);
- *Formative evaluation* (evaluation conducted at a formative stage before training is delivered for the purpose of improving the training); and
- *Summative or confirmative evaluation* (evaluation conducted following training or performance interventions to make judgments about their efficacy, durability of outcomes, impact, and ROI).

These types of analysis will be covered in detail in the following chapters in the rest of the chapters of the book. For whatever analysis is needed, data will be collected. Data collection and analysis for performance improvement, when done well, uses multiple methods and data sources to explore ''how'' and ''why'' workplace tasks are currently performed the way they are and how to improve performance. Analysis is used to make informed decisions on performance contexts, processes, inputs and outputs, so the goal of the data collection is to systematically gather information that is reasonably trustworthy so the decisions will be sound.

The general outline for the process of data collection and analysis is:

1. Determine the questions that need to be answered.
2. Select data sources.
3. Select data collection methods.
4. Create data collection instruments (surveys, interview forms, observation forms, etc.).
5. Collect data that will provide a well-rounded picture of the performance problem (or training results, in the case of an evaluation).
6. Analyze data: Mining facts from a body of evidence involves deciding how to classify, weigh, connect, compare, organize, and display information. These choices are directed by the analysis questions, input from stakeholders, and the data available.
7. Report results and recommendations.

These steps can be used for any of the stages of the instructional design process in which data is collected. This chapter covers data sources and data collection methods for analysis of workplace performance.

APPROACHES TO ANALYSIS

There are several models for analysis. Most analyses in this field use a straightforward decision or goal-oriented approach, meaning they begin with goals, develop measures, collect data to measure the gap in attainment of those goals, and facilitate decisions on performance problems and interventions. This is the approach taken in this chapter. One other approach that bears mentioning is a backward or goal-free approach (Isaac & Michael, 1981; Scriven, 1974). In this approach, without knowing organizational goals, the analyst (often an external consultant) examines processes and outcomes (planned or not) and relationships that may have profound effects on an organization. Data collected in this manner can be used to validate findings from a goal-oriented analysis.

WHAT DO YOU PLAN TO DO WITH YOUR INFORMATION?

Two primary considerations for an analysis are *usefulness* (utility) and *feasibility*. The point of data collection and analysis is for data-based decision making. Best practice tells us that making useful decisions about performance problems includes looking at the organization as a whole as well as the individual performance problem. Those responsible for training and performance improvement need to be sure that their activities are strategic to their organization and contribute to achieving business goals (Finnegan & McCampbell, 1982, p. 16; Noe, 2008, p. 51). This can be done by collecting data that measures this contribution in general in addition to specific projects, bringing us to the issue of feasibility. The costs of the analysis and training design and implementation will need to provide value to the organization. The training department will need to focus on results and take a systematic approach to performance improvements in the larger context of the organization by setting priorities for action and addressing issues that will bring value to the organization. The benefit must be greater than the costs of the project. One way to achieve this is to determine the utility of closing the performance gap by examining: (1) the number of people the performance affects, (2) the length of time the need has existed and is likely to continue into the future, and (3) how critical the performance is (Burton & Merrill, 1977, p. 36). For instance, we do front-end analyses to reduce risk of intervention failure. We also need to consider the cost of ignoring the problem or of building an ineffective or suboptimal solution compared to the cost of addressing it (Kaufman, 1972, p. 43). Calculating costs versus benefits has the added benefits of reminding us to avoid the trap of spending too much time gathering too much data, which can result in paralysis by analysis.

If there are many stakeholders involved in the priority-setting process, it may be valuable to use a nominal group technique (described later) or similar process to achieve consensus on priorities. The goal is to identify the large, high-value performance deficiencies, address those with the appropriate training or other intervention, evaluate results, and report high-value results to stakeholders.

WHAT IS THE QUESTION?

The first step in the application of the data collection process for performance and needs analyses is to determine the questions that need to be answered, which depend on what led to the performance concern in the first place. Two prompts typically lead to performance and needs analyses; one is reactive and one is proactive. The sales scenario is an example of a reactive prompt—a response to a perceived performance problem. Current performers are not achieving the desired results. The second is a proactive prompt—anticipation of a change such as the introduction of new technology at the job site. Given the scenario above, we know we will need to collect data on the performance problem. One symptom of the problem, according to the client, is that sales figures are down, but there are many things we don't know. Is there really a performance problem? What is the ideal performance? How much of a problem is there? Whose performance falls short of the ideal? Are the sales persons new or did they previously perform well? What is the cause of the performance problem, obstacle, or barrier to performance? Has something significant changed, such as the technology they use to perform their jobs? Further conversation with the sales manager will allow us to make educated guesses, perhaps even hypotheses, and define the questions that need to be answered to verify or reject the guesses.

The output of a performance analysis for our fictional sales manager would include a description of the gap between the ideal and the actual performance, performance problem causes, and the target audience. The aim is to determine the performance need. Data will need to be collected and analyzed and solutions recommended. This may involve finding answers to several questions shown in Exhibit 4.1, which would be selected as they apply to the project. *Note:* These are questions to ask yourself, not questions to ask stakeholders.

The questions that guide the data collection and analysis need to be reasonable, answerable (or measurable), brief, clear as possible to the respondent (that is, using their language and frames of reference), and should have reasons, based on the conversation with the client, for being considered worth assessing. An example of such questions is listed in Table 4.1. Determining what we don't

Exhibit 4.1 Typical Performance Analysis Questions

1. What are people not doing that they should be doing?
2. What are people doing that they should not be doing?
3. What is the gap between the actual and ideal performance?
4. What are the causes for the performance problem?
   a. Is it a lack of skills and knowledge or some other cause?
   b. What are the obstacles or barriers to optimal performance?
   c. What motivational factors affect performance? What are the consequences or rewards for performance?
   d. Are incentives ineffective or absent?
   e. What environmental factors, tools, or processes affect performance?

If this is a proactive analysis, the following questions would also be asked:

5. What changes in the workplace are anticipated?
   a. Will new equipment or technology be installed? Will a new service be offered?
   b. What job tasks will an equipment, technology, or service change require performers to do in the near future?
   c. When will the change in job tasks take place? When will technology or equipment be installed?
   d. If software, how long will the current release be current?
   e. What is the importance of the product (equipment, technology, or software) to the client?
6. Is there consistency of views about items 1 through 5 above held by content experts, job performers, supervisors?

Table 4.1 Data Collection Brainstorm for Sales Associate Performance Problem

Analysis Question	*Potentially Relevant Data*	*Potential Data Sources*
Ideal sales performance	Sales quota, activity quota, financial quota, sales volume quota by unit such as individual salesperson, a sales territory, a branch office, a region, a dealer or distributor, or a district.	Sales management extant data: Sales unit planning documents, individual sales associate professional development plans, marketing plans. Data to be collected from sales associates, supervisors,

(*Continued*)

Table 4.1 (*Continued*)

Analysis Question	*Potentially Relevant Data*	*Potential Data Sources*
	Supervisor and customer perceptions	customers (observation, survey, interview)
Actual sales performance	Actual sales figures and activities. Supervisor and customer perceptions	Sales management extant data: Sales unit manager dashboards and quota reports, customer surveys. Data to be collected from sales associates, supervisors (observation, survey, interview)
What sales associates are doing or not doing	Data on activities leading to sales: Quota reports, summary call reports, opportunity progress reports, sales activity reports. Sales associate, supervisor, and customer perceptions	Sales management extant data: Quota reports, summary call reports, opportunity progress reports, sales activity reports, customer surveys. Data to be collected from sales associates, supervisors (observation, survey, interview)
Causes for performance problem: Knowledge/ skills; performance obstacles; motivation and incentives; environmental factors, tools, processes	Sales training attendance and evaluation. Observations of performers on the job. Unit planning documents. Sales associate, supervisor perceptions. Performance appraisals	Training attendance lists, reports and evaluations. Data to be collected by observation, survey, interview of sales associates, supervisors, customers

know and brainstorming some sources for that information is the place to start. In all cases, there are three initial considerations for data collection:

- What information is needed or what variable needs to be measured to address the analysis questions?
- How can the variable be measured?
- Where or from whom can information or evidence addressing the questions be found?

INFORMATION SOURCES

There are multiple sources and ways to collect data to answer analysis questions, including documents, people, and observations. Appropriate information sources vary by the type of analysis being conducted. In our sales scenario, one source of information or evidence for a performance analysis is the stakeholders, including the client or sponsor of the analysis, supervisors, customers, and job incumbents. Work products and other physical evidence are also useful sources of information for performance analysis. Should the project continue to a needs analysis, sources could include the client, current performers (high and low), their supervisors, content or subject-matter experts, and customers or others who receive the output (goods or services) from the performers. Internal documentation and published literature can also be useful.

Sponsor/Supervisor

Sponsors or supervisors supervise the people for whom the course is being designed. This group may include project leaders, team leaders, and senior staff in addition to supervisors and managers. By virtue of their position, supervisors should be able to provide information on employees' performance, and indirectly their needs. They should know *what* is expected on the job and be able to provide information on characteristics of target audience: current and desired job performance level, how tasks are or should be done, task frequency, criticality of the job (depending on how ''hands-on'' the supervisor or work expert is), organization-specific procedures, and ''real-world'' examples. However, this may not always be the case. They may not know, or they may have a biased or inaccurate view of current and/or desired performance, so data obtained should be validated by comparing it to that obtained from others. In addition, supervisors only provide one perspective, which may not be representative of all key informants.

Job Incumbents: Target Audience and Exemplary Performers

A target audience is a group of people who will be taking the training—people who have all the prerequisite skills and knowledge, but no or little knowledge about the content of the course. This group will be able to provide all of the characteristics of target audience information and their perceptions of what they think they need to learn. They are likely to have a biased view of their own skills and knowledge, and their perceptions of what they think they need may be inaccurate or nonexistent.

Performance of job incumbents may range from inadequate to highly productive. Observation of incumbents at various levels, particularly both high and low performance, is useful for needs analysis. Observing the high-level performers provides information on how a job might be done well. However, sometimes one extraordinarily talented individual can skew the data. It is best to observe at least three high-level, expert performers and determine what they have in common. It is useful also to observe novices, and competent ''journeyman'' performers (Jonassen, Tessmer, & Hannum, 1999, p. 260), to determine differences in work performance from the high-level performances. Tasks that are performed by knowledge workers and others using mostly cognitive processes are more difficult to observe than physical motions. Task listings or information processing analysis, described later in this chapter, are useful to capture the cognitive processes required for a successful job performance.

Expert performers are currently the exemplary performers in real-world settings of the job tasks. They may or may not be the same people as the content experts, and they may work for your company or for another company using or supplying the technology. They can provide information on how tasks are actually done on the job, real-world examples, deviations from standard operating procedures (SOPs), organization-specific procedures, and frequency and difficulty of tasks and learning data. Like other experts, they may be unable to articulate procedures or may not know ''why and how'' things work. In addition, they may not know the most current information, SOP, which procedures are generic and which are organization-specific. For new technology, expert performers may not exist or may be in such high demand they are not available to you.

Subject-Matter or Work Experts

Subject-matter experts (SMEs) are useful informants for task analyses. Their experience has given them domain knowledge on particular jobs that can also contribute to formative evaluation by validating course content. They often contribute to data collection through interviews or nominal group processes (covered later in this chapter). One disadvantage of SMEs is that their expertise, developed through learning and much practice, allows them to do many job tasks without occupying much of their minds. As a result, in describing job tasks, they may forget key steps because they are unaware that they are doing them automatically. Interviewing a team of SMEs in a group or using the nominal group technique is often an effective method to elucidate all the steps in a job task because they all have different viewpoints, but it is still important for the analyst to take the perspective of the novice and ask questions when steps are not clear.

Content Experts

Content experts are people who really know the content, who could ''write the textbook'' on the content. They may or may not be now doing job tasks; rather, they may have designed the product. Content experts are often good sources for most current content information, ''textbook'' procedures, definitions, examples, and descriptions of ''how and why things work.'' Like exemplary performers and work experts, content experts frequently have so internalized key concepts and procedures that they are unable to articulate them. Also, they are frequently unaware of practitioners' commonly accepted deviations from SOP or organization-specific procedures.

Training Supervisors/Trainers

Training supervisors and trainers provide or coordinate delivery of training to the target audience or may work for other companies providing similar training. They can provide information on ''characteristics of target audience,'' difficulty of learning, prerequisite concept or course data, and ''marketability'' information. Risks with training personnel may include a biased view of needs or a view of the world from a ''training'' rather than ''job performance'' orientation.

Customers

Customers are an excellent source of perception information on the outcomes of products or services. They generally have little insight to offer on processes. One key problem that has arisen in the last decade is obtaining data from consumers. The American Association for Public Opinion Research and others have noted that response rates have declined precipitously. Whether from being over-surveyed or other reasons, the refusal rate has gone up dramatically, requiring more effort and making surveys more expensive to complete successfully. It is sufficiently problematic that researchers are reexamining the notion that high response rate is the best indicator of validity and reliability (low non-response bias). Experiments have indicated that other indicators of quality should be considered when surveys have short periods and low response rates (American Association for Public Opinion Research, 2007). There is the danger that we may increasingly hear only from those with complaints. At this point in time, there is no evidence that the decline in response rate is equally dramatic for business customers.

Extant Data and Work Products or Artifacts

Advantages of extant data are that it provides cues to performance problems and may be quantifiable. The data exists already and can be examined repeatedly without disruption to performers' work processes. It is often focused on outcomes and may cover a breadth of time and events. However, obtaining access is sometimes difficult and may reflect bias of authors or those keeping the records. Also, it may be more helpful in analyzing past rather than existing work processes or anticipated changes in work tools or processes. Sometimes data is very technical, and an analyst may need to consult with a subject-matter or work expert when interpreting the data.

Work samples or artifacts are physical items actually produced during a particular task. Work samples, like extant data, may already exist and may provide valuable insights on performances when they are difficult to observe because tasks are mostly mental processes. They can provide an indication of outputs of tasks, examples of positive and negative results of doing a task correctly and incorrectly, and common errors made. However, there is no indication of how the output was produced, of the environment in which it was produced, or of tools or references used.

It is often important to consult with subject-matter experts on criteria for exemplary work products before examining them.

In addition to being sources for information on the performance under consideration, obtaining input from stakeholders is useful for gaining access to data sources and buy-in for a resulting performance intervention. Management also typically decides the value and priority of the problem in relation to all other performance improvement projects so that the focus of performance interventions such as training is on solving the high-value problems.

DATA COLLECTION METHODS

Data collection methods should always be linked to the analysis questions and suited to the available data sources. Commonly used methods in instructional analysis include observational methods (observation of performers, processes, and products), collection of perceptions and ratings (surveys, interviews, focus groups), and extant data analysis. When choosing data collection methods, Borg and Gall recommend balancing the following:

- **Utility:** The process must yield clear, timely, useful information on the analysis questions in order to make necessary decisions.
- **Feasibility:** The process must consider the availability of sources and be carried out within reasonable time, effort, and cost constraints.
- **Propriety:** The analysis must be conducted lawfully, ethically, and with context sensitivity so that it is acceptable by the individuals and organizations involved.
- **Accuracy:** The analysis must credibly reflect the performance situation and be as free as possible from error. Reliability and validity are measures of accuracy. Reliability is the degree to which a data collection tool (such as a survey, questionnaire, observation checklist, or audit protocol) yields consistent results over repeated observations. Validity is the degree to which a tool measures what it is supposed to measure. (1989, p. 739)

Effective analyses in this field are essentially case studies or multi-perspective analyses. Ideally, we would collect data from multiple sources, using multiple methods. There are many ways to learn more about a performance problem, including surveys, focus groups, or interviews of stakeholders; direct observation of performers (Stake, 1995, p. 44); examination of extant documents (including archival records) and physical artifacts (Yin, 1994, p. 80). It is unlikely that one information source will have answers to all of the questions that need to be answered. Also, there are strengths and weaknesses for each data source and data collection method, so using more than one method helps overcome the weaknesses of each. In addition, looking for consistency among multiple data sources will help ensure that the information is trustworthy.

Data Collection Process

Generally, the steps to creating data collection instruments are as follows:

1. Define analysis questions.
2. Define the target population, the accessible population, and the sampling method.
3. Locate similar existing instruments to use as a guide for developing your items.
4. Write items or questions that answer all the analysis questions (and contain no ''nice-to-know'' questions) while taking a minimum of time. Create and format a draft of the instrument.
5. Ask content experts to review the instrument for content, format, and appropriateness for the audience. Revise as necessary.
6. Pilot test.
    a. To pilot test for validity: try out the instruments with a few members of the target population and ask for feedback on wording, clarity, utility, length, and suitability. Revise until the instruments garner the information needed to answer analysis questions and until questions and directions are clear and unambiguous, so that everyone using the instrument interprets the questions or items the same way every time; this results in reliable data.
    b. To pilot test for reliability: try out the instruments with a few members of the target population and then administer the instrument again in a week to ten days.
    c. Compare scores for each question. A high percentage of agreement (70 percent would be acceptable) indicates reliability.
    d. Revise or eliminate questions receiving low agreement (Brinkerhoff, Brethower, Nowakowski, & Hluchyj, 1983, p. 103).

e. Note the length of time it takes to complete the procedure for any data collection involving respondents. Revise to reduce the time and burden on the person responding to the instrument until it is minimal while still answering the analysis questions. This can improve reliability and the response rate. A long or burdensome instrument may be set aside or refused by those who feel they have more important things to do. A high response rate is one indication of validity because the responses probably reflect the general population fairly and without bias.

f. Include the length of time required in the interviewer script, instrument directions, or a cover letter.

7. Conduct preliminary data analysis. This will allow a test of analysis procedures and estimation of the variability.

   a. Revise analysis procedures as necessary.

8. Collect the data.
9. Tabulate and analyze the data.
10. Report analysis and recommendations to client and other key stakeholders.

HOW MUCH DATA IS NEEDED?

It is important to avoid the time, expense, and analysis-paralysis that result from collecting too much data. So how much data is needed? The answer is simple; the smallest sample possible that will provide enough data to answer the necessary questions and ensure that the data provides valid and reliable answers to the analysis questions. Ultimately, we want to make data-based decisions on designing and implementing training and other performance interventions that show results which exceed the total costs of analysis, design, and delivery.

In the sales performance scenario, you might have several hundred sales associates across the country. Is it necessary to collect data from all of them? When you have a large number of data sources like this, one way to keep data collection time and expense manageable is to collect data from a sample of the sources rather than from all sources. The aim is to describe the group of performers in need of training as accurately as possible. This means that sampling needs to be done well so that there is similarity on the questions you are studying between the sample and the whole group (population) of a particular type of data source (performers, documents, and so forth). If there is a similarity, we can safely make generalizations about the whole population. In other words, a survey of a sample of performers should give you similar results to surveying all performers, and a content analysis of a sample of documents should give similar results to an analysis of all documents.

It is not possible to eliminate all error when sampling is used, but in general, the larger the sample, the lower the error. This means that the probability that the data is untrustworthy or unreliable is small. There are several approaches to sampling that might be considered. The first is a simple *random sample* in which each individual is chosen entirely by chance (randomly), such that each individual (or each document or other unit of analysis) has the same probability of being chosen at any stage during the sampling process. *Stratified sampling* is sometimes used when the population includes a number of distinct categories. A random sample is selected from each category separately. The main reason for using stratified sampling is to ensure that particular groups are adequately represented in the sample. *Convenience sampling* is the method of choosing items arbitrarily and in an unstructured manner from those available. It is a commonly employed method, although it is the least rigorous. *Snowball sampling* is a similar technique, where existing study subjects are used to recruit more subjects into the sample. Both convenience and snowball sampling may introduce biases.

How large a sample is needed also depends on several things, including:

- The nature of the data collection methods;
- The nature of the population; and
- The consequences if the decisions made based on the data are faulty.

Most data collection instruments are created by the instructional designer, because measures of the performance that have been tested for reliability and validity are seldom available or appropriate to answer your analysis questions. For this reason, a fairly *large sample* is preferable whenever possible. They are especially desirable when the sample must be divided into subgroups and the population is heterogeneous on attributes being studied.

In the scenario, if our sales force is highly heterogeneous, for instance, they have a variety of skills and experience, come from diverse cultures, or live in widely dispersed geographical areas, we may need a fairly large sample for a survey, observation, or other measure. This would also apply if we need to break the sales force into subgroups, such as those sales persons responsible for selling to businesses and those responsible for selling to consumers. Finally, we might need a larger sample if it is critical to ensure that the patterns we are seeing in the analysis are real and not the result of chance.

Small samples are used when the *data collection method* includes role playing, in-depth interviews, projective and other time-consuming measures. This might be done if, by these means, you can get data not obtainable by other measures. Small samples are also used when the *nature of the population* so dictates. For instance, there are often only a small number of managers who can provide data on the management perspective of a given performance problem. If they can spare the time, an interview might be an appropriate data collection method. Estimate the needed sample size by looking at similar studies done in the organization, if any, to see if the audience was very heterogeneous. If so, use a larger sample. If no data is available on the audience, estimate the heterogeneity when doing a pilot test.

PILOT TEST, VALIDITY, AND RELIABILITY

The term ''pilot test'' in this case refers to a small-scale version of the full-scale data collection, as well as the specific pre-testing of a particular research instrument such as a questionnaire or interview guide. A pilot study is an essential element of a good analysis. Many instructional designers have created instruments that seemed clear, but, upon testing, have found that there were ambiguities that could have invalidated their studies. Conducting a pilot study does not guarantee success, but it does increase its likelihood. During a pilot study, data collection instruments and procedures are tested and revised. This step is important no matter what type of data collection method is chosen. Edwin Van Teijlingen and Vanora Hundley (2001) list several reasons for a pilot test, including the following:

- Developing and testing data collection tools for validity and reliability;
- Determining if a planned study is feasible;
- Creating guidelines for data collection;
- Determining whether the guidelines are realistic and workable;
- Determining if logistical problems might occur using proposed methods;
- Determining if the sampling method will be effective;
- Estimating outcomes variability to determine if sample size is large enough. If the audience appears heterogeneous on the study factors, a larger sample can help establish that the variations are real rather than a result of selecting a sample not representing the whole target population (sampling error);
- Collecting preliminary data;
- Determining how much time, staff, and other resources are needed for a full-scale study;
- Evaluating the proposed data analysis methods for potential problems; and
- Persuading sponsors and other stakeholders that the full-scale study is worth supporting. (Van Teijlingen & Hundley, 2001, p. 2)

One primary aim of a pilot study is to test and revise the tool(s) and procedures to reduce bias, reactivity, and other threats to validity and reliability. *Bias* is a systematic distortion of information that can result in poor decisions. It can enter into the analysis process in several ways. For instance, when a survey sample is self-selected, rather than selected randomly, those who choose to participate may be different in some way than persons who chose not to participate. *Reactivity* happens when data collection procedures influence the outcome. For example, when supervisors collect focus group or survey data from employees, their answers may reflect a wish to humor the supervisor, rather than their true feelings. *Validity* is the result of designing an instrument that measures what it is intended to measure and is relevant to the analysis question it is intended to measure. *Reliability* is the result of designing the instruments and procedures to measure consistently. The pilot test allows us to head off problems before they can invalidate the study by generating findings that do not represent the true situation. Details on how to create and use particular types of instruments will be included with a discussion of each methodology.

For tools such as observation checklists, one way to increase validity and reliability is to make sure that coding of tasks is done consistently. For instance, if an instructional designer in the telecom sales scenario checked ''ongoing costs'' as a separate category from ''monthly service expenses'' in one paragraph, then coded it ''monthly service expenses'' when it occurred in the next paragraph, the data and the interpretations drawn from that data would be invalid. In the same way, if more than one person is involved in coding, all of them should consistently code the same, so it is important they are very familiar with the codes. Inter-rater reliability is a measure of the extent to which different raters agree when using the same tool to measure a concept. A simple method for estimating inter-rater reliability is comparing a sample of coding by each of the people involved and computing the percent of agreement. The practical reason for concern about high reliability is that it increases the possibility of making good recommendations and decisions based on the data analysis.

SOURCES OF DATA

There are many sources of information that you can consider: extant data, work products, observation of performances and processes using both obtrusive and unobtrusive measures, and people's perceptions and ratings gathered through surveys, interviews, and group processes. When selecting data sources, you will need to consider the advantages and disadvantages for each one. These will be described in the following pages, along with how to collect data from each source and an example of where these sources . . . where these sources were used in an instructional design project.

Extant Data

Allison Rossett defines extant data that might be used in an analysis as the ''records and files collected by an organization reflecting actual employee performance and its results (for example, sales figures, attendance figures, help desk [records], callbacks for repair, employee evaluations) (1999, p. 225). R. Murray Thomas goes on to add that this form of analysis involves ''searching through one or more communications to answer questions that the investigator brings to the search'' (2003, p. 57). Extant data analysis is also called document or content analysis (Pershing, 2002, p. 36). Records might be in the form of paper or electronic documents, or other media, and may include quantitative or qualitative information. They are not limited to verbal or textual records; images or video records can also be informative. An examination of records can help us focus on outcomes and accomplishments rather than on behaviors, and it should be done early in the process and repeatedly (Rossett, 1987, p. 48). For instance, it is possible to examine whether the criteria being used to measure department effectiveness are aligned with the overall organization mission and goals (Rossett, 1987, p. 51; Stemler, 2001, para. 4).

Information from extant data should be analyzed to identify patterns or trends over time and can be used to verify findings from other data sources and methods. This information may help shape later analyses by focusing inquiries on actual outcomes or, after gathering other data, another analysis of records may be conducted to verify findings from those sources or methods. For instance, an analysis of telecom sales over the past two years indicated the fall in sales began in January of the second year. This pattern suggests that the instructional designer should look for changes that might have precipitated the fall in sales in the work, the workplace, or the sales associate group near the end of the previous year. Also, content analysis can involve a variety of analysis approaches whereby communication content is categorized and classified (Gilmore, 2006, p. 820). When done well, extant data analysis allows us to sift through large volumes of data to produce objective, systematic, and quantitative description of the content of communications. The aim of this technique is to make inferences by identifying characteristics of the communications in the documents (Holsti, 1969, p. 14).

Advantages

- Speed: records stored at the workplace can be examined quickly.
- Inexpensive: there is no need to gather new information.
- Unobtrusive: the instructional designer can examine the records without interrupting work routines or environment of the performers. (Zemke & Rossett, 1985, p. 10)
- Often quantifiable and focused on outcomes.

- Coverage: may cover a long span of time and many events.
- Stable: can be viewed repeatedly, once access is granted. (Rossett, 1987, pp. 52–53)

Disadvantages

- Access may be restricted to authorized personnel. May require justification and negotiation for access. (Rossett, 1987, pp. 52)
- Records may have been kept poorly or selectively (Zemke & Rossett, 1985, p. 10), resulting in bias.
- Records may reflect bias of author(s).
- May not be informative when the performance problem is new technology, products, services, or systems.

How To. Extant data analysis begins with planning, locating the records, and gaining access. Then, the information is systematically sampled, the records are examined and the pertinent information collected (Rossett, 1987, pp. 56–57). The data collection process is as follows:

1. *Plan the analysis.* Begin preparation by writing an overview of the analysis objectives, procedures, and specific questions to be answered. Next, determine what information is needed to answer the questions. Ascertain what documents hold that information and how you will obtain access to the information. Establish how much information will be needed, how materials will be sampled, and how the analysis will be approached. Then an outline for the report can be prepared.

   a. *Determine what information is needed.* This step begins with the job in question and its outcomes. Rossett recommends focusing attention on problematic duties or tasks, ''what employees do, might do and the opportunities and challenges with which they are confronted'' (1987, p. 54). In what forms can the information be found? Job descriptions might seem like a place to start, but they are seldom detailed enough to provide the information needed for a performance or needs analysis. Information on the job procedures and expected outcomes of performance (both terminal and supporting or in-process outcomes) might be found in training materials, job aids, reference materials, recent memorandums to the performers, meeting agendas or minutes, progress reports, knowledge and skills test results, customer feedback, manager's dashboard figures (management information systems providing data on key performance indicators). Performance appraisals may contain relevant information, but as Ron Zemke and Thomas Kramlinger note, personnel files are ''less readily available'' (1982, p. 222) because they are confidential and it is difficult to obtain access to them, even with promises of confidentiality. It is important to establish a clear link between the objectives and analysis questions and the data in the list of likely documents. Some of the information will come in quantitative form and other information will be in qualitative form. Quantitative information is usually objective and measurable, focusing on how much or how many of a relevant attribute may exist. Qualitative information tends to be more subjective and often focuses on characteristics of something being described such as opinions and beliefs of stakeholders.

b. *Determine how much information is needed.* The next decision will be whether all the information will be scrutinized; if there are too many documents to examine them all, a sample of the documents should be examined. If a sample is selected, how will sampling be done? Will a simple random sample be selected, or some other sampling procedure?

c. *Select the coding and analysis approach.* The aim of extant data analysis is to identify themes or patterns in the data. It is critical to determine how the analysis will be carried out before beginning the data collection so that data can be selected in an appropriate form, but there are few fixed blueprints for selecting the analysis approach. Records are examined for the existence of certain words, facts, themes, or concepts (that is, performers and their actions, events) drawn from the analysis question and coded in order to make inferences about the performance in question. Categories are usually established and described before examining the records, but following some preliminary examination of the data, additional codes may emerge. After describing the categories, an instructional designer will create rules that will allow streamlining and organizing the process to code for exactly what is relevant. Developing a set of rules helps the instructional designer ensure that he or she is noting concepts consistently throughout the document.

d. *Create forms or spreadsheets.* Forms or spreadsheets can then be created to record the coding. With a paper form or spreadsheet, the instructional designer would place check marks by appropriate descriptions. Alternatively, when the documents exist in electronic format, codes can be entered into computer software designed for this type of work such as ATLAS.ti or NVivo. By entering one's categories, content analysis programs can ease the coding process and assist in examining large amounts of data quickly and efficiently. Concepts may be implied in a document as well as explicit, so it is important to clearly define both explicit and implicit concepts before the beginning of the coding process. Following coding, the data will be examined for patterns, trends, consistencies, and inconsistencies.

Objective data can be tallied and compared (Pershing, 2002, p. 40). This can include both qualitative and quantitative information. *Quantitative data* (numerical or has elements such as events or words that can be counted or coded into numerical form) may lend itself to statistical treatment (Yin, 1994, p. 103) such as that used with survey research (frequency analysis, cross tabulation analysis, grouping, or segmentation analysis). *Qualitative information* (data describing qualities) may need a more heuristic approach to identify themes and patterns. The information may be sorted into various classifications, put into matrices, flow charts, tables, timelines, or other schemes (Yin, 1994, p. 103). Table shells or other devices for specific data displays can be constructed at this point in planning. The goal of the analysis process is to come to know the data thoroughly, notice patterns, consistency, and inconsistency in evidence from different sources.

2. *Pilot test.* With code definitions written and forms created, the coding procedures are now ready to pilot test.
3. *Locate records and gain access.* At this point, those stakeholders who can allow access to the records need to be identified. Clear explanations of how the records will benefit the project are usually necessary. Rossett notes that it is often easier to gain access to *in-process* outcomes records than records of terminal outcomes, so it may be necessary to be very persuasive in order to obtain information critical to the analysis (1987, p. 55). Confidentiality agreements may be needed for personnel records, proprietary or sensitive information.
4. *Examine and analyze records.* Using the definitions for content analysis categories and any checklists or recording methods, sample the data and code the information. Then analyze the data using the analytical approach determined in Step 1c. It may be necessary to recode some items consistent with the revised coding scheme if additional categories are added during the coding process. After coding, examine the data for patterns. Place data in tables or other displays planned in Step 1. Note the frequency with which particular issues are found in the records and trends you have identified. Is the data consistent in all documents, or are there inconsistencies? Erika Gilmore recommends at this point to reflect on patterns identified in the data (2006, p. 827). With the luxury of a few days to reflect and synthesize, the instructional designer will begin to notice additional patterns, trends, consistencies and inconsistencies and

to explore alternative explanations for the findings. The resulting analysis and recommendations will be more thorough and rigorous.

5. *Report the analysis results and recommendations.* Keep in mind the audiences for the report and how they expect to receive the information. Reports are not always made in written form: managers may prefer a presentation or executive summary, but colleagues may need a full written report when they are expected to carry out report recommendations. Using the report outline prepared during the planning stage, draft the report and arrange for a review by key informants as a way of corroborating essential facts and evidence. The informants may disagree with the instructional designer's conclusions and recommendations, but they should agree on the facts. Their criticisms are invaluable in ensuring a valid report (Yin, 1994, p. 145).

Example. After a corporate acquisition and merger, telecom training team members accustomed to an audience of domestic employees and customers found themselves faced with an audience of global sales associates. Different faces in the classroom did not unnerve them, but puzzling student responses did. Learners did not take notes like domestic audiences and seldom asked questions during class. They also seemed uncomfortable with the informal environment demanded by U.S. domestic audiences, but more to the point, they were not consistently accomplishing the training objectives. Concerned that language was a factor, the training team began with an examination of sales training brochures to see whether they specified that the training would be delivered in English. The brochures did, indeed, say that training was in English, and departmental memos between sales and the training center noted that the trainees were all fluent in English. This problem was apparently more complicated than the training team first imagined. Several other factors could be at issue for trainees, including culture shock, culture, and learning style differences. The team decided to do a content analysis of documents, letters, memos, meeting minutes, student guides, instructor guides, evaluations, and other extant documents to determine the cause of the problem. The performance analysis indicated that the training session with the fewest intercultural communication barriers and the most success had been localized for delivery in Taiwan. The project was recommended for a full needs analysis because of the large number of trainees who were affected. Table 4.2 shows some of the content analysis categories used.

Analysis of Work Products

Analysis of work samples is similar to extant data analysis and follows the same basic procedures. Like extant data, work samples are inexpensive because they already exist; however, assessment may require subject-matter expertise. Work

Table 4.2 Content Analysis Categories

Document:		
Intercultural dimensions	*Questions*	*Tally*
Relationship of individual to the group: Individual/ Collective	Do students appear to prefer to work in teams?	
	Do students appear to prefer to work individually?	
	Do students appear to prefer to answer questions as a group?	
	Do students appear to prefer to answer questions as individuals?	
	Do students deny compliments or praise?	
	Do students accept compliments or praise?	
	Are courses evaluated well in spite of problems?	
	Are students frank in course evaluation about problems?	
	Are individualized instructional methods acceptable to students?	
	Do students prefer group instructional methods?	
Power distance	Do students show discomfort when instructor and trainee participate as equals?	
	Do students prefer an authoritarian teaching style?	
	Do students prefer when instructor and trainee participate as equals?	

samples are the products of work completed by the organization, whether end products or not. Products may include written materials, designs, layouts, blueprints, computer programs, web pages, training materials, manufactured items, and other concrete products.

Advantages

- May provide valuable insights on difficult-to-observe performances resulting from primarily mental processes.
- Speed: records stored at the workplace can be examined quickly.
- Inexpensive: there is no need to gather new information.
- Unobtrusive: the instructional designer can examine the artifacts without interrupting work routines or environment of the performers. (Zemke & Rossett, 1985, p. 10)
- Stable: can be viewed repeatedly, once access is granted.

Disadvantages

- Only appropriate where tasks produce visible work products. (Zemke & Kramlinger, 1982, p. 229)
- May not be informative when the performance problem is new technology, products, services, or systems.

How To

1. Plan the analysis
   a. *Select work products for examination.* Choose what and how many work products will be observed. These products might be produced during the process of work or may be the end products of the work task.
   b. *Define the characteristics or attributes to observe or measure.* These characteristics might be qualitative or quantitative. Sponsors, work experts, or high performers can often provide criteria for the analysis of the work product's requirements.
   c. *Select how data will be recorded.* Checklists or tally sheets are often used in an analysis of work products.
2. Conduct the observation.
3. Report results.

Observation of Performers and Processes

Observation data is collected through direct observation of behaviors and interactions under actual working conditions. These techniques are useful with an established job where existing job incumbents are currently doing the job and there are high performers doing the job to standards, but many others who are not (Zemke & Kramlinger, 1982, p. 184). It might also provide information on causes as well as symptoms of the performance problems among low performers. It can also be used to identify specific issues such as differences between high (exemplary) and low (deficient) performers on the basis of performance criteria such as processes or procedures used, effectiveness of communications and relationships with other personnel or customers, use of time and resources, and adherence to standards (Zemke & Rossett, 1985, p. 9). Observation of performers in their normal work environment can also provide information about broad issues such as work climate, organization culture or group dynamics. Observational data can be used to supplement and validate data from other sources, collected using other methods such as interviews, surveys and records analysis.

Advantages

- Observations, if unobtrusive, can allow the instructional designer to see the performance in real time and in context at the work site.
- Is best used for job tasks in which the performance is observable.
- The method is flexible, allowing the instructional designer to pursue interesting clues to performances that are not necessarily reflected in data collection instruments.
- Direct observation can be useful to get a picture of what actually goes on during the workday. Most other methods gather information indirectly, and what people say on surveys or in interviews or focus groups often do not accurately reflect actual behaviors.

Disadvantages

- May be obtrusive; employees who are aware of the observer may intentionally or unintentionally alter their behavior (Hawthorne effect).
- Biases of observers can influence data. Observers may see behaviors they expect to find and fail to perceive behavior they do not expect (observer bias).
- Recording everything that happens with several people can be an overwhelming task.
- Observations can be time-consuming, especially if the work is cyclical on a weekly, monthly, or seasonal basis.
- Observations are less effective with knowledge workers who spend much of their work time in thought.

How To

1. Plan the analysis.
   a. *Define the characteristics or attributes* (variables) to observe or measure. Narrow the focus by referring to analysis questions for pertinent variables. Variables may be:
      i. *Descriptive.* When the observer simply tallies behaviors by placing checks by a description, little inference is required so the observation tends to be reliable.

ii. *Inferential or evaluative.* If observers will need to guess about an internal state (such as feelings or motivations) from a behavior, they will need to do so with indirect evidence. For instance, an instructional designer may want to determine if a sales associate speaks with confidence or uncertainty. Judging the quality of a performance requires both inference and evaluation (e.g., quality ratings). Observations are more difficult while evaluating, and it is harder to get reliable data on a performance when using indirect evidence than when using simple descriptions of the actual behaviors.

b. *Determine how much observational information is needed.* Much like other forms of data collection, it is important to collect enough information to answer the analysis questions without collecting so much that the analysis is not cost-effective. However, it is important that we don't rely on observational data from a single performer that may not represent the group (high, average, or low performers).

c. *Select analysis approach.* At this point in the planning, it is possible to begin outlining the analysis report. Write an overview of the analysis objectives, procedures, and specific questions to be answered. Data analysis needs to be tied to the questions the analysis is to address. Selection of analysis methods may be done concurrently with selecting how the data will be recorded. Typical methods of analyzing observational data include frequency counts, sorting behaviors into classifications, matrices or arrays, flow charts, tables, timelines, or other appropriate schemes.

d. *Select how the data will be recorded.* Recording may use paper and pencil forms, one-way screens, camera, audio, or videotape (with respondent permission). Video recording permits later coding or audit for accuracy of data collection. Paper-and-pencil recording may use methods such as using checklists or writing a narrative description during or after the observation (when the observation was recorded). Observers may also use devices to assist in the observation, such as mechanical counters or stopwatches to capture data on frequency or duration of behaviors. To create forms for tabulation:

i. Define the task in *behavioral units* (sales associate uttered x# sentences to explain new cell phone concept) so that it is specific, observable, and measurable and differentiated from behaviors not under study. Locations and times for the observation may also be specified. Methods for tallying the behavior should also be determined. For instance, an observer may use a mechanical counter or mark behavioral units on a form to count behaviors.

ii. *Define time unit* (four hours of observed sales or ten observations, five minutes each, per hour). The observer may use a stopwatch to help keep track of observations that include frequency counts, rates, or duration.

a. A *frequency count* tallies the number of times a task is performed (i.e., number of positive reinforcements over a sample time period).

b. If *speed* is important for a given task, the observer measures the *rate* as shown in the sample observation form Table 4.3. This is similar to a frequency count but includes the length of observed time in the calculation. It is good practice to score only one behavior at a point in time.

c. *Duration* is a measure of the time it takes to complete a task.

d. *Continuous notation* of all behavior is probably the most common observation method in this field. It involves writing a chronological narrative of behaviors during the observation period. If the observations are done over an extended period of time, the observer effect (respondents being aware of observer, which affects behavior) may be reduced. Zemke and Kramlinger recommend that observers avoid sacrificing performer comfort by keeping a low profile, the observation unobtrusive, and remaining quiet. (1982, pp. 80–82)

e. Determine how the analysis will be reported.

2. *Pilot test* all data collection procedures and revise as necessary.
3. *Familiarize observers* with descriptors and practice observation until there is agreement among observers for consistency in observation coding reliability.

Table 4.3 Observation Form

Employee:

Task: Sales associate asks questions of prospects to determine problems that can be solved by cell phone products or services

Date	Time		Observation tally (number of questions)	Total
	Start	Stop		
Rate (count/length of time) =				

4. *Conduct the observation and analyze the data.* Analyze data using the methods planned in Step 1.
5. *Report the analysis results and recommendations.* Complete the report as outlined in Step 1.

Time Studies. Time studies are also called time and motion studies, productivity measurement, or work study. This is a type of observation that involves an analysis of a particular task in order to find the most efficient method of completing it, information for developing productivity standards, costing, staffing, scheduling, and a variety of other uses. This type of industrial engineering study is often used in businesses practicing lean manufacturing and is effective for analyzing work flow, processes, and repetitive tasks such as those on assembly lines. Sometimes these studies also use work-sampling, a method in which a worker self-reports what he or she doing at time intervals. This may be useful for studies in service industries like health care where workers move about a building (Finkler, Knickman, Hendrickson, Lipkin, & Thompson, 1993, p. 579). As concerns for the cost of health care have increased, time studies have been used to examine time use in the health care field to plan for better efficiencies.

An effort has been made to use a variety of technologies to track and record the data, such as videotape, proprietary handheld devices, and cell phones so that studies are less disruptive to the employees than methods that use physically present observers. The method has evolved to include innovative uses such as tracking traffic flow in public places like airports and transportation. Some industries have time studies that can be a good source of information for a task analysis (Zemke & Kramlinger, 1982, p. 32).

Example. Shortly after moving to a new facility, supervisors of customer service representatives (CSRs) at a telecommunication company noticed that calls on customer questions were taking too long before satisfactory answers were provided. This trend was alarming because it could affect product sales. A preliminary observation of the workspace identified one problem: not all CSRs had easy access to a full set of documentation. Only the most common issues were addressed in the online documentation available to the CSRs. Some of the documentation was provided in paper form in binders, one set for every four CSRs. An observation of a normal workday showed that most calls were addressed through the electronic resources, but occasionally a CSR had to get up and walk to another desk to use the printed documentation. It was also noted that the current linear arrangement of desks discouraged sharing knowledge and group problem solving among CSRs. Until full searchable documentation was available online, a simple rearrangement of the room into ''pods'' of four desks facing each other with shared documentation was used. This solved 90 percent of the problems, as measured by customer feedback, and no further analysis was undertaken. Figure 4.1 shows the observation form used.

Employee: ____________________

Performance: ☐ High ☐ Average ☐ Low

Task: Customer Questions Answered or Problems Resolved

Minutes Observed: __________

Date	Call Duration		Observation Tally								Total Calls	Comments
	Start time	Stop time	Listens carefully	Interrupts	Expresses empathy	Asks questions	Apologizes	Has client agree to	Recaps	Thanks customer		
Rate (count/length of time) =												

Figure 4.1 Behavioral Count Observation Sheet.

Unobtrusive and Obtrusive Measures

Also called non-obtrusive measures, unobtrusive measures don't require your presence in the context in a way that alters performer behavior. This method of data collection reduces the ''guinea pig'' effect and role selection. Some methods of data collection or measurement cause performers to react or change their behavior in response to their interpretation of what you expect. This will result in contaminated data. Reactive measures may also make some people more sensitive to certain issues. The tendency to agree doesn't happen with unobtrusive or non-reactive measures. Extant data analysis and analysis of work samples are examples of unobtrusive measures. Indirect measures also have the advantage of being non-reactive or unobtrusive. Indirect measures are measures of something that occurs naturally within the work context. For instance, to measure paper recycling in a workplace, an instructional designer doesn't need interviews or surveys. He only needs to observe the number of pounds of recycling being sent to the recycler. However, if he wants to know attitudes toward recycling, asking people through focus groups, surveys, or interviews would be superior. Unobtrusive measures are often useful to compare with data gathered using surveys, interviews, or focus groups.

Task Listing. Task listing is a commonly used method of collecting job task information for task analysis from a panel of subject-matter experts in which a hierarchy of tasks is described and enumerated (Zemke & Kramlinger, 1982, p. 35–37). Other approaches to task analysis include S-R tables and behavioral algorithm task analysis. S-R tables, developed by William Deterline, involve identifying overt (actions) and covert (thinking) behaviors through observation, and then interviewing high performers while they work to verify the rules or cognitive guidelines they use for their work. Behavioral algorithms, like S-R tables, are an approach to task analysis that seeks to identify all rules used to solve problems. It is an information-processing approach to tasks that may involve complex mental processes. The result is a flow chart of the processes, and decisions used in completing a task (Zemke & Kramlinger, 1982, p. 48).

PERCEPTIONS AND RATINGS

Surveys, interviews, and group processes are ways of collecting perceptions and ratings from people. Creating effective questions for these sources of data is key to collecting valid data to support instructional design decisions.

Surveys: Paper, E-mail, and Web

A questionnaire or survey is a data gathering instrument through which respondents answer questions or respond to statements in writing. If the data sought is opinions rather than facts, it is frequently called an ''opinionnaire.''

The terms questionnaire and survey can refer to many different types of research instruments consisting of series of questions and other prompts for the purpose of gathering information from respondents. They may gather factual information (as perceived by the respondents) or they may survey opinions, beliefs, or attitudes. They may be conducted one time or several times in succession, and they are useful when the populations are too numerous to be directly observed (Watson, 1998, p. 31). Tests or inventories are sometimes considered as a special category of survey. Surveys are among the most often used methods of data collection because they can provide invaluable information on perceptions, attitudes, and beliefs. However, they are often done poorly, which can invalidate your findings. Good surveys gather information with practical significance to the analysis objectives, gather information that cannot be gathered in another way or from other sources such as extant document analysis, are neatly laid out and easy to use (have clear and complete directions and clear questions), contain no leading or biased questions, have questions in a reasonable order, and finally, are easy to tabulate and interpret. Willem deVries notes that the data collection needs of business and government agencies place a burden on respondents. This response burden has resulted in pressure to make the data collection less burdensome (deVries, Keller, & Willeboordse, 1996, p. 199). Using short, concise surveys that require a minimum of time to fill out is more likely to get the data you need from overburdened employees.

Advantages

- Good at gathering perceptions and attitudes
- Gathers data from a widespread audience relatively quickly
- Anonymity allows frank answers

Disadvantages

- Difficult to get at problem causes
- Answers cannot be probed
- Response burden
- Poor response rates

How To

1. *Develop survey objectives.* Examine the analysis questions and use them to guide the development of survey objectives.
2. *Identify the population.* Who has the information you need? Whom do you want to gather data from? Unless there is a small population, like a class of students where all members can be surveyed, sampling may be necessary. How many people to survey, how the sample will be

selected, and how the population will be accessed are decisions that you will need to make.

a. Determine sampling method.

b. Access the population. Access to the population or sample may use mailing lists or e-mail lists.

c. Outline report and select analysis approach. At this point in the planning, it is possible to begin outlining the analysis report. Write an overview of the analysis objectives, describing procedures and listing specific questions to be answered. Data analysis needs to be tied to the questions the analysis is to address. Selection of analysis methods may be done concurrently with selecting how the data will be recorded. Typical methods of analyzing survey data include frequency counts; sorting responses into classifications, matrices, or arrays; creating tables, graphs, or other appropriate schemes.

3. *Develop item pool.* The third step is developing a pool of questions (items) that will meet the survey objectives. Select the items that will best meet all of the objectives while taking a minimum of the respondents' time. To be sure all the objectives are covered, note which survey items meet which objectives. Avoid any items for which the survey respondents have no information, and make the items relevant and significant to the respondents, whenever possible. Include directions for completing the items. More information on how to write survey and interview questions is shown in Exhibit 4.2.

Exhibit 4.2 Tips for Writing Effective Survey or Interview Items

Closed-form questions are preferable to most respondents. Open-ended questions place more of a burden on respondents, response rate is poor, tabulation cost is high, and coding can be unreliable. You can avoid falling into the trap of using open-ended questions in an effort ''not to miss anything'' by conducting a focus group discussion (discussed later in this chapter) and using the results to help you formulate your survey. Then pilot test the survey. These steps will reduce the number of open-ended questions you will need to use.

Closed-form items require a short response from a limited or restricted set. Examples of closed-form items are (1) select response from a list of alternatives, (2) select yes or no, or (3) rated item or ranked list. Their major advantage is that they are easy to tabulate. If you need to provide for unanticipated responses use ''other'' and then ask for a description.

Example of a Closed-Form Question

1. How did you hear about this training program?

   ______a. The company website

   ______b. The training catalog

   ______c. A training department mailing

   ______d. A friend

   ______e. I have taken other management training courses from the training department

   ______f. Other (please describe)________________

Open-form items ask for a response in the respondent's own words. The advantage of this type of questions is that it allows for a response that you may not have expected or greater depth of response. A disadvantage is that it requires more of the respondent, so it may result in fewer returns. This type is also harder to tabulate and interpret.

Example of an Open-Form Question

1. How did you hear about this sales training program? ____________

Most surveys combine both open- and closed-form questions.

4. *Write introduction and directions.* The next step is to develop an introduction and directions, put the survey in order, and then lay out the survey form.

   a. The introduction includes the purpose of the survey and overall survey directions, including information such as to whom to return the survey. This goes at the top of the survey, unless this is a mail survey with the information in a cover letter.

   b. Put the survey in order. A survey should be in reasonable order to the respondent. This often means begin with general items and move to specific, which helps the respondents organize their thoughts. Also, avoid placing sensitive or intimate questions first, and try to avoid embarrassing questions entirely. More information on sensitive questions is described in the survey and interview writing style tips in Exhibit 4.3.

   c. Lay out the survey so it is attractive, and design it so it is easy to follow and easy to tabulate. Numbering items and alternative responses helps respondents or interviewers follow the form. More information on designing interview and survey questions is found in Exhibit 4.4.

Exhibit 4.3 Tips for Survey and Interview Writing Style

- Avoid double negatives.
- Underline words needing emphasis such as *not* and *or*, which change the meaning of a sentence.
- Define terms that could be easily misinterpreted (for example, *biweekly* could mean once every two weeks or twice a week).
- Do not talk down to respondent or interviewee.
- Write at readability level 6 (simple words, keep it short).
- Keep references general and in third person.
- Each item should be about a single idea. Combining ideas results in double-barreled questions.
- Avoid leading questions.
- Avoid unwarranted assumptions (for example, asking "Are you satisfied with your ability to gain support from your supervisor for your recommendations?" assumes that the respondent is in a position to make recommendations to the supervisor, which may not be the case).
- Phrase questions so they are appropriate for the target respondents.

Exhibit 4.4 Tips for Survey and Interview Question Design

SURVEY AND INTERVIEW QUESTION DESIGN

As you design your questions, you need to think about how responses will be summarized. For instance, the type of question that requests respondents to select all that apply are hard to summarize. One solution is to ask respondents to rank items. Then items can be weighted by rankings. When there are open-ended question responses to analyze, you will need to code the question and create a code list for the answers. Start the coding with 1 rather than 0. If no response is given for a question, it can be tabulated as a 0.

Example

In this two-part question, have 1 = YES and 2 = NO, and if a respondent answers yes, you can code the second part and it can be included in the results:

1. Do you determine which problems you can solve for your prospects?

   Yes No

   1 2

______If yes, please describe.

Design questions that will obtain a complete response.
Write clear and complete directions for item.

Writing Alternatives

- Stems should be single phrase relating to all alternatives.
- Alternatives should use parallel construction.
- Group ideas together, but make sure there is no overlap in categories.
- Items should be mutually exclusive.

Example

1. Age ___ 1) 5–10 ___ 2) 11–15 ___ 3) 16–20
2. Salary ___ 1) $1,000–$5,000 ___ 2) $5,001–$10,000
3. Years ___ 1) 1–5 ___ 2) 6–10 ___ 3) 11–20 ___ 4) 21–25

- Provide adequate and plausible alternatives.
- Keep the number of alternatives consistent where possible.
- Provide a line for comments.
- Make the response alternatives mutually exclusive (which do I check, "5,000 to 10, 000" or "10,000 to 50,000" if my work location has 10,000 employees?).

A long list of alternatives (more than six) can make it difficult for the respondent to remember and select an alternative. On the other hand, when asking the respondent to select responses from too few alternatives, you may choose to use an open-ended question and classify responses after surveys are returned.

Writing for Response Tabulation

For paper surveys and interview guides, write questions for easy tabulation of responses:

- Number items and response alternatives, especially if a person will be tabulating the survey responses rather than a computer. It is easier to keep track of letters than numbers.
- All responses can be coded using coding sheets or some other convenient scheme.
- Clearly indicate how the respondent should mark his or her response. For instance, on a paper survey, a blank line before the responses indicates where a respondent should mark from a selection of alternatives.

Example

1. Question . . . . . .

(*Continued*)

Exhibit 4.4 (*Continued*)

_______1 . . . . . .

_______2 . . . . . .

_______3 . . . . . .

_______4 . . . . . .

If respondents should circle an answer, people tend to circle, check, or X things quite large. Try to keep the area well spaced.

1. Question . . . . . .

   1. choice a 2. choice b 3. choice c 4. choice d

d. Each survey or individual record should be uniquely identified in some way, such as a number so it is possible to track who to send reminders to, but make sure the tracking system is destroyed when the survey is complete to maintain confidentiality.

5. *Write a cover letter*, including the significance of the survey and any directions not on the survey itself. If possible, ask someone respected by the respondents, such as a senior manager, to sign the letter. An effective cover letter contributes to an adequate response rate:
   - Explain study purpose and importance.
   - Help the respondent feel that the study is significant.
   - Reference respondent's knowledge, professional status, or group affiliation.
   - Mention the value of the information his or her group can provide.
   - Offer to send results.
   - State the length of time the survey will take to complete.
   - Assure the respondents that their answers will be held in confidentiality.
   - Request a return date (mailing time plus one week).
   - Thank the respondent.
   - The letter should look individually word processed (use mail merge to personalize) and signed.
   - Copy letter on appropriate letterhead.
   - Include a self-addressed, stamped envelope when using a paper survey (Borg & Gall, 1989, pp. 427–428).

6. *Pilot test and revise the survey.* Pilot test the survey with a few members of the target population. For the most valid survey, we need to ask the right questions (those which help achieve the objectives of the survey) in the clearest, least ambiguous way. All terms must be clear or clearly defined so that they mean the same thing to all respondents. Use the feedback to revise so the survey is clear and unambiguous. Also note the length of time it takes to complete the survey for the cover letter. The pilot test will also give you an opportunity to conduct preliminary data analysis.
7. *Administer the survey.*
8. *Follow up.* If you coded each mailed survey with a unique number that can be tracked to the mailing list, then follow-up letters or postcards can be sent to those needing reminders. In a secondary analysis of survey research, Kim Sheehan (2001, para. 1) found that since 1986 response rates have significantly decreased, so follow-up reminders are probably going to be necessary to generate an adequate response rate. A few days after the last date indicated in the cover letter, send out a reminder using the same tone and describing the importance of the respondent's contribution and study and indicate that, perhaps because of some oversight or error on the part of the instructional designer, it was overlooked. Another copy of the survey may be included. Postcards have also been successfully used, but in general, the follow-up letter has had about a 7 percent better response rate than a postcard (Borg & Gall, 1989, p. 431).

Studies may need as many as three follow-ups. Keep in mind that some studies have had response rates as high as 92 percent for very short surveys and others as low as 5 percent for longer surveys, so although follow-up is important, other factors may also influence response rates. Sometimes another form for the second and third follow-up may be more effective (Borg & Gall, 1989, p. 431). Often, an effective way of increasing the response rate is to send follow-up certified e-mails or mailings or make telephone calls. If your response rate is low, you are faced with potential bias. One needs to ask how the results would have changed if all the surveys had been returned. With a higher return rate, the results could have been considerably different. This would be the case if the non-respondents represented biased sampling, that is, they are different in some measurable way from those who did respond. For instance, we know from experience that:

- Persons who feel positive about the topic of the survey (such as a training course) are more likely to respond than those who feel that their course is poorer quality.
- Higher achieving persons are more likely to respond than less successful persons.

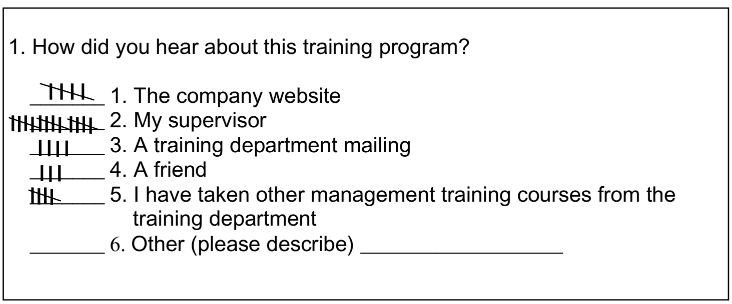
If more than 60 percent of the surveys have not been returned, a portion of the non-responding group should be checked to see whether their responses differ significantly. Twenty is a reasonable number to check. If they differ significantly from the responding group, a sampling bias has occurred. Ideally, you should select a random sample of the group and interview them by telephone or in person. If the differences are significant, this should be noted and discussed in reporting the results of the survey.

One way to improve response rates is through personal administration of a survey. This requires you to locate respondents in groups or individually and give them the survey. This allows the instructional designer to personally explain the survey purpose and importance to respondents and to clarify any ambiguity. This method is often used in educational research, where groups of respondents may be in one place at the same time.

Analyze and Interpret Data

After the surveys have been collected, it is time for organizing and analyzing. Data may have been collected in paper form for manual data entry or on paper Scantron ''bubble sheet'' forms for computer data entry. It might also have been collected with a web survey. Whatever method had been chosen, data will need to be tabulated and prepared for analysis method(s) selected. Careless data entry can invalidate findings, so take care in tallying results. Shown in Figure 4.2 is a manual tally with a frequency count and percent of total responses.

Especially when open-form questions have been used, the next step will be designating appropriate, logical, and mutually exclusive categories for the tabulation. If someone besides the instructional designer is doing the data entry, it is especially important that the open-ended responses be pre-coded and a code sheet provided for the data-entry people. A great deal of time can be saved at this phase if data tabulation has been thought through before data is collected. Numerical codes can be assigned to each of the important variables collected for demographics: For instance, in a survey of management trainees, gender can be code 1


1. How did you hear about this training program?

1. The company website
2. My supervisor
3. A training department mailing
4. A friend
5. I have taken other management training courses from the training department
6. Other (please describe) ________

Figure 4.2 Sample Manual Survey Tally.

for female and 2 for male. Attitude data may be broken down into a scale and ranked or weighted. Departments that house respondents can each be assigned a number (1 for widget department, 2 for gadget department, and so on).

Data may be entered into a spreadsheet such as Excel. Entering the data may be labor-intensive, but calculation is quick and easy. Unfortunately, spreadsheet programs neither tell you whether a chosen statistical formula is appropriate, nor how to interpret the result. Many people using these programs are tempted to use fancy statistics, but it is generally unnecessary. Simple frequency counts, percent of total respondents, or means are usually sufficient. One way of examining the data to look at it is through the use of tables and graphs or charts. You may first choose to describe the group as a whole, but you should also look for patterns in the data such as between group comparisons. For instance, if there are important subgroups, you should examine the responses by each group to see whether they differ significantly. For instance, you might need to examine responses by high and low performers or by employees in different geographic areas. Taking this a step further, you may wish to look for relationships between patterns of responses in your data through a bivariate tabular (crossbreak) analysis. The purpose of this analysis is to see whether:

- There is a relationship between patterns of answers in the data;
- How strong a relationship there might be; and
- The direction of the relationship.

In addition to looking at the patterns in data, you might also consider analyzing the findings with respect to outside criteria. These criteria or benchmarks may consist of things such as:

- Performance or practices of comparable organizations;
- Best practices or research as found in the literature;
- Standards set forth by standards bodies, professional organizations; and
- Laws or rules of government or regulatory bodies.

As you analyze and interpret your findings, be careful to avoid the following problems:

- Confusing statements with facts: perceptions are not the same thing as reality;
- Failing to recognize that a small number of returned surveys may reduce representativeness and limit your ability to generalize;
- Making faulty assumptions;
- Failing to consider alternative explanations or interpretations;
- Confusing frequency of reports with importance;

- Unconscious bias on the part of the researcher or any persons participating in data collection;
- Failure to recognize response sets (a response set reflects a general predisposition rather than carefully thought-out answers to your questions; these may occur with surveys of topics about which people feel strongly). (Borg & Gall, 1989, p. 482–483).

E-mail and Web Surveys

E-mail and web surveys are designed in essentially the same ways as paper-based surveys, following the same steps. Initially, researchers reported that e-mail surveys had a response rate that was equal to or higher than paper surveys (Sproull, 1986, p. 159), but more recent studies suggest that may no longer be the case (Cook, Heath, & Thompson, 2000; Couper, 2000; Sheehan, 2001). Some researchers have also found differences in the quality of responses (Kiesler & Sproull, 1986), while others have found none (Fraze, Hardin, Brashears, Haygood, & Smith, 2003). There seems to be no consensus at this point, and more research needs to be done. Still, there are several advantages to electronic delivery related to speed and cost. Web survey services such as SurveyMonkey and Zoomerang also make web-based survey creation easy.

Advantages

- Reduced costs of postage and duplication
- Less time to distribute
- Reduced time to compile data
- Perceived as environmentally friendly

Disadvantages

- The disadvantages are basically the same as for paper surveys.

Example. One organization transformed a traditional training function into a corporate university to improve performance and productivity through a stronger link to business strategies, increased employee engagement, and greater transfer of training. In order to assess the project, the human resources director undertook a series of surveys using both paper and online methods to collect information from managers and learners in classroom, online, and peer learning groups. The surveys were followed by interviews to validate survey information and to gain more in-depth insight. Table 4.4 shows means for each survey on a selection of questions. For many of the questions, on-site training and peer learning have been very successful. Online training was not found to be as successful when learners took it at their own computers during the workday. One recommendation was to set aside a computer room with blocks of time when learners could be away from their desk to take online courses (Christianson, 2008, p. 84).

Table 4.4 Executive Summary for Survey Questions

	General OPU Learner Survey	*Online Support Survey*	*Peer Learning Survey*	*Manager Survey*
Co-Worker Social Support Questions	2.92	3.20	3.13	–
Manager Support Questions	3.10	2.89	2.81	–
Self-Learning Questions	3.50	3.55	3.45	–
Peer Learning Support Questions	–	–	3.43	–
Management Involvement Questions	–	–	–	3.41
Manager Support Questions			–	3.31

Note. From Implementation of a Corporate University: A Case Study. Unpublished master's paper (p. 69), by K. Christianson, 2008, St. Cloud, MN: St. Cloud State University. Reprinted with permission.

Interviews, Telephone, and Videoconference Interviews

Interviews can be formal or informal, structured, semi-structured, or unstructured. They can be conducted in person, on the telephone, or by videoconference. Skilled interviewers can learn about interviewees' attitudes, perceptions, causes of problems, and possible solutions. Interviewees can express themselves in their own words, and the interviewer can probe and explore issues that may not have been anticipated. Telephone surveys and, when available, videoconference surveys are essentially structured interviews following the same steps as a structured interview.

Advantages

- Many people would rather talk than write, and an interviewer can establish rapport and glean information that a respondent might not want to put in writing. This is particularly true for sensitive topics. However, interviewers must be chosen carefully with regard to gender, race, or other characteristics when the study is about sensitive topics such as sexual harassment or race relations.
- An interviewer can also explain the study, clarify questions, and probe a respondent's answer.

- An interviewer can also ask questions in more than one way, to stimulate the subject's insight or check sincerity or truthfulness.
- The interview is appropriate for people with a variety of issues that present barriers to responding in writing, such as persons with visual impairments.
- Interviewer presence can result in high response rate.
- Telephone surveys can save time and money without being as intrusive as a face-to-face interview. (Piper, 1988, p. 189)

Disadvantages

- The time and labor it involves; interviewing is a labor-intensive process.
- A tendency of respondents to give particular kinds of responses (response effect). These may be inaccurate or not thought out. For instance, exemplary performers often leave out important steps because they have performed the task so often that some of the steps become so internalized that these performers do not recognize every action they are taking.
- Errors from the predispositions of the respondents: if they are hostile, suspicious, indifferent, unmotivated, lack the information the interviewer is seeking, want to please the interviewer or present themselves in a favorable way, the data will lack validity.
- Errors from the predispositions of the interviewers: if they are uncomfortable, ill at ease, cannot establish rapport, opinionated, or have stereotypes, the study can be invalidated.
- Problems with study procedures: the way the study is explained to the respondents, methods for gaining cooperation, length of the interview, place where the interview is held, and presence of other people during the interview may be sources of validity or reliability issues.
- Response rates are lower for telephone surveys.
- Use of visuals is not possible with telephone interviews.

To reduce errors you can:

- Select respondents carefully, including consideration of predispositions.
- Design the study interview guide well.
- Train interviewers.

How To. An interview guide, trained interviewers, and effective recording methods all contribute to successfully conducted interviews.

Interview guide. An interview guide includes notes or a script for explaining the interview purpose and directions. It lists questions to be asked and possible probes in a desired sequence. Typical responses or categories are listed to make it easier to document interview responses. Interviews may be highly structured, semi-structured, or unstructured. *Structured interviews*, where questions and order are the same for all respondents, reduce the impact of context effects by holding question order the same for all respondents. Context effects come into play when the answers given to a survey question depend on the characteristics of previous questions. *Unstructured interviews* may have some core concepts to ask about but no formal structured form or procedure. The interviewer may shift the conversation toward interesting topics that may arise, but each interview tends to be unique for each respondent, so it is more difficult to analyze this data across respondents. Structured interviews are the least troublesome for novices. Like any other data collection method, the interview guide should be pilot tested.

Interviewer training. Interviewers should be comfortable and be able to establish rapport without giving any cues to ''desired'' answers. As with extant data analysis, interviewers should be able to take notes accurately and consistently with other interviewers so that inter-rater reliability is achieved.

If a team will be involved with collecting interview data, the interviewers should be trained to

- Explain the purpose, significance, and directions for the interview.
- Maintain the respondent's interest and motivation to complete the interview. This interest can be engendered when interviewers nonverbally communicate their own commitment to the importance of the project.
- Ask all questions informally but as written, in proper sequence.
- Clarify questions.
- Ask for clarification of the answer, elaboration, or repeat the answer as necessary (Are you saying . . . ?).
- Avoid finishing sentences for the respondent. Pause and wait for respondents or encourage them to continue with a nod.
- Record the answers (including probes) immediately. If this makes the respondent uncomfortable, simply listen carefully and record the responses out of the respondent's view immediately after the interview.
- Avoid interview bias caused by suggesting a particular response by voice or body language cues. It is very important in conducting an interview to avoid indicating what type of answer is expected or leading the respondent in any way.
- Interact with the respondent as an equal.
- If questions are sensitive, remind respondents that their answers will be held in the strictest confidence.

- Conclude the interview by thanking the respondent. Don't be rushed or abrupt.

Recording the interview. Recording the interview has several advantages: one is accuracy. Audio or video media can also be reexamined at a later date to check for alternative interpretations or to test alternate hypotheses. Express permission is needed to record respondents. Computers can be used to record the interview responses as they are given, This is often done with telephone interviews. The data can be placed into a database or qualitative data analysis program for later analysis.

Critical incident technique. Critical incidents can be gathered in various ways, but typically in a structured interview or questionnaire where the interviewer asks respondents to tell stories or anecdotes about experiences they have had that have critical significance to performance problems. These are often called ''war stories'' and are often swapped by SMEs. Usually it is focused on behaviors that are very effective or ineffective at accomplishing the goal of the work performance in question (those that differentiate between successful and unsuccessful accomplishment of the task objective). It is useful to ask for problems at varying levels of difficulty. The interview method has the advantages and disadvantages of other interview methods, and the questionnaire method has the advantages and disadvantages of surveys.

There are a variety of approaches, but this method proceeds more or less like this:

1. Determine the focus and objectives of the data gathering.
2. Gather information on the issue.
3. Create critical incident forms with directions for capturing the information. Directions should request respondents to describe anecdotes about experiences they have had that had critical significance to a performance problem (those that made a difference between successful and unsuccessful accomplishment of the task objective). Ask them to focus on the event, not the person, but include activities, conditions, place, and people. Also ask the respondents to describe the job experience level of the person involved (or themselves, if they were the one involved in the incident), and why the action was particularly effective or ineffective.
4. Decide who to interview and how many.
5. Conduct interviews.
6. Analyze data, using methods discussed under interviews or surveys.
7. Report the information.

Focus groups. A focus group is a group interview method in which a group of people are asked about their attitude toward something. In our field it might be a procedure, process, or performance problem. Questions are asked in a group setting in which participants interact with other group members. It is a good method of identifying issues that should be included on a survey or other data collection instrument.

Advantages

- Low in cost
- Can be done relatively quickly with a group of people
- Can generate many ideas through interaction
- May allow synthesis of many ideas
- May tell why something happens

Disadvantages

- Data may be difficult to synthesize or quantify
- Does not describe how much, how often, or how critical
- May be dominated by one or two strong-willed participants
- The interaction among the group members may create the impression that a problem is more widespread than it is, or less so
- Validity of results depends on the representativeness of the sample

Consensus groups. Consensus groups are used to come to a group decision of judgment on a question, problem, or topic. Two types of consensus group techniques may contribute to the instructional design process: the nominal group technique and the Delphi technique.

Nominal group technique. This technique is a group judgment or decision-making process done in a face-to-face setting or in a virtual setting such as a teleconference or videoconference. There are no hard-and-fast rules, but it is conducted more or less as follows: (1) participants are given a question or problem, (2) they generate responses silently, in writing, (3) responses are collected and posted anonymously, (4) responses are clarified in round-robin format, (5) further iterations may follow, (6) and a final set of responses is established by voting or ranking (Price, 1985, p. 13). A similar group process, when many experts are involved, is the Delphi technique.

Delphi technique. This technique is an iterative group judgment technique in which consensus is sought through the following steps: (1) a survey or questionnaire is sent to expert participants who are unaware of others' participation, (2) the responses are collected and summarized, (3) the summary of responses is sent to the participants, who are asked whether they wish to revise their responses, (4) after multiple rounds (at least three), a consensus should emerge and a final group response can be reported. Keith Zoski and Stephen Jurs recommend establishing a cutoff point in cases when consensus takes so many rounds that the study suffers participant attrition (1990, p. 215).

DIVERSITY ISSUES IN DATA COLLECTION

Most organizations have experienced the effects of diversity through serving diverse customers domestically or globally or working with diverse employees. Culture, language, and other diversity factors affect assessment, instructional design, and implementation in several ways, including learner characteristics (language, culture, physical characteristics) and the selection of assessment and instructional methods. If you are collecting data, you will need to consider your target population carefully when you choose your data collection methods and design of instruments. For instance, a training team planned to use a variety of methods including extant data analysis, surveys, and critical incident interviews to do a needs analysis that included a variety of their international sales force. All went well until the pilot test with their Asian sales team. They found that the learners were concerned that someone would lose face, so they would not write any criticisms on a survey. The team also had problems with the critical incident method; they could not find the words to explain clearly the type of anecdote they needed. Finally, they decided to use structured interviews because the learners agreed they would express their concerns verbally. The instructional designers waited until they returned to their offices to make their notes.

The training team learned two things from the process: (1) it is important to be aware of how their own values, beliefs, and attitudes guide the design and development process and decisions; and (2) a better understanding of the values, beliefs, and attitudes of the cultural groups that are included helps them make better decisions about the process and methods chosen. It was a good decision to incorporate those learners in the assessments because they provided critical feedback on the assessment design and methods, as well as on the courses that were ultimately redesigned. It was valuable to be flexible in the choices of data collection sources and methods because the team ended up building better buy-in.

Over the years since that project, the training team learned to anticipate the problems encountered in that first analysis and have developed some strategies to address them, including:

- Use multiple methods for data collection;
- Build in more time for data collection when language, culture, or disabilities present communication challenges;

- Use as little jargon as possible, and translate written surveys where necessary;
- Consider language, cultural values, attitudes, and expectations in the needs analysis;
- Gain the trust of the respondents; and
- Observe performance over an extended period. (Hites, 1990, p. 232; 1996, p. 70)

SELECTING THE RIGHT METHODS (WHICH AND HOW MANY)

A good instructional design or performance improvement project requires good analysis, but the analysis can be no better than the data collected. There are many data source and data collection methods that we can select, each with strengths and weaknesses. So how do we choose? First, we don't settle on a single source or method. A more valid set of findings and recommendations is often made when multiple data sources and methods are used. A rule of thumb is to triangulate sources and methods—always choose at least three of the best-matched sources and methods. For instance, beginning with an analysis of extant documents, an instructional designer should be able to determine job tasks. Job incumbents, supervisors, and subject-matter experts then can be consulted through surveys, interviews, the critical incident technique, or group processes such as consensus groups about the nature of each task, including the frequency, difficulty, and criticality of the tasks. Customer surveys can be used to measure their perceptions of product quality and service performance. More on task analysis will be discussed in Chapters Seven and Eight. We need to keep in mind the four criteria of utility, feasibility, propriety, and accuracy when selecting data collection sources and methods, making sure they are matched to the problem. We want to collect only enough information to answer the analysis questions with information that is both valid and reliable and to avoid the problem of paralysis by analysis that drives up the costs of the training or performance intervention. However, every organizational context places different pressures on the selection criteria. There always seems to be a balancing act taking place among time, budget, and accuracy, with the instructional designer juggling them on the high wire. For instance, budget pressures are sometimes more severe than time constraints or the other way around. And both of these can impact the accuracy of the data. In other words, there is not a cookbook formula for selecting analysis sources and methods that apply to every organization, so you will need to prioritize your selection criteria in accordance with your own organizational context. Table 4.5 summarizes selection factors.

Table 4.5 Selection Factors

Data Collection Method	*Time*	*Cost*	*Response Rate*	*Advantages*	*Disadvantages*
Extant Data					
Analysis	Medium	Low	N/A	Not obtrusive	May reflect author bias
Work samples	Medium	Low	N/A	Outcome-focused	May have restricted access
				Speed	Reflects past situation
Surveys					
Paper	Medium	Medium	Low	Can reach widespread audience quickly	Answers cannot be probed
Electronic	Low	Low	Low	Anonymity allows frank answers	Difficult to get at problem causes
					Response burden
Observation					
	High	Medium	High	Minimally obtrusive	Not as useful for knowledge workers
				Flexible	Observer may influence behavior
					Provides picture of an actual workday
Interviews					
In-person	High	High	High	Many people prefer to talk rather than write	Labor-intensive
Telephone	Medium	Medium	Medium	Interviewer can probe, clarify	Response sets

Critical incident	Medium	High	High	Synthesis of ideas	Interviewer may influence respondent
Focus groups	Low	High	Medium		Data may be difficult to synthesize
Consensus Groups					
Nominal group technique	Low	Medium	High	Synthesis of ideas	Labor-intensive
Delphi study	High	Medium	Medium	Synthesis of expert ideas	Data may be difficult to synthesize

CONCLUSION

After collecting and analyzing data on the analysis questions from each data source and method, you will note important findings from each and then begin the process of synthesizing findings from all sources and methods to come to an overall understanding of the performance problem. The report of findings will be more persuasive if evidence is clearly linked to conclusions.

Conclusions and recommendations result from the process of interpreting the findings. For a performance analysis, you will recommend whether or not to do a needs analysis. More on performance and needs analysis will be covered in Chapters Five and Six. For a needs analysis, this step requires making sense of the pattern findings and then determining what solutions match with the causes. Finally, the analysis findings should be reported to stakeholders in a timely and impartial fashion. Using graphics to explain patterns in the data can be an aid to understanding on the part of your readers, but must be done clearly and ethically (Wainer, 1997, p. 2).

References

American Association for Public Opinion Research. (2007). *Response rates—an overview*. Retrieved March 26, 2008, from www.aapor.org/responseratesanoverview.

Borg, W. R., & Gall, M. D. (1989). *Educational research*. New York: Longman.

Brinkerhoff, R. O., Brethower, D. M., Nowakowski, J., & Hluchyj, T. (Eds.). (1983). *Program evaluation: A practitioner's guide for trainers and educators*. Boston: Kluwer-Nijhoff.

Burton, J. K., & Merrill, P. F. (1977). Needs assessment: Goals, needs and priorities. In L. J. Briggs (Ed.), *Instructional design principles and applications* (pp. 21–48). Englewood Cliffs, NJ: Educational Technology Publications.

Christianson, K. (2008). Implementation of a corporate university: A case study. Unpublished master's paper. St. Cloud State University, St. Cloud, MN.

Cook, C., Heath, F., & Thompson, R. L. (2000). A meta-analysis of response rates in web- or Internet-based surveys. *Educational and Psychological Measurement*, *60*, 821–826.

Couper, M. P. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly*, *6*, 230–253.

deVries, W., Keller, W., & Willeboordse, A. (1996). Reducing the response burden: Some developments in the Netherlands. *International Statistics Review*, *64*(2), 199–213.

Finkler, S. A., Knickman, J. R., Hendrickson, G., Lipkin, M., Jr., & Thompson, W. G. (1993, December). A comparison of work-sampling and time-and-motion techniques for studies in health services research. *Health Services Research*, *28*(5), 577–597.

Finnegan, G. T., & McCampbell, J. F. (1982). The role of value analysis. *NSPI Journal*, *21* (5), 16–18.

Fraze, S. D., Hardin, K. K., Brashears, M. T., Haygood, J. L., & Smith, J. H. (2003). The effects of delivery mode upon survey response rate and perceived attitudes of Texas agri-business teachers. *Journal of Agricultural Education*, *44*(2).

Gilmore, E. R. (2006). Using content analysis in human performance technology. In J. A. Pershing, H. D. Stolovitch, & E. J. Keeps (Eds.), *Handbook of human performance technology: Principles, practices, and potential* (3rd ed.). San Francisco: Pfeiffer.

Hites, J. M. (1990). Design and delivery of training for international students: A case study. Doctoral Dissertation, Northern Illinois University, 1990. *Dissertations Abstracts International*, *11A*(51). 3599. Retrieved September 11, 2008, from First Search/Dissertation Abstracts International database.

Hites, J. M. (1996). The design and delivery of training for international trainees: A case study. *Performance Improvement Quarterly*, *9*(2), 57–74.

Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Reading, MA: Addison-Wesley.

Isaac, S., & Michael, W. B. (1981). *Handbook in research and evaluation*. San Diego: EdITS Publishers.

Jonassen, D. H., Tessmer, M., & Hannum, W. H. (1999). *Task analysis methods for instructional design*. Mahwah, NJ: Lawrence Erlbaum Associates.

Kaufman, R. A. (1972). *Educational system planning*. Englewood Cliffs, NJ: Prentice Hall.

Kiesler, S., & Sproull, L. S. (1986). Response effects in the electronic survey. *The Public Opinion Quarterly*, *50*(3), 402–413.

Noe, R. A. (2004). *Employee training & development*. NY: McGraw Hill.

Pershing, J. A. (2002). Using document analysis in analyzing and evaluating performance. *Performance Improvement*, *41*(1), 36–42.

Piper, T. D. (1988). Telephone surveying: An effective and efficient assessment and evaluation technique. *NASPA Journal*, *25*(3), 185–190.

Price, W. T., Jr. (1985). The nominal group technique: A needs assessment methodology for vocational education. *Journal of Vocational and Technical Education*, *2*(1), 11–25.

Rossett, A. (1987). *Training needs assessment*. Englewood Cliffs, NJ: Educational Technology Publications.

Rossett, A. (1999). *First things fast: A handbook for performance analysis*. San Francisco: Pfeiffer.

Scriven, M. (1974). Prose and cons about goal free evaluation. In W. J. Popham (Ed.), *Evaluation in education* (pp. 34–43). Berkeley, CA: McCutchan.

Sheehan, K. B. (2001). E-mail survey response rates: A review. *Journal of Computer-Mediated Communication*, *6*(2). Retrieved March 2, 2008, from http://jcmc.indiana.edu/vol6/issue2/sheehan.html.

Sproull, L. S. (1986). Using electronic mail for data collection in organizational research. *The Academy of Management Journal*, *29*(1), 159–169.

Stake, R. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.

Stemler, S. (2001). An overview of content analysis. *Practical Assessment, Research & Evaluation*, *7*(17). Retrieved March 8, 2008, from http://PAREonline.net/getvn.asp?v = 7&n = 1.

Thomas, R. M. (2003). *Blending qualitative and quantitative research methods in theses and dissertations*. Thousand Oaks, CA: Corwin Press.

Van Teijlingen, E. R., & Hundley, V. (2001, Winter). The importance of pilot studies. *Social Research Update*, *35*. Retrieved March 6, 2008, from http://sru.soc.surrey.ac.uk/SRU35.html.

Wainer, H. (1997). *Visual revelation: Graphical tales of fate and deception from Napoleon Bonaparte to Ross Perot.* New York: Springer.

Watson, S. C. (1998). A primer in survey research. *The Journal of Continuing Higher Education*, *1*(46), 31–40.

Yin, R. (1994). *Case study research: Design and methods* (2nd ed.). Thousand Oaks, CA: Sage.

Zemke, R., & Kramlinger, T. (1982). *Figuring things out: A trainer's guide to needs and task analysis*. Reading MA: Addison-Wesley.

Zemke, R., & Rossett, A. (1985). *Be a better needs analyst*. Alexandria, VA: American Society for Training and Development.

Zemke, R., & Kramlinger, T. (1982). *Figuring things out: A trainer's guide to needs and task analysis*. Reading, MA: Addison-Wesley.

Zoski, K., & Jurs, S. (1990). Priority determination in surveys: An application of the screen test. *Evaluation Review*, *14*(2), 214–219.

CHAPTER FIVE

From Performance Analysis to Training Needs Assessment

Kerry J. Burner

What are performance analysis and training needs assessment, and how are they conducted? This chapter provides an introduction to both processes in two ways: a discussion of the elements and procedures involved and a complementary narrative that explores ways in which the two processes, performance analysis and training needs assessment, can play out in an organization. Often conflated because of their fundamental reliance on analysis, performance analysis and training needs assessment are two separate processes; in fact, one emerges from the other. But they are similar in the steps they take, the systematic way in which they are performed, and the data collection processes applied. In a sense, a training needs assessment is a tighter iteration of a performance analysis. Performance analysis reveals areas of need in an organization, and assuming training is part of the recommended set of solutions—and it is not always—a training needs assessment is conducted to articulate, to verify, and to determine very specifically what training is needed and how best to bring that training to the performers who need it. A training needs assessment should be triggered by a performance analysis. Training that is able to be traced to a performance analysis is more easily justified both rationally and financially.

DEFINITIONS OF PERFORMANCE ANALYSIS AND TRAINING NEEDS ASSESSMENT

One of the most well-known authorities on these two processes, Allison Rossett (1999), offers the following definitions in her highly regarded *First Things Fast: A Handbook for Performance Analysis*:

> *Performance Analysis:* Process by which you partner with clients to identify and respond to problems and opportunities, and to study individuals and the organization and to determine an appropriate cross-functional solutions system. A systematic and systemic approach to engaging with the client; this is the process by which you determine when and how to use education and information resources. (p. 227)

> *Training Needs Assessment:* Systematic study that incorporates data and opinions form varied sources in order to create, install, and evaluate educational and informational products and services. The effort commences as a result of a handoff from the performance analysis and should concentrate on those needs that are related to skill, knowledge, and motivation. Also known as a needs assessment. (p. 230)

But Rossett's (1999) definitions are not the only ones in the parlance of performance analysts. Roger Kaufman (1985, 1995; Kaufman & Valentine, 1989) distinguishes between a needs assessment and a needs analysis, keeping cost at the forefront of needs assessment and expanding costs to include social as well as financial considerations.

> *Needs Assessment:* The identification of needs [gaps between current results and desired results], and the placing of them in priority order in terms of what it costs to meet the needs versus the cost for ignoring them (where ''cost'' is both social and financial). (Kaufman, 1985, p. 88)

> *Needs Analysis:* The breaking down of an identified need to determine its basis and causes and the relationships among identified needs. (Kaufman, 1985, p. 88)

This chapter relies on Rossett's (1999) distinction between a performance analysis and a needs assessment, acknowledging that the distinction is at times academic as, in practice, both types of analysis happen at once.

PERFORMANCE ANALYSIS

What does it mean to conduct a performance analysis? If the head of an organization walked in on a performance analysis in progress, what would she see happening? Since performance analysis is data driven, it's difficult to describe ostentatiously observable acts involved, but, depending on the stage of the analysis, she might see presentations, group discussions, interviews, observations, or the performance analyst working with the gathered data to generate a clear understanding of the performance issue and corresponding solution systems.

Performance analysis must be planned ahead of time and be directed by a clearly envisioned goal, initially defined by the stakeholder requesting the performance analysis and then refined with the performance analyst through questioning. Formally or informally, before a performance analysis is undertaken, a needs or opportunity analysis should be conducted. A needs analysis examines the current situation at any level (society, organizational, process, or work group) to identify external and internal influences; the output is a statement describing the current state, the projected future state, and the rationale or business case for action or non-action—to conduct or not conduct a performance analysis (Standard 5, ISPI's Performance Technology Standards). This first step in any analysis revolves around understanding the impetus for the analysis: who requested it and for what reason(s). By asking increasingly detailed questions about perspectives on the problem, its component elements, and the desired outcome, the performance analyst is working to refine the focus of the analysis. Once the goal, or outcome, of the performance analysis has been identified, be it bringing online a new process, system, or technology or identifying the cause of a poorly functioning subunit of the organization, the next step is to conduct a systematic inquiry into the current situation, the reality of the organization, and the features of the desired outcome, the ideal the organization's leaders have envisioned.

A performance analyst may employ any number of well-accepted strategies or models to better understand the real in order to develop a plan to move toward the ideal (Harless, 1973, 1975; Kaufman, 1985, 2005; Rossett, 1987, 1999; Rummler & Brache, 1990; Zemke & Kramlinger, 1982). Good approaches or models guide the performance analyst through a systematic process of inquiry and analysis and delineate what data needs to be gathered for that analysis. The selection of a particular model depends on the situation, and the one selected is the one the performance analyst determines will best help understand the situation-specific problem and generate solutions. The expert performance analyst has multiple methods at hand.

Likely the performance analyst is responsible for conducting all levels of analysis, from performance analysis to training needs assessment, should one be called for by the performance analysis. Think of a training needs assessment as the response to information uncovered in a performance analysis. Training needs assessments should ensue only after a thorough performance analysis, and only if indicated by the results performance analysis. If a performance analyst is also tapped to conduct any needed training needs assessment, there exists a continuity based in a broad perspective that lends itself to developing systemic solutions.

Simultaneous Performance Analysis and Training Needs Assessment

A critical reality of conducting performance analysis and training needs assessment is that frequently the analyst must go deep and wide at the same time. That is, as the performance analysis unfolds and gaps emerge, the performance analyst—faced with real pressures of time, cost, and opportunity—may have to conduct a training needs analysis *in situ*. This rapid prototyping approach to conducing a training needs analysis during the larger performance analysis takes advantage of the gathered resources, from documents to access to employees, in order to most effectively gather the greatest amount of information about the problem that seems to be solvable with training. The training needs assessment will either verify it as such, thus setting the stage for the design of the training that will follow from the training needs assessment. Or the training needs assessment will eliminate training as a solution, thus informing and redirecting the overarching performance analysis. Figure 5.1 illustrates this relationship.

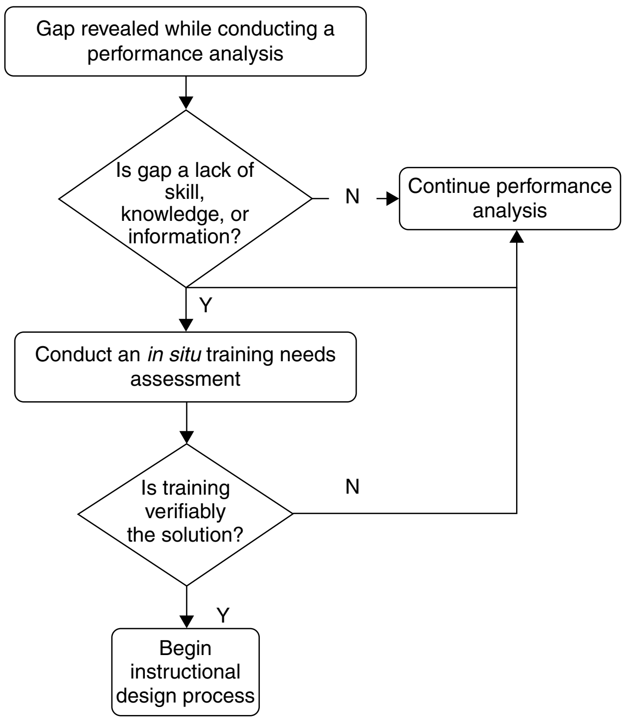


Figure 5.1 *In situ* Training Needs Assessment.

Composite of Foshay, Silber, and Westgaard, 1986; Richey, Fields, Foxon, 2001

PERFORMANCE ANALYSIS METHODS

With any approach to performance analysis, the organization's goal guides the process. A performance analyst gathers data about each contributing unit of an organization in terms of the goal. To gather data under the guidance of any of the models (specific data-gathering methods, instruments, and considerations should also be tailored to the specific analysis at hand; these are discussed in more detail later in this volume), the performance analyst would ask questions of people at all levels of the unit and organization, probe organizational artifacts like processes and reports, and research or access research about industry standards and customer expectations. Then the performance analyst examines and analyzes the set of data to develop a solution system that addresses all of the necessary groups in the organization. Such a solution system would necessarily be broadly configured; since organizational groups are not a set of silos, independent of each other, the ideal can only be attained if the organization's groups are functioning in a series of coordinated and interdependent interactions supported by a solution that considers the whole organization and, specifically, the individuals who constitute it.

Every performance analysis is the process of conducting a situation-specific analysis to determine solutions developed from a clear understanding of the situation and constituting elements. Being systematic is key to this process. Being systematic simply means making a plan before the undertaking and being meticulous about following that plan—including revising the plan as the situation demands: identifying the end point and developing a way to reach it, but being responsive to situational aspects that force revision. For example, the plan may include specific populations, but once the analysis has begun, it may become clear that another group in the organization should be tapped for their perspectives. The plan would then be revised to include that group. Revisions, however, should not be undertaken lightly, or the analyst runs the risk of broadening the project beyond its original scope.

Creating a plan should be the first step of any analysis project, regardless of its focus or the models used to conduct it. Not only does a plan give the project shape, but it provides a framework for a systematic inquiry. The first edition of *Instructional Design Competencies: The Standards* (1986) outlines the seven parts of any needs assessment plan (a through g, below) and serves as a guide to performance analysis as well (Foshay, Silber, & Westgaard, 1986). In the third edition of *Instructional Design Competencies: The Standards* (2001), the six elements of a needs assessment, which are also appropriate to a performance analysis, are described in Competency 6 (1 to 6, below) (Richey, Fields, & Foxon, 2001). Integrated, they create a blueprint for an assessment plan and will guide the discussion of training needs assessment later in this chapter.

1. Describe the problem and its dimensions, identifying the discrepancies between current and desired performance.
   a. Objectives
   b. Target audience
2. Clarify the varying perceptions of need and their implications.
3. Select and use appropriate needs assessment tools and techniques.
   c. Sampling procedures
   d. Data collection methods
   e. Specifications for instruments and protocols
   f. Methods of data analysis
   g. Description of how decisions will be made based on the data
4. Determine the possible causes of the problems and potential solutions.
5. Recommend and advocate non-instructional solutions where appropriate.
6. Complete a cost benefit analysis for recommended solutions. (Foshay, Silber, & Westgaard, 1986; Richey, Fields, & Foxon, 2001)

Many books and articles have been published detailing the processes of performance analysis and training needs assessment. The first handbook in this series as well as the third edition of the *Handbook of Human Performance Technology* (Pershing, 2006) offer extensive discussions of models and methods; particularly, the reference listings from Chapters Eight through Ten are a useful resource to begin researching performance analysis.

Since the general acceptance of performance analysis as a regular in the array of organizational tools, not much research has been done evaluating models or techniques; practice seems to drive what's accepted and not accepted. After all, the principle behind performance analysis is not complex: it's the systematic study of the alignment of the real and the ideal in terms of organizational success. Once that's accepted, the ways in which this study is undertaken—the ways in which the systematic analysis is carried out—can and should fit the circumstance, making the creation and validation of monolithic models inappropriate. What follows is a brief overview of approaches to conducting performance analysis.

Front-End Analysis

The importance of conducting analysis before implementing a solution was brought into focus by Joe Harless (1973) when he championed the concept and process of front-end analysis. Terminology around front-end analysis has evolved, and performance analysis and needs assessment have emerged and been identified as two distinct processes, but both are grounded in the approach Harless championed, marked by a continual drilling down, a search for clarity before action. He recommended a systematic approach to determining not just the problems and solutions, but calculating the associated costs; his thirteen ''smart questions'' are a seminal approach to conducting an analysis, particularly a performance analysis or a training needs assessment:

1. Do we have a problem?
2. Do we have a performance problem?
3. How will we know when the problem is solved?
4. What is the performance problem?
5. Should we allocate resources to solve it?
6. What are the possible causes of the problem?
7. What evidence bears on each possibility?
8. What is the probable cause?
9. What general solution type is indicated?
10. What are the alternative subclasses of solution?
11. What are the costs, effects, and development times of each solution?
12. What are the constraints?
13. What are the overall goals? (Harless, 1973)

In a 1997 interview, Harless declared, ''One thing is for sure: Whoever does the front-end analysis should do the back end analysis'' (Langdon & Whiteside, p. 36). This connection between problem analysis and solution, or interventions, generation is a critical success factor in a performance analysis. In discussing the ways performance analysts determine interventions, Harless presents a taxonomy of interventions, laying out five classes of interventions: (1) employee selection and assignment, (2) work process re-engineering, (3) work environment re-engineering, (4) motivation, incentive, and attitudinal, and (5) skills, knowledge, and information, further articulating the last category into (5a) training and (5b) job aids (Langdon & Whiteside, 1997). But Harless cautions: the result of a ''diagnostic front-end analysis'' is never a single intervention (Langdon & Whiteside, 1997). This holistic view of an organizational problem and situations is evident in the next two approaches, the Behavioral Engineering Model (BEM) and the Six Boxes Model (SBM).

Behavioral Engineering Model and Six Boxes Model

The Behavioral Engineering Model. With his behavioral engineering model, Thomas Gilbert (1996) offers the performance analyst a deep understanding of human performance from a behavioral perspective. Based in the study of human behavior and focused on performance and the constituting influences of events (or non-events), the model limns a stimulus–response–reinforcer relationship among three components of behavior, information (stimulus), instrumentation (response), and motivation (reinforcer) on two levels, the environment (the organization) and the individual. Key to Gilbert's model is that the resulting six elements of the model are not intended to be understood in isolation: they are six ways of looking at the same event (Gilbert, 1996). The model is outlined below:

Environmental Supports

Information = Data: performance feedback, performance expectations, and guides to performance;

Instrumentation = Instruments: science-based tools and materials needed for work;

Motivation = Incentives: financial, non-monetary, and career development opportunities;

Person's Repertory of Behavior

Information = Knowledge: scientifically designed training and placement of workers;

Instrumentation = Capacity: considered and tailored approaches to maximizing workers' abilities;

Motivation = Motives: assessing and understanding extant motives, selecting motivationally aligned workers. (Gilbert, 1996)

The principle behind the model is that a direct relationship can be made between performance and certain features of a work environment and/or individual worker. The model is a diagnostic tool that helps pinpoint the most effective way to improve performance, or ''an outline of a performance troubleshooting sequence'' (Gilbert, 1996, p. 91).

A performance analyst using the behavioral engineering model would begin by working through the stimulus–response–reinforce relationship on the environmental level (data–instruments–incentives), then move to the individual level (knowledge–capacity–motives) in order to locate the cause(s) of a performance problem. Gilbert notes that, while the model is effective in locating causes, because it is six ways of looking at a particular behavior, it does not necessarily indicate solutions. Gilbert points to achieving the greatest leverage by implementing the solution most likely to increase worth performance, defined as the ratio of the value of the outcome of performance to the cost of behavior (Gilbert, 1996).

The Six Boxes Model. The six boxes model offered by Binder (1998) is a derivative of Gilbert's (1996) behavioral engineering model. It does not emphasize the Skinnerian terminology and relationship between the components, making it more accessible for those unfamiliar with these concepts and language (Binder, 1998). Like the BEM, the SBM is visually depicted as six boxes, arranged in two rows of three with the first row examining the environment and the second the individual. The three columns in Gilbert's model are labeled information, instrumentation, and motivation, and while the six boxes model does not employ these labels, they are evident in the relationships between the boxes in the columns. The parts of the model are summarized below; for a complete discussion of the six boxes model, including how specifics of the derivation from the BEM, see Binder (1998).

Environment

Box 1—Expectations and Feedback: clearly stated performance objectives and processes plus performance feedback;

Box 2—Tools and Resources: everything the performer needs to meet the stated performance objectives, from the tangible to intangible, for example, workstation to mentoring; and

Box 3—Consequences and Incentives: results of the performer's efforts, typically compensation and incentives, but extends to intended and unintended, positive and negative outcomes.

Individual

Box 4—Skills and Knowledge: what the performer needs to know and do to be successful;

Box 5—Selection and Assignment: performer characteristics that are a prerequisite to organizational membership; and

Box 6—Motives and Preferences: individual performer's drivers and druthers. (Binder, 1998)

Performance analysts using either model work through each of the six boxes, starting with the environment, then moving to the individual. At the core of each unit or department of any organization is a group of people who work in tandem with each other and with the people in other units or departments. The behavioral focus of the BEM and the SBM gets to the core of organizational performance by way of its people.

Performance Analysis Flowchart

The Performance Analysis Flowchart (Mager and Pipe's model) sequences questions to discover the nature of the performance problem and it funnels the analysis to an appropriate training or non-training solution (Mager & Pipe, 1984). As with all of the models, the performance analysis model begins by eliciting a description of the problem. Then it asks the analyst to assess the importance of the problem in terms of its impact on organizational success and the cost to solve or ignore the problem. Once importance is established, the analyst moves through a series of yes or no questions, ultimately looking at the performer's skill set: if a deficiency exists and the performer has not performed the skill in the past, the model recommends formal training; however, if the performer has satisfactorily performed the skill in the past, the model recommends practice for the infrequent performer and feedback for the one whose performance has declined (Mager & Pipe, 1984).

Mager & Pipe's model also offers the performance analyst pathways for problems resulting from something other than a skills deficiency by asking questions about four dimensions of the performance problem: Is the performance punishing? Is it rewarding? Does the performer care about performing well? Are there obstacles to performance? The solutions for these four dimensions would involve removing punishers and obstacles and implementing consequences and feedback loops. Mager and Pipe's model is an effective tool for a quick, targeted analysis but runs the risk of oversimplification (Rothwell, Hohne, & King, 2007).

Mega Planning and the Organizational Elements Model

Roger Kaufman takes a societal perspective with his theories around mega planning (Kaufman, 1985, 2000, 2005). Mega planning focuses on the organization in the larger context of society: ''From this shared societal value-added frame, everything one uses, does, produces, and delivers is linked to achieve shared and agreed-upon positive societal results. This societal frame of reference, or paradigm, I call the Mega level of planning'' (Kaufman, 2005). Key to this perspective is Kaufman's Organizational Elements Model (OEM), which has five dimensions: the Mega (societal), Macro (organizational), or Micro level (unit/group), processes, and inputs (Kaufman, 2000). The first three are seen as the ends—the outcomes, outputs, and products of an organization—while the latter two are the means—the mechanisms and resources that an organization depends on to function.

Analysis that is conduced from a mega planning perspective takes into consideration all five dimensions, the means and the ends, but it is focused on a need—a gap in results—in the first three dimensions (Kaufman, 2000). Once a needs assessment identifies the gaps in results, a needs analysis is conducted to identify the causes and corollary solutions; final solution implementation depends on the costs compared to the payoffs (Watkins & Kaufman, 1996). Mega planning involves using the OEM, but the model can be used independently as a way to document the real and the ideal.

Four Opportunities for Performance Analysis

Rossett (1999) has taken a different approach by mapping four series of stages to undertake a targeted performance analysis. Each series is tailored to one of the four opportunities for performance analysis she identifies: (1) the rollout of a new process, system, or technology, (2) improving the performance of an organization or a sub-unit of that organization—a problem, (3) developing specific personnel, and (4) developing organizational strategy. The four processes, data sources, and sequences are summarized in Table 5.1.

Keeping with a systematic approach is not difficult when following the stages Rossett (1999) has outlined for each of the four opportunities for a performance analysis. The amount of data and data sources can be overwhelming, and the stages clearly state from whom to start gathering what kinds of information for each type of performance opportunity. Each series of stages starts with an interaction with the sponsor of the performance analysis. It is critical that the performance analyst understand the rationale—both organizationally and financially—for the analysis and the impact the sponsor expects it to have. (For further discussion of the cost-benefit tradeoff analysis, see Anderson, Chapter Four in this volume.) Frequently, this conversation works to refine the purpose of the performance analysis.

When a performance analyst is asked to conduct an analysis that organizational leaders hope will help improve the performance of an organization or a subunit of that organization, that is, fix a problem, the problem must first be refined. As with all other opportunities for performance analysis, the stages of the analysis begin with the sponsor (Rossett, 1999). Of particular interest is the nature of the problem, its history, and the efforts made thus far to solve it. After this conversation with the sponsor of the analysis, organizational artifacts should be closely examined for additional information that will help to define the problem and point to root causes. Organizational artifacts include documents, reports, products, and policies related to the problem—anything tangible that represents the organization in part or whole. Next, experts—internal or external—should be consulted to help determine the optimal outcome: If the problem didn't exist, what would be happening? What would the performers be doing? Finding out what the performers say they are actually doing and getting their perspectives on the problem along with their conception of solutions is the next stage, followed by an examination of industry literature to examine barriers and best practices for problems of the sort under analysis. Last, the performer's supervisors should be asked about the problem—whether they see it the same way as the analysis sponsors, what root causes and solutions they may have, and whether they see the same priority as the sponsors do.

Rossett's Five Types of Questions. Rossett (1987, 1999) offers the analyst an approach to inquiry that relies on five types of questions to find out about different dimensions of the performance problem. The questions in Rossett's typology range from a broad understanding of the problem to details of the situation, the performance, and the attitudes; the questions also probe for causes. The first four

Table 5.1 Summary of Rossett's Four Opportunities—Sources and Sequences

Sources of Data	*Sequence of Inquiry Steps for Four Opportunities for Analysis*			
	Rollout of a New Process, System, or Technology	*Improving the Performance of an Organization or a Sub-Unit of that Organization*	*Developing Specific Personnel*	*Developing Organizational Strategy*
Discussion with the Sponsor About the Impetus, Desired Outcome, and Attitudes	1	1	1	1
Organizational Artifacts		2		
Discussion with Organizational Leaders and Managers				2
Discussion with Internal Experts	2	3	3	3a
Discussion with External Experts			4	3b
Discussion with Decision Makers		3		
Discussion with Vendors	4			
Review of Industry Literature	4	2	4	
Discussion with Performers	5a	5a	5a (model performers), 5b (average performers)	5
Discussion with Supervisors	5b	5b	6	6

types of questions investigate a perspective the analyst will need to determine the need for training, while the fifth question returns information about the need for a performance analysis by asking about the causes of the problem (Rossett, 1987, 1999). Optimally, a performance analyst would start with Types 1 and 5, moving to 2 and 4 next, concluding with type 3 (Rossett, 1982).

Table 5.2 shows Rossett's five types of questions and provides a guide for generating questions in each type (Rossett, 1987). Table 5.3 presents a summary of what information to uncover based on what's known and the corresponding question type to use (Rossett, 1987).

Table 5.2 Generating Questions Using Rossett's Typology of Five Questions

Type 1	
Topic	The nature of the problem
Purpose	Seek a general picture of the problem
Question generation prompts	What's happening/not happening? How should it be happening? When is it a problem? When is it not a problem? Who's affected? Positively? Negatively? Where does the problem exist? Why is it allowed to exist?
Type 2	
Topic	The priorities within the problem
Purpose	Seek details of the situation
Question generation prompts	Who's performing the problem? Who's responsible for it? For fixing it? What's the most important aspect of the problem? Least? Rank a list of the problems/solutions. Where should performers get the required skills/knowledge?
	Why is it still a problem? When did the problem become a problem? How do the performers identify/work around/avoid the problem?
Type 3	
Topic	The subject matter or skills
Purpose	Generate proof of what performer knows and how performer enacts job function
Question generation prompts	Ask performer to perform as if there were not a problem. Ask performers to perform as if they had the needed skills/knowledge to perform the task problem-free.

Type 4	
Topic	The attitude toward the problem
Purpose	Seek feelings about the topic/skills/body of knowledge/ training/perception of priority of training/confidence related to it
Question generation prompts	Ask questions that are tailored to the performers: their perspectives, their job functions, their level of responsibility, etc. Use a variety of feeling words in questions.
Type 5	
Topic	The problem's cause
Purpose	Seek the causes of the problem
Question generation prompts	Why is there a problem? What's happening/not happening that causes the problem? Who would fix the problem? Where does the problem come from? Who's causing it? Allowing it to happen? When did the problem become evident? How did the problem develop?

Surveys are one of the most common ways to gather data; the method is inexpensive and less time-consuming than interviews and focus groups. The guiding principles for a good survey question serve too for a good interview question; a performance analyst should ensure that every question:

- Evokes the truth. Questions must be non-threatening.
- Asks for an answer on only one dimension.
- Can accommodate all possible answers.
- Has mutually exclusive options.
- Produces variability of responses.
- Follows comfortably from the previous question.
- Does not presuppose a certain state of affairs.
- Does not imply a desired answer.
- Does not use emotionally loaded or vaguely defined words.
- Does not use unfamiliar words or abbreviations.
- Is not dependent on responses to previous questions.
- Does not ask the respondent to order or rank a series of more than five items. (Walonick, 2004)

Table 5.3 Knowns and Questions to Use to Discover Unknowns

If it's known . . .	*Then the analyst must . . .*	*Using question*
That there is or will be a performance gap	Determine the desired performance or knowledge (optimals)	Type 1/Type 2
What the desired performance or knowledge (optimals) should be	Determine what the actual performance and knowledge (actuals) are.	Type 1/Type 2/ Type 3
What the optimals and actuals are	Determine the individual's attitude or feelings about the job or situation.	Type 4
What the optimals, actuals, and attitudes are	Identify the cause(s) of the problem.	Type 1/Type 2/ Type 3/Type 4/ Type 5
What is causing the problem	Develop solutions to the training-related problems and recommend sources for solving	

See other chapters in this part of the book for further discussion of well-crafted data gathering instruments and protocols.

Performance Analysis in Action: A Case Study

Maria re-read her notes from today's meeting with Ty, the district vice president of the leading network of national call centers. He had provided a lot of information that she now needed to organize and consider. The performance analysis was initiated because a rural call center was financially underperforming and had an incredibly high rate of customer complaints about customer service. Ty wanted Denise to fly to the call center and figure out what was going wrong with the intensive and costly customer service training for the call center employees. After initial training, a system of tutorials was used to respond to customer input. If cumulative feedback indicated the need for a refresher, workers would have to take a tutorial in an area of customer service before being able to access a work screen. Ty said they'd put one of their best customer service people in charge of the site, Denise, a long-time employee.

Right away, Maria noticed that Ty (and apparently Denise) were operating under the assumption that the root cause was a lack of customer service skills and knowledge on the part of the workers, which is why Ty's focus was on verifying the effectiveness of the computer-based customer service training. Maria realized that, ultimately, if she was going to get to the real causes and appropriate solutions, she'd have to get Ty and Denise to uncouple their assumptions from the facts by broadening their understanding of the situation. She'd make sure that, in tomorrow's conference call with Denise, Ty authorized Maria access to a wide range of information once on site. Maria would have to be able to look at data that might seem to Denise out of the scope of the performance analysis. Denise knew that Maria was coming; she welcomed the input on getting the tutorials working correctly, as she freely admitted to being a low-tech person. But Maria wanted to make sure that the tutorials were indeed the problem. This meant possibly working outside of Denise's expectations—asking for data that Denise might not see as relevant, given the trajectory of her thoughts on the problems.

Once on site, Maria spent time with Denise, getting to understand her perspective on both the financial and customer service problems. Both problems, according to Denise, were likely caused by the same issue: poorly trained workers. Clearly, Denise thought little of the locals working at this new, rural call center, and she viewed the computer-based training as inadequate—seeming almost eager to help prove it a failure. Maria asked a few questions about the timing of the financial losses and the drop-off in positive customer service feedback rates. Denise had already crunched the numbers and the two definitely seemed related, suggesting that fixing one might fix the other. Maria suggested that analysis frequently uncovered unanticipated relationships and shared that she'd be asking for a wide range of data and free access to all levels of workers. Before the meeting ended, Maria requested some of data that she knew she was going to need: data on employee performance, attendance, and past training.

After two days on site, Maria had gathered enough information to begin analyzing. One thing she'd learned from being around the call center was that workers would be tricky to write up. The workers had made enough side comments for Maria to get the picture: they knew Denise thought little of them and had little interest in helping her. They did their jobs and cared quite a lot about doing them well. They'd responded well to the computer-based training at first, but were quickly souring on it, as it was mandating training they felt they didn't need. They were happy to help Maria, whom they saw as an ally from Corporate with the same goal of keeping the center open—economically, their area needed the call center. Through this positive rapport, Maria gained the trust of the workers. She knew that buy-in at all levels was essential to an effective performance analysis. What she didn't know was how to address the climate issues at the call center. They were definitely an impediment to both performance and the analysis; for now, she decided the climate issues were outside of the scope of this performance analysis.

The areas with the highest rates of negative customer feedback were related to disconnected calls which, according to the workers, was because the phone system hung up calls. Shift supervisors had reported this to Denise, and when Maria asked, Denise said she thought it was an excuse and that she believed the workers had hung up on customers. Because there was no obvious pattern—no one phone bank or individual phone—and likely because of her unveiled bias, Denise dismissed this critical piece of information.

Maria knew that she wanted to dig deeper to find out whether there was a way to substantiate the experience of the call center workers, who freely admitted the bizarre randomness of the disconnections. If the problem was not centered around a specific bank of equipment in the main call center, what other variables might show clusters of negative customer feedback? In the end, the pattern was not worker or

(*Continued*)

(*Continued*)

equipment related, but time of day. At the same time of day the call center's server exchanged massive amounts of data with the Corporate server, causing hiccups—not enough to impact the data stream but enough to cause a disruption in the Internet-based phone system.

Another clear deficit was revealed during the performance analysis: the call center workers did not receive any formal training on new products and services or on updates to the customer service software. They were expected to read the information in the company newsletter and look over marketing materials, extrapolating the necessary information about new products and services. Frequently, Maria was told, workers had to navigate brand new screens, taking more time than expected for calls. They didn't lack skills, knowledge, and information about customer service; in fact, they were model customer service agents. They lacked support—training or otherwise—to do their jobs. While she hadn't finished her report for Ty, Maria knew that he was efficiency minded and decided to call him the next day, recommending she begin a concurrent training needs assessment to best understand how to meet the real training needs of the call center workers.

RESULTS OF PERFORMANCE ANALYSIS: DERIVING A SOLUTION SYSTEM FROM DATA

The importance of a performance analysis and its connections to a training needs assessment are highlighted in the case study above. A company keeps getting complaints about unhelpful and rude customer service agents from a specific call center. The objective of the performance analysis should be to determine the causes of and solutions to poor customer service at the call center. Making the assumption that call center employees need alternate training to improve customer service skills, then implementing such a training program, might solve the problem, but by skipping altogether or ignoring the results of a performance analysis, the need for that training cannot be logically or financially justified.

More importantly, training may not be the solution needed at all. In fact, the causes of the problem in the scenario include a faulty phone system that hangs up on customers who are put on hold or transferred; a lack of communication among product development, marketing, and customer service teams, resulting in a lack of knowledge about the organization's products and services; and a lackluster approach to interacting with customers on the phone. True causes cannot be inferred from the appearance of the problem; rather, causes must be systematically revealed through a performance analysis, just as training needs are revealed through a systematic assessment.

The overall solution system should emerge from the discovered causes. A faulty phone system should be replaced or repaired, and customer service agents must be trained in new products and services, and on any updates to the customer service software critical to their success. It is likely that the attitudinal and motivational problem of being lackluster in approach will be solved by eliminating the punishing aspects of the customer service agents' job. When training or educational interventions are indicated, they should be designed and delivered, but only after training needs assessment is conducted.

How do the results of the analysis translate into giving the client results? Remember, the ultimate purpose of any performance analysis is to help the organizational leaders align the real with the ideal. The performance analyst must clearly present the causes and solutions for the gap between what is and what could be. The analysis must clearly connect causes and solutions. These exact relationships are the foundation for the solution system and, in cases of gaps of skills, knowledge, and information, what moves the analyst from the broader focus of the performance analysis to the focused investigation, the training needs assessment. After discussions of data analysis, solution systems as a whole and training solutions in particular, the next section of this chapter moves into a more detailed discussion of a training needs assessment.

Analyzing Data

Regardless of the model used, the performance analyst is frequently tasked to begin researching solutions during the performance analysis process. That is, first analytical steps in planning the changes that will close the gap are taken well before the performance analysis is completed, first steps such as conducting a concurrent training needs assessment. Both processes depend on data analysis, which is touched on below and more comprehensively by Jeanne Hites Anderson in Chapter Four of this volume.

The analysis of the information, the causal analysis, is essentially the search for connections between what is happening and what the organizational leaders would like to see happening. During the process of analysis, specific questions are asked of data sets; then accepted, preferably validated, methods are used to discover the answers to the questions. The process is also iterative: along the way, more questions emerge from truths or trends that show up, and sometime analysis reveals the need for more data. It can also be alluring, seducing the analyst into hours of unnecessary examination because, while he may be onto something, he is woefully off-track for the specified purpose of the analysis. The project becomes paralyzed due to ever-expanding data collection. The specified purpose of the analysis as outlined in the plan is the beacon for the analyst tempted to over-do it; for each question, enough data must be analyzed so that the truths are internally validated or triangulated, and then the analyst must move on to the next question. If this is not systematically approached, another form of analysis-paralysis is likely; overwhelmed by the enormity, analysts simply stop performing well or at all. A well-known proverb is commonly invoked in discussions of curing analysis-paralysis: How do you eat an elephant? One bite at a time. Here, too, is where having a clear plan of action is essential; it acts both as a guide and prompt to stay on task. The analyst must be clear both on what information is being sought and on what the consequences are of an incorrect inference.

Issues in data analysis center around the selection of appropriate data analysis methods and the ethical calls made when performing analyses. In general, the way in which a data set is manipulated must fit with the kind of data it is and the kinds of questions being asked. If one were looking for trends in employee absenteeism and had numerical as well as written survey data and ways to parse the data into appropriate subsets, it would be appropriate to employ both descriptive, statistical, and qualitative analyses.

Ethical calls would initially seem to be out of the analyst's realm, but they're not. Data sets can be manipulated and massaged in quite acceptable ways. Validated methods for replacing missing data or repartitioning data subsets exist in order to compensate for imperfections in the data collection process—incomplete surveys or imbalanced populations. The application of these compensatory processes requires judiciousness on the part of the analyst. The biggest danger is coaxing out truths that are not there from data sets that essentially have been mistreated. This mistreatment is not limited to numerical data; written responses can also be subject to bad ethical calls and end up poorly analyzed. Discounting any submission from a model performer that is rife with misspellings and bad grammar because model performers should be literate, for example, results in a data set not representative of the model performers, but of the subset of model performers who have the ability to submit written information in a format deemed acceptable to the performance analyst's biased filter.

The Whole Solution System

Once the data has been gathered and the analysis completed, the performance analyst must create a solution system. Rossett (1999) defines a solution system as "an array of interventions that, when strategically combined, increase human performance in the workplace. Decisions about the nature of a solution system are based on causes and drivers and determined during the performance analysis" (p. 228). The solutions should develop from the truths revealed in analysis. If a group of employees is habitually tardy because they take a bus that arrives later than the start time of their work day—their choice being that or arrive an hour early for work—a solution might be to offer flexible work hours, permitting a fifteen minutes later or one hour earlier start and end to the work day.

A systems scholar, Gharajedaghi (2006) outlines critical features of a solution design, features that are applicable to a solution system designed for a performance analysis: the idealized solution design is constrained by technological feasibility, operational viability, and learning and adaptation, while recognizing the boundaries, environment, and purpose of the organization with the intention of improving functions, structure, processes, and performance criteria and measures. In the above example, if the analyst recommends that employees are allowed flexibility in work hours, the analysis must consider two things: that this policy must extend to all employees, and the impact of the irregular numbers of workers available at the beginning and end of the day. In order to be effective, the solution—flexible work hours—must be part of a larger system of solutions that considers the organization as both a whole and a series of interdependent parts.

Figuring out solutions is made easier with Rossett's (1987, 1999) distillation of Gilbert's (1996) and Binder's (1998) models into the concept of drivers. At the heart of performance is a set of drivers: ''Everything that it takes to enable performance to *grow*'' (Rossett, 1999, p. 38).These drivers, according to Rossett (1987, 1999), fall into one of four categories and are similar to the individual elements of the six boxes model. Rossett (1999) identifies a solution for each kind of driver, summarized in Table 5.4.

Harless's taxonomy of interventions is another useful tool in aligning solutions with the problem revealed by the performance analysis: (1) employee selection and assignment, (2) work process re-engineering, (3) work environment re-engineering, (4) motivation, incentive, and attitudinal, and (5) skills knowledge, and information (Langdon & Whiteside, 1997). Regardless of the approach to determining the appropriate solution, the performance analyst must ensure that it is the most appropriate as indicated by the data.

Training Solutions

Solutions for gaps in skills, knowledge, and information increase capacity and provide information that is readily usable to the performer (Rossett, 1999). Typically, training, job aids, and education are effective solutions to gaps in

Table 5.4 Summary of Drivers and Solutions

Driver	*Primary Probable Solution*
Skills, Knowledge, Information	Training, Job Aid, Education
Motivation	Training, Education
Environment	Job Process; Technology Revision
Incentives	Improved Recognition; Management Development

Table 5.5 Description of Situations and Solutions

If the performers . . .	*Use the following solution . . .*
Don't know it, can't do it	Training
Can do it but need support	Job aid
Can't anticipate it	Education

Synthesized from Rossett, 1987, and Clark and Estes, 2002

skills, knowledge, and information. Table 5.5 provides a set of descriptions that can help guide solution generation when it comes to training, job aids, and education. Often the performance analyst begins to investigate training needs as a gap in skills, knowledge, and information becomes evident during the causal analysis. This *in situ* approach to training needs assessment is appropriate only when training is the most likely solution to the problem revealed by a performance analysis. Without a training needs assessment, there is a great risk of addressing the gap in the wrong way, most commonly by creating training when it's unnecessary or by creating the wrong training.

Performance gaps are usually unique to a specific job function, requiring specifically designed training solutions; furthermore, the gaps revealed by performance analysis are typically not able to be addressed via extant training. Because of the uniqueness of performance gaps, a great deal of what's usually found in the corporate training directory and in a great many management development programs is not a viable solution; the performance analyst must move forward with a training needs assessment in order to develop a gap-specific training solution. When the gap calls for technical training, it's a fairly straightforward process to develop the unique training from a thorough training needs assessment.

More difficult are context-specific gaps in skills such as interpersonal interactions, communication, teamwork, or impromptu thinking. In general, if performers are not performing prescribed, serial tasks that require, at most, a decision tree, rather than being expected to interact with novel situations and create innovative solutions, the performance gap most likely resides in the performers' ability to solve ill-structured problems. A discussion of performance improvement needs that require problem solving is found later in this volume (Jonassen, Chapter Eleven). Non-training solutions are discussed in the second volume of this series. When training is part of the solution system, regardless of the category of performance problem, another analysis must be undertaken: a training needs assessment. This next process articulates the gap revealed by the performance analysis.

TRAINING NEEDS ASSESSMENT: A CORE PROCESS

A training needs assessment is the next step when a performance analysis produces a solution system that recommends training or educational interventions. This type of analysis is conducted by a wide variety of analysts: human resource professionals, training professionals, and instructional designers. At a minimum, a training needs assessment should pinpoint the problem or set of problems, confirm or flesh out the problem, and determine solutions (Gupta, 1999). Once training has been identified as a solution, learner, setting, and task analyses should be conducted to better guide the design and development of the training. Because of the same pressures of time and resources that trigger the training needs assessment *in situ*, the beginnings of these three analyses are often begun during the training needs assessment, thus providing the instructional design team with a solid foundation of data to augment.

As with a performance analysis, there is no one monolithic methodology—the way in which the needs assessment is conducted is dictated by the circumstances and is responsive to those circumstances. Just as in a performance analysis, the analyst must begin with a plan for the training needs assessment, one that outlines the steps and acts as a compass. Combining the six parts of conducting a needs assessment (1 to 6, in Table 5.6) (Richey, Fields, & Foxon, 2001) with the steps of a needs assessment plan (a through g in Table 5.6) (Foshay, Silber, & Westgaard, 1986) creates a blueprint for a training needs assessment process that is easy to follow and ensures a systematic approach; Table 5.6 provides an overview of the dimensions of the plan and the input sources for each stage. The next sections discuss each of the steps in the process from three perspectives: making the plan, carrying out the process, and explicating an example.

Describe the Problem

The first step in a training needs assessment, like with a performance analysis, is understanding and delineating the real and the ideal specifically in terms of the lack of skill and knowledge of the performers, the target audience. Both the objective of the analysis and the target audience should be identified before beginning the training needs assessment and should be indicated by the results of the performance analysis. Just as with a performance analysis, a clearly stated objective keeps the analysis on track; the objective should clearly indicate the group or groups of performers directly related to the problem.

Continuing with the call center customer service example, where customer service agents (CSAs) will be using a new phone system and have a lack of knowledge about an organization's products/services and a lack of skill navigating the customer support screens, the analyst conducting a concurrent training needs assessment would identify objectives of the training needs

Table 5.6 Needs Assessment Plan Dimensions and Sources

Needs Assessment Plan Dimension	*Source*
1. Describe the problem and its dimensions, identifying the discrepancies between current and desired performance.	This comes directly from the performance analysis, most likely a job analysis.
a. Objectives	The objective for the training needs assessment should clearly state the desired outcome of the assessment.
b. Target audience	This is the performer or group of performers who may need training and is also indicated by the performance analysis.
2. Clarify the varying perceptions of need and their implications.	Performers directly and indirectly related to the performance gap should be consulted in order to calibrate the needs assessment in a realistic understanding of the vantage points of those employees involved.
3. Select and use appropriate needs assessment tools and techniques.	Largely circumstance driven, the performance analyst should select familiar and accessible tools and techniques.
c. Sampling procedures	Performance analysis should indicate who should be sampled.
d. Data collection methods	Appropriate to data needed
e. Specifications for instruments and protocols	Appropriate to data needed
f. Methods of data analysis	Appropriate to data gathered
g. Description of how decisions will be made based on the data	An extant decision model or one created for the assessment should drive final determinations
4. Determine the possible causes of the problems and potential solutions.	Based on the results of the data collection and analysis
5. Recommend and advocate non-instructional solutions where appropriate.	Based on the results of the data collection and analysis
6. Complete a cost/benefit analysis for recommended solutions.	The performance analysis final results may provide the cost of the problem, and the cost of the recommended training solutions should be calculated based on the costs of training design, development, and delivery.

Synthesized from Foshay, Silber, and Westgaard, 1986; Richey, Fields, and Foxon, 2001 assessment before continuing. The two-fold objective of the training needs assessment is to verify that formal training for CSAs covering products and services as well as how to use their computer system is needed. Ideally, CSAs should (1) have accurate knowledge of the organization's products and services and (2) know how to navigate the customer service screens.

Clarify the Varying Perceptions of Need and Their Implications

Before the data collection begins in earnest, understanding the affective and practical contexts of the problem can be a great asset to the analyst. Rossett (1987) refers to this process as one that seeks feelings from the target audience, specifically how they feel about the problem—its importance and value to their work and to the company—and how they perceive others to feel—their understanding of atmosphere surrounding the problem. This information helps to shape the needs assessment plan by identifying possible barriers to training.

Depending on the size of the organization, a series of short interviews—causally conducted—may suffice or a more robust endeavor may be undertaken—a questionnaire or survey that would serve to both provide data about the perceptions of training needs and introduce the upcoming training needs assessment. This information harkens forward to the affective component of a setting analysis, so it may be prudent to combine these steps. This information may already be available from the performance analysis. The implications for the analyst include how the training needs assessment is conducted, that is, the methods selected. For example, sending out a survey by e-mail may be more successful with a group of employees who have been repeatedly tapped for one-on-one interview and focus groups with little realized results.

For example, if customer service agents at the poorly performing call center have been subjected to multiple training initiatives on good customer service while knowing all the while that a large portion of the problem is technical rather than practical and that the practical parts are related to products/services and computer updates, their attitudes or feelings about training and even providing great customer service may be negatively impacted. If the call center leaders are ill-informed or Ludditish about the major role technology plays in the effective performance of good customer service, their attitudes or feelings about the problem may be biased toward a human performance solution, such as training, when an environmental solution coupled with training would better solve the problem.

Select and Use Appropriate Needs Assessment Tools and Techniques

Data gathering, as in performance analysis, is a process of systematically gathering data from extant sources and individuals directly related to the performance problem as defined by the objective. A training needs assessment plan should identify sources of data—the sample set of people and organizational artifacts—that will reveal the exact nature of the training or educational needs. Sources typically include the subset of performers involved in the performance problem; other groups in the organization impacted by the problem or who might have insight into training needs; work processes—documented and observed; policies defining and constraining the performance; and support systems. Based on what is available from the performance analysis, the analyst decides what additional data should be gathered. Next the analyst selects tools and techniques, then determines the methods for analysis and decision making appropriate to the data, tools, and techniques. A more detailed discussion of sampling procedures, data collection methods, instruments and protocols, and methods of data analysis are found in Chapter Four of this book.

During this step, the analyst under pressure of time and resources will frequently plan for broad-scope learner, setting, and task analysis, concurrently gathering data useful for both the determination and the development of the unique performance gap and the possible training. Table 5.7 outlines the most common data sources for a training needs assessment—the words in parentheses represent the perspective of the concurrent analyses—and Table 5.8 outlines the most common methods used.

Performer/Learner Analysis. Learner analysis refers to an examination of characteristics and individual differences that may impact on the design of the training solution (Dick, Carey, & Carey, 2001; Rothwell & Kazanas, 1998). This profile can help the analyst verify the need for training. If, for example, performers are found to lack required entry behaviors, not only does the analyst have valuable information that helps determine the need for training, the design

Table 5.7 Data Sources and Descriptions

Data Source	*Description*
Results from performance analysis	Identified gap and organizational level data that triggered training needs assessment
Performers (learners) (task)	Model and average performers who enact job functions associated with the performance problem
Supervisors, co-workers, customers, vendors (task)	People who directly or indirectly interact with performance problem
Environment (setting) (task)	Where the performance problem occurs
Records/reports	Documents directly related to the performance, performers, or the problem
Work samples (task)	Output of performance

team begins the process with a clearer understanding of the scope of the subsequent, more detailed analyses needed to design the specific training intervention. (See Table 5.9.)

Setting Analysis. Setting analysis examines factors in the performance environment so that instruction will be designed in a way appropriate to the performance environment (Dick, Carey, & Carey, 2001; Rothwell & Kazanas, 1998). Two general areas the broad analysis should focus on are the physical and the affective environments. The physical environment may hold clues to the performance gap not uncovered in the performance analysis. Additionally, this information can serve the design team, helping them to design instruction that is appropriate to the setting. For example, if performers do not have access to calculators, learners should not either. (See Table 5.10.)

The affective environment can influence the ill-structured performance problems discussed previously in this chapter. If teamwork skills are under analysis, specifically hand-off between team members, it is critical to be aware of the tone of the work environment and discover the factors involved in creating that tone. Workers who are lackluster in their application of hand-off protocols may be emulating the slip-shod approach of their superiors. Workers may not, in fact, need training on process; rather, their superiors may need training and possibly development.

Task Analysis. Task analysis details the way in which the job function should be performed (Dick, Carey, & Carey, 2001; Rothwell & Kazanas, 1998). More specific than a job analysis but not as detailed as a content analysis, a task analysis performed at this stage highlights the difference between the actual performance and the ideal. The analyst conducting the

Table 5.8 Method and Descriptions

Method of Gathering Data	*Description*
Observation	Document the ways in which performers complete their jobs and what happens when they do
Interviews/focus groups	Meet with relevant employees who directly or indirectly interact with performance problem either individually or in groups
Questionnaires	Text-based way to gather a standard set of information
Selection of organizational artifacts	Informed choice to analyze records, reports, and work samples

Table 5.9 Analysis of Performer/Learner Characteristics

Performer/Learner Characteristic	*Influence on Design*
Existing skills, knowledge, and attitudes	Orients training to the right set of performance objectives; informs the scope of the training solution (number of course, breadth of content, content organization, chunking, cognitive load, complexity of interaction)
Education and ability	Orients training to the right set of performance objectives; informs the scope of the training solution (number of course, breadth of content); informs content choices; informs delivery choices (face to face, web)
Demographic characteristics (age, ethnicity, and gender)	Collect only if necessary; need for data dictates the way the data informs training solution
Physiological characteristics (general physical condition)	Informs level of dexterity needed for training (a ropes course, for example)
Past successful learning experiences	May inform content; may inform delivery
Geographical location	May inform content; may inform delivery
Value system	May inform content
Life/career stage	May inform content; may inform delivery
Motivation	Informs content in terms of imbedded learner-motivation strategies that help to create ownership on part of learner which impacts transfer
Attitudes toward organization	May inform content
Group characteristics	May inform content; may inform delivery

training needs assessment needs to have a faceted understanding of the task in order to recommend that the training be developed. Like with learner and setting analysis, the instructional design team will conduct subsequent analyses—perhaps a finer-grain task analysis and definitely a content analysis—in order to design the precise training solution. (See Table 5.11.)

In the ongoing example, the two-fold objective of the training needs assessment is to verify that formally training CSAs in products and services as well as using their computer system is needed. Ideally, CSAs should (1) have accurate knowledge of the organization's products and services and (2) know how to navigate the customer service screens. Two strands of data need to be collected: one on the lack of knowledge about the organization's products and services

Table 5.10 Setting Analysis

Setting Analysis Factor	*Influence on Design*
Physical environment	Compatibility of training environment with workplace; informs delivery choices (face to face, web)
Organizational climate	Orients training to the right set of performance objectives; informs the scope of the training solution (number of courses, breadth of content); informs content choices; informs delivery choices (face to face, web)

and one on the lack of skills navigating the customer service screens. In addition to the learner, setting, and task analysis, data that details the current process of training CSAs on new products/services and computer use should be gathered. Model performers should be identified and their work processes and

Table 5.11 Task Analysis

Task Analysis Factor	*Influence on Design*
List of subtasks	Orients training to the right set of performance objectives; informs the scope of the training solution (number of courses, breadth of content); informs content choices; informs delivery choices
Frequency of task	Informs solution selection, if the task is completed infrequently, a job aid may be the better choice; informs content choices
Difficulty of task	Informs the scope of the training solution (number of courses, breadth of content); informs content choices
Importance/consequence of task	Informs the scope of the training solution (number of course, breadth of content); informs content choices
Task triggers (initiation/ completion)	Informs content choices
Standards	Orients training to the right set of performance objectives
Tools/technology needed	Informs content choices; informs delivery choices
Participants in task completion	Orients training to the right set of performance objectives; informs the scope of the training solution (number of courses, breadth of content); informs content choices; informs delivery choices

experiences should be documented for comparison to those of their less successful counterparts.

Determine the Possible Causes of the Problems, Potential Solutions, and Recommend Non-Instructional Solutions (Where Appropriate)

After the needs assessment plan has been executed, in general, the data is examined for the exact deficit in skills and knowledge and corresponding causes of the lack of skills and knowledge. If the purpose of the training needs assessment is to verify the need for training on specific points, the data gathered in the previous step should clearly answer the question. Planning for analysis doesn't involve a lot of resources; logistically, it relies on the systematic way in which data is collected and the availability of any specialized software. More than anything, the analyst needs to plan for the time to conduct the analysis.

At the onset of a training needs assessment, the solution has generally been narrowed to a training or educational one, but if the data analysis begins to suggest otherwise, the analyst should consult the solution system from the performance analysis. Going back to the broader analysis informs the next step: expanding the training needs assessment, revisiting the performance analysis data to explore the possibility of other types of solutions, or perhaps even redirecting the performance analysis.

The cause of the call center CSAs' lack knowledge about new products or services is the way in which the organization distributes information. In this case, all new products and services are introduced in the monthly newsletter. The CSAs are not given any specific training on the new products or services, nor are they formally shown updates made to the computer system when new products or services are introduced. In essence, they have to learn while attempting to help customers.

Job aids would work since the changes to products and services are small and happen frequently. The current method of announcing these changes in a newsletter primarily written for the sales staff is not functional for the customer service agents. Training would be too big an endeavor for this problem; however, training is appropriate to bring the customer service agents up to speed about the quarterly updates on the computer system that are currently unannounced and unsupported, as shown in Table 5.12.

Complete a Cost/Benefit Analysis for Recommended Solutions

After a determination is made as to the appropriateness of training or educational solutions to the performance problems, the analyst needs to conduct a cost/benefit analysis before making the final recommendation for training. A cost/benefit analysis shows in quantitative terms how the organization will benefit from the training or educational initiative.

Table 5.12 Causal Analysis of Call Center Example

Problem	*Cause*	*Solution*
Lack of knowledge of products	Indirect communication of updates	Job aids that show changes: a crosswalk from old to new
Lack of skills with computer system	No communication of updates	Training module on changes to CSA software

Essentially, it finds and compares two figures: (1) the cost of the performance problem and (2) the difference between the cost of the performance problem and the direct and indirect costs of training. Occasionally, the *status quo* is less expensive than implementing training, for example, if the training would be rendered moot by an impending policy, procedural, or technology change. As with most organizational decisions, time must be a factor in determining the benefits of training; the immediate costs of training may be more than the status quo; however, the compounded costs of an ongoing problem must be considered.

In the case of the call center, this long-range perspective weighed heavily on the implementation of a training solution. The initial costs of creating a training module for the computer system can be amortized across subsequent training modules as, in this case, subsequent modules will largely cannibalize previous ones.

The spiral of analysis, from performance analysis to learner analysis, ensures that the right problems are being addressed in the right way. Training is justified only if it targets a clearly identified gap in skills or knowledge and only if the benefits outweigh the costs. At the end of this chapter are two exhibits generated to correspond to the TNA process outlined. One is a needs analysis checklist and the other a worked example of that checklist (see Exhibits 5.1 and 5.2). An expert performance analyst generates the tools he needs for the analysis and allows inquiry to inform and shape the process and the content. What is found from an interview with the performer shapes the questions asked of the supervisor.

Exhibit 5.1 Needs Analysis Checklist Job Aid

This core checklist can be tailored for the specific performance or training inquiry.

Describe the problem and its dimensions, identifying the discrepancies between current and desired performance.

(*Continued*)

Exhibit 5.1 (*Continued*)

Objectives

Target audience

Description of problem: *At the onset of the analysis, document a robust description of the problem.*

Objectives of the analysis: What are the envisioned results?

Target audience for analysis: Who will be at the center of the inquiry?

Clarify the varying perceptions of need and their implications.

Pertinent perspectives on the apparent need: *Gather data from appropriate sources. For a performance analysis, reference models like the six boxes model (Binder, 1998) or Rossett's (1999) four opportunities to narrow and focus.*

Who/What	How	Why
Identify data source	Identify data collection method	Document relationship to inquiry
____________	____________	____________
____________	____________	____________

Clarify implications for analysis: *Perceptions shape affect and logic; consider how the various perspectives help shape and define both the problem, the problem context, and the search for a solution.*

Select and use appropriate needs assessment tools and techniques

Sampling procedures

Data collection methods

Specifications for instruments and protocols

Methods of data analysis

Description of how decisions will be made based on the data

Sampling procedures: *Determine how the various sets and subsets of stakeholders, performers, experts, artifacts, and literature will be selected. For a performance analysis, reference models like the six boxes model (Binder, 1998) or Rossett's (1999) four opportunities to narrow and focus.*

Who/What	**How**	**Why**
Identify data source	Identify data collection method	How will the data be understood?
__________	__________	__________
__________	__________	__________

Description of how decisions will be made based on the data: *Document the plan for the next steps, triggers for revision, ex: A more detailed analysis will be planned and initiated if . . .*

Determine the possible causes of the problems and potential solutions.

List specific causes and possible solutions for each identified problem. The level of detail is scope-specific. Identify the type of driver associated with the cause in order to identify the category of probable solution.

Cause →	**Problem →**	**Solution**
What's making the problem manifest?	Name the problem (not describe) for specificity.	Derived from both the cause and the manifestation of the problem.
__________	__________	__________
__________	__________	__________

Driver →	**Primary Probable Solution**
Skills, Knowledge, Information	Training, Job Aid, Education
Motivation	Training, Education
Environment	Job Process; Technology Revision
Incentives	Improved Recognition; Management Development
__________	__________
__________	__________

(*Continued*)

Exhibit 5.1 (*Continued*)

Recommend and advocate non-instructional solutions where appropriate.

Training is not always the answer, and a lack of knowledge and skills is not always the problem. When it isn't, the analyst must identify and determine appropriate solutions.

Driver →	***Primary* Probable Solution**
Skills, Knowledge, Information	Training, Job Aid, Education
Motivation	Training, Education
Environment	Job Process; Technology Revision
Incentives	Improved Recognition; Management Development
________________	____________________________
________________	____________________________

Complete a cost/benefit analysis for recommended solutions.

A cost/benefit analysis compares the cost of the solution to the anticipated benefits from it—the primary benefit being the elimination of loss because of the problem. Greer's (1992) ID Project Management: Tools and Techniques *can help the analyst determine an estimate for the cost of a training solution from a project perspective. Weighed against the cost of the problem, the training solution may or may not be fiscally justified.*

Exhibit 5.2 Worked Example of Needs Analysis Checklist Job Aid

Describe the problem and its dimensions, identifying the discrepancies between current and desired performance.

Objectives

Target audience

Description of problem: *At the onset of the analysis, document a robust description of the problem.*

The performance analysis that investigated complaints about unhelpful and rude customer service agents (CSAs) from a specific call center revealed that the agents were using a faulty phone system that hung up on calls, resulting in the perception of rudeness. A new phone system will replace the faulty one. The customer service agents were also evaluated as being unhelpful, which the performance analysis traced to a lack of knowledge about an organization's products/services and a lack of skill navigating the customer support screens.

Objectives of the analysis: What are the envisioned results?

This training needs assessment will verify that two units of formal training for CSAs are needed

Products and services

Using their computer system.

Target audience for analysis: Who will be at the center of the inquiry?

The customer service agents.

Clarify the varying perceptions of need and their implications.

Pertinent perspectives on the apparent need: *Gather data from appropriate sources. For a performance analysis, reference models like the six boxes model (Binder, 1998) or Rossett's (1999) four opportunities to narrow and focus.*

Who/What	**How**	**Why**
Identify data source	Identify data collection method	Document relationship to inquiry
CSAs	Interview, survey	The CSAs are at the core of the problem.
CSA supervisors	Interview	The supervisors can speak to the training efforts to date and the impact of the problem/ ROI on a training solution.

Clarify implications for analysis: *Perceptions shape affect and logic; consider how the various perspectives help shape and define both the problem, the problem context, and the search for a solution.*

Perceptions: Customer service agents are not eager for training because they have been subjected to multiple face-to-face training initiatives on good customer service. The manager of this call center, suspicious of computer-based training, has supported half-day training sessions but is beginning to become leery about training as a way to address what she has seen as a problem with customer service skills. The CSAs' supervisor expressed tenuous lines of communication with management and has not reported that the problem is technical, not practical, and that the practical parts are related to products/ services and computer updates. The CSAs expressed a desire for some formal

(*Continued*)

Exhibit 5.2 (*Continued*)

conduit of information, but not a pull-out training session, about updates to products/services and their computer systems.

Implications: History of training solutions may indicate a bias toward training solutions of a certain type (face to face, instructor-led). The potential resistance to training means that if training is needed, alternative delivery methods, possibly intranet based, should be explored both for efficiency and for CSA buy-in.

Select and use appropriate needs assessment tools and techniques

Sampling procedures

Data collection methods

Specifications for instruments and protocols

Methods of data analysis

Description of how decisions will be made based on the data.

Sampling procedures: *Determine how the various sets and subsets of stakeholders, performers, experts, artifacts, and literature will be selected. For a performance analysis, reference models like the six boxes model (Binder, 1998) or Rossett's (1999) four opportunities to narrow and focus.*

Who/What	How	Why
Identify data source	Identify data collection method	How will the data be understood?
CSAs	Observation, interview, focus group (if needed)	Qualitative analysis: show actual gaps
Supervisors	Observation, interview	Qualitative analysis: gaps, context
Management	Interview	Qualitative analysis: influence solution development
Product/service types, specifically changes to	Internal documentation	Quantitative analysis: number and scale (small < 5) of changes
Customer support screens, changes to	Internal documentation	Quantitative analysis: number and scale (small < 5) of changes

Description of how decisions will be made based on the data: *Document the plan for the next steps, triggers for revision, ex: A more detailed analysis will be planned and initiated if . . .*

Content analysis should be conducted if the initial results of the data from the CSAs indicate a lack of skill.

Determine the possible causes of the problems and potential solutions. *List specific causes and possible solutions for each identified problem. The level of detail is scope-specific. Identify the type of driver associated with the cause in order to identify the category of probable solution.*

Right now, CSAs expected to find and then read a copy of the company newsletter, extrapolate the pertinent information, and integrate it into their workflow. Often, the trigger for searching out a newsletter is a change to the customer support screens they use in the course of a support call.

Cause →	**Problem →**	**Solution**
What's making the problem manifest?	Name the problem (not describe) for specificity.	Derived from both the cause and the manifestation of the problem.
Informal distribution of information	CSAs unaware of changes to products and services	Small changes: job aid
Large-scale changes, new products/services: product overview session (training)		
Unannounced/ unsupported updates	CSAs unable to fluidly navigate customer support screens	Small changes: job aid
Large-scale changes, new support screens: computer-based training, and JIT support module		

Driver →	**Primary Probable Solution**
Skills, Knowledge, Information	Training, Job Aid, Education
Motivation	Training, Education
Environment	Job Process; Technology Revision
Incentives	Improved Recognition; Management Development

Recommend and advocate non-instructional solutions where appropriate.

(*Continued*)

Exhibit 5.2 (*Continued*)

Training is not always the answer, and a lack of knowledge and skills is not always the problem. When it isn't, the analyst must identify and determine appropriate solutions.

Driver →	***Primary* Probable Solution**
Skills, Knowledge, Information	Training, Job Aid, Education
Motivation	Training, Education
Environment	Job Process; Technology Revision
Incentives	Improved Recognition; Management Development

The analysis revealed that the lack of knowledge in some cases is a simple lack of information, that is, small changes to existing products/services or the customer support screens. In the case of small changes, where small is defined as five or fewer, a job aid that provides a crosswalk from the old product/service to the new ones is more efficient than a training session. Minor changes to the customer support screens can be addressed with informational pop-ups/rollovers.

Complete a cost/benefit analysis for recommended solutions.

A cost/benefit analysis compares the cost of the solution to the anticipated benefits from it—the primary benefit being the elimination of loss because of the problem. Greer's (1992) ID Project Management: Tools and Techniques *can help the analyst determine an estimate for the cost of a training solution from a project perspective. Weighed against the cost of the problem, the training solution may or may not be fiscally justified.*

Cost of performance problem: The estimated cost of the loss of customers due to poor customer service averaged $45,000 over the last two quarters.

It is impossible to disaggregate the technical (faulty phone system) causes from the skill/knowledge/information-based ones in terms of parsing the $45,000 to determine the portion that job aids/training might provide. Further, the cost of the new phone system plus service contract, a solution from the performance analysis, far exceeds $45,000. The estimated cost of job aids and training development, based on the number and scale of changes made in the last year, is $20,000.

Management might choose to wait for quarterly results following the installation of the new technology before deciding to implement the solutions recommended by the TNA in order to better calculate the ROI on the training.

CONCLUSION

Performance analysis is the process of understanding a situation with a specific goal in mind, be it improved performance or increased profits. A performance analysis must first establish the desired ideal, then build an honest understanding of the real situation. From that, a solution system that is causally linked to identifiable gaps is derived. A possible solution is training, but developing training should also depend on a clear understanding of the need for training. A training needs assessment should arise from a performance analysis and should be in response to a verifiable lack of skills, knowledge, or information.

References

Binder, C. (1998). The six boxes: A descendent of Gilbert's behavior engineering model. *Performance Improvement*, *37*(6), 48–52.

Clark, R. E., & Estes, F. (2002). *Turning research into results: A guide to selecting the right performance solutions*. Atlanta, GA: CEP Press.

Dick, W., Carey, L., & Carey, J. O. (2001). *The systematic design of instruction* (5th ed.). New York: Addison-Wesley, Longman.

Foshay, W., Silber, K., & Westgaard, O. (1986). Instructional design competencies: The standards. Iowa City, IA: International Board of Standards for Training, Performance, and Instruction.

Gharajedaghi, J. (2006). *Systems thinking: Managing chaos and complexity: A platform for designing business architecture*. Boston: Butterworth-Heinemann.

Gilbert, T. F. (1996). *Human competence: Engineering worthy performance* (tribute ed.). Washington, DC: International Society for Performance Improvement.

Greer, M. (1992). *ID project management: Tools and techniques for instructional designers and developers*. Englewood Cliffs, NJ: Educational Technology Publications.

Gupta, K. (1999). *A practical guide to needs assessment*. San Francisco: Jossey-Bass.

Harless, J. (1973). An analysis of front-end analysis. *Improving Human Performance: A Research Quarterly*, *4*, 229–244.

Harless, J. (1975). *An ounce of analysis is worth a pound of objectives*. Newnan, GA: Harless Performance Guild.

ISPI's Performance Technology Standards. (2002). Silver Spring, MD: International Society for Performance Improvement.

Jonassen, D. H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology: Research and Development*, *45*(1), 65–95.

Jonassen, D. H. (2000). Toward a meta-theory of problem solving. *Educational Technology: Research & Development*, *48*(4), 63–85.

Kaufman, R. A. (1985). Needs assessment, needs analysis, objectives and evaluation. *Performance & Instruction*, *24*(6), 21.

Kaufman, R. A. (2000). *Mega planning: Practical tools for organizational success*. Thousand Oaks, CA: Sage.

Kaufman, R. A. (2005). Defining and delivering measurable value: A mega thinking and planning primer. *Performance Improvement Quarterly*, *18*(3), 6–16.

Kaufman, R. A., & Valentine, G. (1989). Relating needs assessment and needs analysis. *Performance & Instruction Journal*, *28*(10), 10–14.

Langdon, D. G., & Whiteside, K. S. (1997). The performance technologist's role in interventions: An interview with Joe Harless. *Performance Improvement*, *36*(10), 36–38.

Mager, R., & Pipe, P. (1984). *Analyzing performance problems* (2nd ed.). Belmont, CA: Fearon.

Pershing, J. A. (2006). Human performance technology fundamentals. In J. A. Pershing (Ed.), *Handbook of human performance technology: Principles, practices, potential* (3rd ed.). San Francisco: Pfeiffer.

Piskurich, G. M., Beckschi, P., & Hall, B. (Eds.). (2000). *The ASTD handbook of training design and delivery: A comprehensive guide to creating and delivering training programs—Instructor-led, computer-based, or self-directed*. New York: McGraw-Hill.

Richey, R. C., Fields, D. C., & Foxon, M. (2001). *Instructional design competencies: The standards* (3rd ed.). Syracuse, NY: ERIC Clearinghouse on Information and Technology.

Rossett, A. (1982) A typology for generating needs assessment. *Journal of Instructional Development*, *6*(1), 28–33.

Rossett, A. (1987). *Training needs assessment*. Englewood Cliffs, NJ: Educational Technology Publications.

Rossett, A. (1999). *First things fast: A handbook for performance analysis*. San Francisco: Pfeiffer.

Rossett, A. (2006). Analysis and more. In J. A. Pershing (Ed.), *Handbook of human performance technology: Principles, practices, potential* (3rd ed.). San Francisco: Pfeiffer.

Rothwell, W. J., Hohne C. K., & King, S. B. (2007). *Human performance improvement: Building practitioner competence (improving human performance)* (2nd ed.). Boston: Butterworth-Heinemann.

Rothwell, W. J., & Kazanas, H. C. (1998). *Mastering the instructional design process: A systematic approach* (2nd ed.). San Francisco: Pfeiffer.

Rummler, G. A., & Brache, A. P. (1990). *Improving performance: How to manage the white space on the organization chart*. San Francisco: Jossey-Bass.

Svenson, R. (2006). Requirements: The bridge between analysis and design. In J. A. Pershing (Ed.), *Handbook of human performance technology: Principles, practices, potential* (3rd ed.). San Francisco: Pfeiffer.

Walonick, D. S. (2004). *Survival statistics*. Minneapolis, MN: StatPac.

Watkins, R. (2006). Aligning human performance technology decisions with an organization's strategic direction. In J. A. Pershing (Ed.), *Handbook of human performance technology: Principles, practices, potential* (3rd ed.). San Francisco: Pfeiffer.

Watkins, R., & Kaufman, R. A. (1996). An update on relating needs assessment and needs analysis. *Performance Improvement*, *35*(10), 10–16.

Zemke, R., & Kramlinger, T. (1982). *Figuring things out: A trainer's guide to needs and task analysis*. Reading, MA: Addison-Wesley Longman.

CHAPTER SIX

Behavioral Task Analysis*

Herbert H. Bell
Dee H. Andrews
Wallace H. Wulfeck II

INTRODUCTION

Behavioral task analysis is a fundamental tool of the human systems professional for a variety of processes. In one form or another, task analysis plays an important role in mission analysis, organizational design, job design, system design, quality improvement, personnel selection, training, and evaluation.

Simply put, task analysis helps us understand what people do when they successfully accomplish their work. Mager (1988) described task analysis as ''a collection of techniques used to help make the components of competent performance visible'' (p. 29). Behavioral task analysis focuses on the behaviors people perform while doing their jobs. Typically, these behaviors are documented as discrete tasks or procedures individuals must accomplish to successfully perform a job (Jonassen, Hannum, & Tessmer, 1989). For training design, the task analysis process continues to identify conditions, actions, and standards for each task to be trained and each performance to be assessed.

It is impossible to cover all the details of task analysis in a single chapter. Indeed, there are a number of books devoted entirely to task analysis (Annett & Stanton, 2000; Carlisle, 1986; Hackos & Redish, 1998; Jonassen, Hannum, & Tessmer, 1989; Jonassen, Tessmer, & Hannum, 1999; Kirwan & Ainsworth, 1992). Because task analysis covers such a wide range of processes and involves a number of different disciplines, there are a variety of techniques for performing

* *Note:* The opinions expressed in this chapter are those of the authors and do not necessarily represent the official views or policies of the Department of Defense or the Services.

task analyses and a number of different ways to describe the resulting products. This complexity is reflected in the FAA Human Factors Workbench, which lists thirty-one different tools to support task analysis (Federal Aviation Administration, n.d.). Given the scope of task analysis, the large number of tools available for task analysis, and the different theoretical foundations and terms used by practitioners, task analysis may seem overwhelming to a newcomer.

The primary focus of this chapter is on behavioral task analysis for training. This chapter and Chapter Seven in this volume on cognitive task analysis by Villachica and Stone will provide the reader with a working knowledge of task analysis and an understanding of how task analysis helps people have the right tools, knowledge, and skills to perform their work successfully.

The military nurtured many of the task analysis concepts and methods used today in order to design complex systems and train large numbers of people to operate those systems. As a result, many of our examples involve military applications of behavioral task analysis. These same concepts have been adapted and used in a number of different areas such as organizational design and human-computer interaction across both public and private sectors.

This chapter begins by defining a few key concepts that are critical to understanding the nature of tasks and task analysis. Next, we present a brief overview of the development of task analysis and the distinction between behavioral task analysis and cognitive task analysis. Then we list the general questions that behavioral task analysis helps answer and describe the typical stages involved in conducting a behavioral task analysis. Following this general discussion, several cases illustrating specific applications are presented as well as some limitations and pitfalls. Finally, a brief overview of tools that are available to assist in conducting a task analysis is given.

FUNDAMENTAL CONCEPTS

What Is Behavioral Task Analysis?

Task analysis is applicable throughout the life cycle of a system. Meister (1985) identified various areas within the system life cycle where task analysis is especially useful and the types of questions addressed by task analysis. Table 6.1 presents an abbreviated summary of those areas and questions. As illustrated, the information gathered from a particular question is often applicable to several areas. For example, the system designer must provide operators timely and correct displays that cue the need to perform a particular task. Similarly, training developers must ensure that correct relationships between cues and control behaviors are established during training, and both the system designer and the training developer must identify criteria for successful performance.

Table 6.1 System Life Cycle Stages and Relevant Task Analysis Questions

Area of Interest	*Sample Questions*
Design	What tasks need to be performed and how should they be performed?
	What are the consequences of failing to perform a task or performing it inadequately?
	What is the order in which tasks must be performed?
	What information is needed to perform the task?
	What actions must the operator perform to accomplish the task?
	What coordination is required?
	What are the perceptual, cognitive, psychomotor, and physical demands?
	What errors are likely?
Staffing	How many people are required to perform the task?
	What knowledge, skills, abilities, and experience are required to perform the task?
Training	What behaviors underlie each task?
	How difficult or complex is the task?
	What information is necessary to perform the task?
	What are the criteria for successful performance?
	What are the consequences of not performing or inadequately performing a task?
	What is the relationship between various tasks?
	How frequently is the task performed?
Performance Evaluation	What are the criteria for successful performance?
	What are the consequences of poor performance?

Based on Meister, 1985

Tasks are performed to accomplish work for specific purposes under specific conditions. Behavioral task analysis involves collecting, abstracting, organizing, and reporting information about what people do in performing work. There are a number of different ways to collect data about tasks, abstract that data into organized categories, and present that data. Because of these differences, the terminology associated with task analysis varies across practitioners. Table 6.2 lists many of the terms commonly associated with task analysis and provides a brief definition of each term.

Table 6.2 Common Task Analysis Terms

System *acq*.	A composite of equipment, skills, and techniques capable of performing or supporting an operational role, or both. A complete system includes all equipment, related facilities, material, software, services, and personnel required for its operation and support to the degree that it can be a self-sufficient unit in its intended operational environment.
Task analysis *acq*.	A systematic method used to develop a time-oriented description of personnel/equipment/software interactions brought about by an operator, controller, or maintainer in accomplishing a unit of work with a system or item of equipment. It shows the sequential and simultaneous manual and intellectual activities of personnel operating, maintaining or controlling equipment, in addition to sequential operation of the equipment. It is a part of system engineering analysis where system engineering is required. The following taxonomy is used to inventory or analyze tasks, with mission and scenario conditions stated by the procuring activity and the remaining levels dependent on the current phase of system development and purpose (e.g., gross analysis of task analysis of critical tasks) for which the analysis is being conducted.
Mission	What the system is supposed to accomplish, e.g., combat reconnaissance.
Scenario/ conditions	Categories of factors or constraints under which the system will be expected to operate and be maintained, e.g., day/night, all weather, all terrain operation.
Function	A broad category of activity performed by a system, e.g., transportation.
Job	The combination of all human performance required for operation and maintenance of one personnel position in a system, e.g., driver.
Duty	A set of operationally related tasks within a given job, e.g., driving, weapon servicing communicating, target detection, self protection, operator maintenance.
Task	A composite of related activities (perceptions, decisions, and responses) performed for an immediate purpose; written in operator/maintainer language, e.g., change a tire.
Subtask	An activity (perceptions, decisions, and responses) that fulfills a portion of the immediate purpose within the task, e.g., remove lug nuts.
Task element	The smallest logically and reasonably definable unit of behavior required in completing a task or subtask, e.g., apply counterclockwise torque to the lug nuts with a lug wrench.

Department of Defense, 1999, pp. 30–32 in hierarchical order

This chapter uses the generic model and terms shown in Figure 6.1 to provide an organizing structure for the concepts associated with task analysis. This model reflects a synthesis of many theorists and practitioners, most importantly Annett (2004), Meister (1985), and Miller (1953, 1962). Its purpose is to provide a context for understanding behavioral task analysis and a means of describing behavior in a meaningful way.

System. As shown in Figure 6.1, work occurs within a system context to affect or control some specific part of the environment.

Vincente (1999) defines a system as ''a set of interrelated elements that share a common . . . [p]urpose'' (p. 9). There are four important ideas embedded within this definition.

1. A system exists for a purpose.
2. The system can be decomposed into elements or subsystems.
3. The overall purpose of the system determines the interrelationship of system elements.
4. The system can be viewed hierarchically.

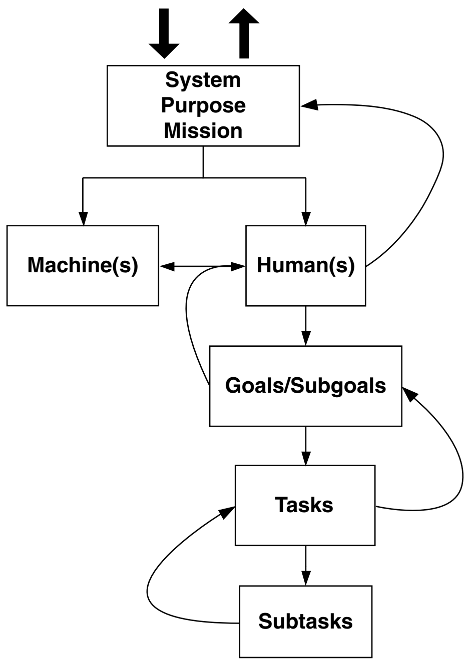


Figure 6.1 Generic Task Domain.

The system perspective allows us to understand the reason for performing a particular task, how it relates to other tasks, and the consequences of various levels of performance. Work is performed as part of a specific system comprised of multiple elements in order to enable the system to achieve its purpose.

A system perspective is necessary to understand tasks because it contextualizes human behavior in terms of purpose, identifies factors that influence the performance of tasks, and describes how specific activities contribute to the successful completion of work (Meister, 1985, 1989). Without the system view, it would be very difficult to do anything other than simply list the sequence of task behaviors. The system perspective provides a means to understand why and how specific tasks contribute to the successful completion of work.

Purpose. As noted above, a system exists to achieve a particular purpose. That purpose determines the functions various subsystems must perform. For example, the purpose of a particular combat aircraft might be to achieve and maintain dominance over land forces by disrupting their ability to continue fighting. This requires that, as a system, the combat aircraft include functions that enable specific altitudes, airspeeds, and maneuvers; processing certain radar information; and acquiring, tracking, and engaging ground targets. While these functions or various subfunctions may be assigned to either humans or machines, they must be present in order for the system to fulfill its purpose.

Mission. Although a system is designed for a particular purpose, that purpose is often stated in relatively global terms. A mission is a specified goal-oriented activity the system is intended to accomplish. For example, an aircraft designed to enable us to disrupt the fighting capability of enemy land forces could achieve its purpose by destroying their supply centers or by attacking enemy troops that are in close proximity with friendly ground forces. These represent two different missions, each of which is consistent with the system's purpose. However, these two different missions may change the specific weapons carried on the aircraft as well as the tactics, techniques, and procedures used by the aircraft's pilot. Therefore, the behavioral analyst must consider not only the system and its purpose but also the specific missions or various ways in which the system is used.

Job. A job is a collection of duties and tasks that are performed by *one individual*. The job is the basic unit used in carrying out the personnel actions of selection, training, classification, and assignment. A job may consist of several, or even many, different duties and tasks, and they need not be related to each other. For example, an electronics technician may also serve as a labor organizer, benefits counselor, and radar operator in addition to primary duties of electronic equipment troubleshooting and repair.

Task Environment. The task environment consists of the cues and conditions that influence how the operator performs a particular task. It includes environmental conditions such as temperature as well as the particular tools, displays, and controls the operator uses to perform the task.

Task. At the simplest level, a task is ''A single unit of specific work behavior with clear beginning and ending points that are directly observable or otherwise measurable'' (Department of Defense 2001a, ¶6.5.35). As indicated in Table 6.2, tasks are performed for a purpose. Tasks are viewed as system related activities, performed by humans, to achieve specific *goals* and *subgoals* that must be achieved for the overall system to fulfill its mission. Typically, these goals and subgoals are hierarchically nested within a mission and determine the specific tasks that must be performed at a particular time. For example, a commercial airline pilot has the overarching goal of safely transporting passengers to a specific destination. Nested underneath that goal are subgoals such as avoiding air turbulence and maintaining fuel efficiency.

Tasks typically have performance conditions, performance requirements, and performance criteria or standards that must be met in order for them to contribute successfully to mission accomplishment. Performance conditions are the environmental cues that initiate and guide action and factors that enable or constrain action. Performance requirements describe the types of actions and manner of task execution. Performance criteria or standards define the acceptable level of task performance necessary for successful task completion.

Miller (1962) distinguished between task description and task behaviors. Task descriptions describe the general nature of the work and specify the interactions between the operator, the other system elements, and the work environment within a systems framework. They ''describe what humans are expected to do'' (Nemeth, 2004, p. 187) and are typically stated as high-level functions such as ''to detect'' or ''to maintain.'' Task behaviors, on the other hand, describe what the operator must actually do in order to accomplish functions such as detect or maintain. Task behaviors describe in detail how work needs to be performed to accomplish a particular function and serve as the primary means for describing the specific instructional content and its sequencing. For the sake of brevity, a systematic distinction between task descriptions and task analysis will not be made. More detailed discussions of this difference are available in Jonassen, Tessmer, and Hannum (1999) and Meister (1985, 1989).

Evolution of Behavioral Task Analysis

Behavioral task analysis is hardly a new process. It is the product of industrial engineering, behavioral psychology, and systems analysis. The industrial engineering influence can be traced back to Frederick Taylor, Frank and Lillian Gilbreth, and others who systematically analyzed worker behavior during the early 20th century in an effort to increase factory output through the application of scientific management (Annett & Stanton, 2000). Their goal was to identify the most efficient way to accomplish manual work in order to increase worker productivity in the context of the nascent manufacturing technology of the time. For example, the Gilbreths developed a notational system called ''therbligs'' for ''analyzing the motions involved in performing a task . . . . as well as . . . delay'' in order to eliminate inefficient motions and wasted time (Ferguson, 2000, p. 1).

This early work demonstrated that jobs could be described as being comprised of a number of distinctive elements, and these elements could be decomposed into tasks and subtasks. It established the foundation for both job analysis and task analysis and identified a number of questions that are still important for task analytic methods. These questions include:

- What is the work performed?
- How do we measure the quality of the work?
- Under what conditions is the work performed?
- How is the work performed?
- What is needed to perform the work?
- How is work performance measured?

During the first half of the 20th century, psychologists, particularly in the United States, focused much of their attention on the analysis of observable behavior. Although competing schools of behaviorism differed in many ways (Bower & Hilgard, 1981), John B. Watson, Hull, and Skinner and their followers believed the foundation for a scientific understanding of behavior rested on the analyses of observable and measurable outcomes associated with specific stimuli, and they eschewed any appeal to unobservable mental processes. Much like the proponents of scientific management, attempts to apply behaviorism to the analysis of work focused on the conditions for and the consequences of behavior. For example, applied behavioral analysis focused on task performance as a chain of overt stimuli and responses (Gilbert, 1962). The result was a conception of human performance as a series of stimuli, responses, or actions, and consequences or outcomes in which these outcomes became the stimuli for subsequent responses. It essentially viewed a task as a linear stimulus–response–stimulus sequence and ignored the role of cognitive processes and knowledge in performing a task. The goal was to determine how to ensure the appropriate response to the specific set of stimulus conditions. This behavioral influence, which dominated American psychology for half a century, viewed work as the assembly of interchangeable parts and human performance as the assembly of specific actions in response to specific stimulus conditions.

Occupational/Job Analysis

Modern job and task analysis began during the 1940s and early 1950s. It was initiated during World War II when large numbers of people had to be trained quickly to operate complex equipment under difficult, life-threatening conditions. Engineers, psychologists, and military subject-matter experts were brought together to develop methods for improving the design of equipment and the methods of training people to use that equipment. Understanding how people used this complex equipment to accomplish specific tasks and how to train people to use that equipment led to modern ergonomics, human factors, and instructional systems theory (Koppes, 2006; Nemeth, 2004).

The first step in meeting these needs was to determine what tasks would have to be taught and to whom. This required a detailed listing of the tasks comprising a specific individual's job or duty. In the years after World War II, these techniques were found useful in organizational analysis (for example, what should a collection of people, such as a brigade or platoon, consist of and what tasks should be performed by which persons in the collection to achieve an overall mission?), and in the process of planning for new hardware systems (for example, what combat and support tasks must be performed and how should these be distributed over the number of crew personnel afforded by a ship's size and berthing facilities?). Once the tasks are identified, the analyst collects task information such as that shown in Table 6.3.

These occupational analysis functions were then consolidated in the 1950s and 60s into organized programs in the military, especially the Air Force. This resulted in establishment of the Air Force's Comprehensive Occupational Data Analysis Program (CODAP), which became a computational effort involving large-scale surveys, data collection, and analysis. The Navy had a similar effort called the Navy Occupational Task Analysis Program (NOTAP). The civilian sector also started similar efforts. Perhaps the largest of these is the activity of the Vocational-Technical Education Consortium of States (VTECS, n.d.), a consortium that conducts job analytic efforts for the purpose of designing career and technical education and training curricula.

Task Analysis and Instructional Systems Development

Behavioral task analysis plays a major role in the needs analysis portion of the training development process. This linkage between task analysis and training is most apparent in the formalization of what has come to be called instructional systems development (ISD).

The post-World War II formalization of human performance led to proceduralized methods for ISD (Branson, Rayner, Cox, Furman, King, & Hannum, 1975), also known as the ''systems approach to training,'' which by the late 1970s had been adopted by all U.S. military services and by many universities,

Table 6.3 Typical Information Collected During Task Analysis

Who performs each task?
For how long? What is the percent of time an individual spends performing the task?
How frequently is the task performed?
To what standard of performance?
In what sequence or combination with what other tasks?
What are the cues for initiation of the task?
What are the hazards and environmental and safety constraints on task performance?
What is the criticality of performance? That is, how essential is correct and complete task performance to overall mission success?
What is the task delay tolerance? That is, what is an acceptable interval between cues and the need to perform the task? For example: bleeding should usually be stopped immediately, but paperwork can wait. Is it possible to bring in other people to perform delayed tasks?
How difficult is the task, or what is the probability of inadequate performance?
How difficult is the task to learn?
Does the task have to be performed correctly by an individual upon first assignment to a job, or is there an opportunity for the task to be performed under supervision?
Are there long periods of non-performance when forgetting may occur?
What are the tools and aids for performing the task? Can job aids be developed which will simplify the tasks, reduce the requirement for training, or increase resistance to forgetting?

corporations, and other training or education-oriented consortia. These instructional systems development/systems approaches to training methods included task analysis as a critical phase in developing instruction and training. Montemerlo and Tennyson (1976) noted that from 1951 to 1976, there were over a hundred different ISD manuals written. Although differing in basic assumptions, emphases, or approaches in the design of training, most included some sort of task analysis. Andrews and Goodson (1980) reviewed sixty models of instructional design and found that 75 percent had an analysis phase that included some process and/or product recommendations concerning the analysis of tasks for which the instruction would be designed.

The trend in the U.S. Department of Defense toward extensive procedural documentation noted by Montemerlo and Tennyson (1976) has not abated. Each of the services has revised its ISD guidance documents several times in the last thirty years, and the Department of Defense itself has consolidated much of this guidance. As of this writing, the current task analysis methodological guidance is contained in *Performance Specification Training Data Products* (Department of Defense, 2001a) and is supported by the *Handbook for Instructional Systems Development/Systems Approach to Training and Education* (Department of Defense, 2001b). The appendix to this chapter lists some of the current guidance from each of the military services.

These later variants of instructional systems development/systems approach to training guidance provide clarification (if not much improvement) over the earlier methodologies. Although they are often updated to address more modern interactive media, training analysis methods are relatively constant, since training media decisions are typically made after the task analysis is completed. In general, the major activities involved in performing behavioral task analysis are the same as those that began after World War II: Observe performance; try to describe it in words; unpack the description hierarchically into sub-procedures; continue the process until some assumed elemental level of description is reached; identify conditional antecedents and measurable outcomes for each element; and finally consolidate commonalities across the hierarchy.

Cognitive Task Analysis

Cognitive task analysis is rooted in cognitive psychology, which investigates the mental processes involved in activities such as perceiving, remembering, thinking, and problem solving. Between 1950 and 1970, cognitive psychology emerged because of a growing dissatisfaction with behavioristic accounts for such complex activities, and the rapid growth of systems engineering and information theory. Cognitive psychology views the human as an information processor and emphasizes higher-order mental processes as critical components of skilled behavior (see Neisser [1967] for a review of the early stages of information processing). The result is a characterization of human tasks as involving various processes such as perception, pattern recognition, intention, memory storage, knowledge retrieval, mental computation, reasoning, and choice as well as overt action. This led to a new view of task analysis, and a new set of task analysis techniques for identifying cognitive components of task performance.

Relationship Between Behavioral and Cognitive Task Analysis

Behavioral task analysis is primarily concerned with the observable tasks that operators perform successfully to accomplish a particular job as part of a specific system (Kirwan & Ainsworth, 1992). It focuses on what should or must be done to accomplish work (Vincente, 1999) and its primary focus is on identifying specific input–task–output sequences that, if correctly performed, allow the individuals to achieve specific goals. The typical output of a behavioral task analysis is an ordered listing of tasks, subtasks, inputs, activities, outputs, environmental conditions, and performance standards that are heavily dependent upon the specific components of the system (Benyon, 1992).

In contrast, cognitive task analysis attempts to describe or analyze the mental phenomena that are thought to engender specific behaviors. The focus is on the mental representations, underlying knowledge structures, and information processing activities necessary to make decisions and perform actions. Cognitive task analysis helps contextualize behavior for those aspects of the job that are ambiguous, difficult, or involve multiple inputs. It expands traditional or behavioral task analysis by capturing ''information about the knowledge, thought processes, and goal structures that underlie observable task performance'' (Schraagen, Chipman, & Shalin, 2000, p. 3). It helps to understand and contextualize behavior for work that is complex, ill-defined, or difficult by describing how experts use their knowledge to structure relatively complex ill-defined work and accomplish that work effectively and efficiently.

However, there are significant differences between traditional or behavioral task analysis and cognitive task analysis. Early task analytic methods such as those of Miller (1953) and Flanagan (1954) placed their emphasis on the behaviors or actions that the worker must perform as part of a human-machine system. This view described both the human and the machine portions of the system as subsystems that receive inputs, perform internal operations on those inputs, and provide outputs. Even though these pioneers of task analysis recognized that cognitive processes were inherent in performing tasks, they included these cognitive processes within an overarching system framework and focused on describing the required output or performance as a function of the input or stimuli.

As Stanton (2006) points out, beginning in the 1960s, Annett and colleagues expanded on the importance of cognition in task analysis and provided a direct link to cognitive psychology by emphasizing the importance of goals and subgoals in their development of hierarchical task analysis. They recognized that analyses focused primarily on simple observable behaviors were unable to capture the dynamic nonlinear nature of what people did as work increasingly shifted from hands-on manufacturing to process and supervisory control. Hierarchical task analysis proposed that work consists of hierarchically organized clusters of goals and that workers perform tasks to meet specific goals within a particular goal hierarchy. During this same time, a similar view of the importance of hierarchical organized goals and their importance in human-machine systems was also evolving within the field of control theory (for example, Kelly, 1968).

Technology enables smart machines to perform many highly predictable, procedural tasks, leaving the worker to cope with cognitive tasks requiring inferences, judgment, diagnosis, and decision making (Howell & Cooke, 1989). As a result, it becomes more difficult, if not impossible, to specify procedures ''for every possible situation, especially in a world filled with unexpected events'' (Norman, 1988, p. 156).

In practice, there is no hard-and-fast line separating behavioral and cognitive task analysis. Modern behavioral analysis includes cognitive tasks because the scheme for making sense of, or inferences from, the trace of behavior is really a characterization of the cognition underlying task performance. Conversely, cognitive task analysis must start with an observable purpose, mission, and overt performance, and involves observing and analyzing verbal and nonverbal behavior. Contemporary cognitive science is doing very fine-grained task tracing of cognitive events and constructing tasks that elucidate what alternative rules or problem-solution paths people are using while neurocognitive techniques such as functional Magnetic Resonance Imaging (fMRI) offer the promise of making the neural elements of cognition observable (National Research Council, 2008).

Contemporary task analyses are eclectic, involving both behavioral and cognitive techniques. Behavioral task analysis focuses on the identifiable behavioral activity that an operator must perform. Most practitioners recognize that monitoring, detecting, recognizing, and deciding are essential components of any task analysis. Therefore, all successful task performance involves at least some cognitive components ''in the sense that perception, decision, knowledge, and judgment are required'' (Welford, 1968, p. 21).

Job-Task-Cognition Continuum

There is a continuum of techniques for analyzing jobs and job performance. Task analysts must be familiar with both behavioral and cognitive task analysis in order to understand and describe what is required to accomplish complex work successfully.

If one needs to determine primary job tasks and their characteristics (frequency, criticality, difficulty, conditions under which they are performed, time required to complete, and so forth), then one conducts a job analysis. If one is interested in creating a hierarchical list of job performances from the task to the operant level of performance (discrete steps), one uses a behavioral task analysis. If one is dealing with observable performances that are the result of complex cognitive processes involving interpretation, troubleshooting, decision making, and other forms of problem solving, then one uses a cognitive task analysis to elicit the knowledge, analyze it, and represent it in ways that enable the closure of performance gaps. Analysts frequently use a combination of behavioral and cognitive methods and balance their relative investment in each method based on the nature of work and the final goals of the analysis (Gordon, 1994).

BEHAVIORAL TASK ANALYSIS PROCESS

This section describes the typical stages involved in conducting a behavioral task analysis with several examples of specific applications of behavioral task analysis as well as a number of limitations and pitfalls.

Behavioral task analysis is the process used to identify critical tasks and identify the standards, conditions, performance measures, and other criteria associated with the performance of those tasks. While this section emphasizes the use of behavioral task analysis to support training, the basic principles underlying behavioral task analysis are applicable to the broad spectrum of human system integration activities. It is a critical part of the human systems integration process and is used throughout the life cycle of the system to help allocate functions between humans and machines, identify necessary staffing levels, design human–machine/human–computer interfaces, and assess human performance as well as to develop training and job performance aids.

A training task analysis is typically conducted in an iterative fashion and involves mission analysis, job analysis, and task identification, as well as behavioral and cognitive task analyses. The results of the behavioral task analysis serve as the basis for the development of a training program. In courses that tie the content directly to preparing students for the performance of a mission or job, the analyst documents the performance requirements and develops a task list for the mission/job that may include higher-level tasks such as problem solving, leadership, and management. The analyst then hierarchically decomposes the required performance by looking at the mission, job, or the task itself and cataloging its parts. A result of this phase is the identification of the knowledge, skills, and abilities, aptitudes, or attitudes (KSAs) required for the mission/job/task performance. Then the analyst compares the KSAs the actual jobholder must possess with the KSAs already possessed by the incoming students. The difference between what the students already know and can perform and what the mission/job/task requires them to know and be able to perform defines a training requirement.

Mission Analysis. The goal of the mission analysis is to identify all the major tasks and functions necessary for accomplishment of the overall organizational mission. All instruction should be based directly on mission, job, or education requirements. Mission/job analysis uses data from many sources, including mission statements found in regulations or locally developed statements.

For the military, the Universal Joint Task List (Chairman of the Joint Chiefs of Staff, 2008) provides broad task action descriptions, specifies the conditions affecting performance of the tasks, and provides measures and criteria for performance that comprise the task standards. Each Service and Defense Agency also has developed Service or Agency Mission Essential Task Lists, in some cases down to the level of individual job/task analyses, which also include specific conditions and standards. Each task description typically consists of an action verb, an object of the action, and qualifiers that provide additional detail concerning conduct of the action, conditions, and/or standards.

For nonmilitary organizations, mission, job, and task analyses are often conducted by human resources departments, trade/labor organizations, or governmental labor/commerce departments. Analysts and curriculum developers also make use of management engineering reports, occupational data, and direct observation to determine the actual mission and job requirements. The products of many military analyses are also applicable to nonmilitary jobs; for example, the Air Force has provided a number of task lists for jobs such as Airfield Management and Maintenance Data Systems Analysis to VTECS for use in their state workforce development and assessment programs.

Job-Task Analysis. Job-task analysis is typically done for purposes of job description for personnel functions such as determining hiring qualifications, allocating tasks to various levels of responsibility, specifying promotion paths, etc. Many examples of job-task analysis can be found on the Internet. See, for example, the website of the Oregon Department of Public Safety Standards and Training (DPSST), the website of the *Journal of Nuclear Medicine Technology* (Prekeges, 2003), and the Texas Commission on Law Enforcement Officer Standards and Education (1997). (Website addresses are in the References section.) In the military, analyses have been done (and redone) for most jobs. These are found in job descriptions such as the Air Force Occupational Standards (AFOS), Army and Marine Corps Military Occupational Standards (MOS), and the Navy Enlisted Occupational Standards (NEOCS).

As part of this process, *Instructional Systems Development/Systems Approach to Training and Education* (Department of Defense, 2001b) recommends development of a mission matrix, denoting who (what team or individual) is responsible for each task. Typically, this is done in a hierarchy of more subordinate tasks at each lower level of command. Higher-level tasks are often ''collective'' tasks, when they require more than one individual to complete, and lower levels in the hierarchy identify discrete parts of collective tasks assignable as standalone tasks to individuals. Analysts identify tasks within a system and mission context to provide the purpose of the task and the goals that the operator is trying to achieve in performing that task. They then decompose tasks to subtasks based on a number of factors such as the purpose of the analysis, the complexity of the task, and the consequences of poor task performance.

Figure 6.2 illustrates the general structure of collective and individual tasks associated with most complex systems. As mentioned earlier, a system frequently has a number of missions, and those missions usually require a combination of collective and individual tasks.

The point of the collective task analysis is to (1) identify the conditions, standards, and actions for work-group-level tasks so that they can be assigned to the appropriate level of command or supervision and (2) separate individual

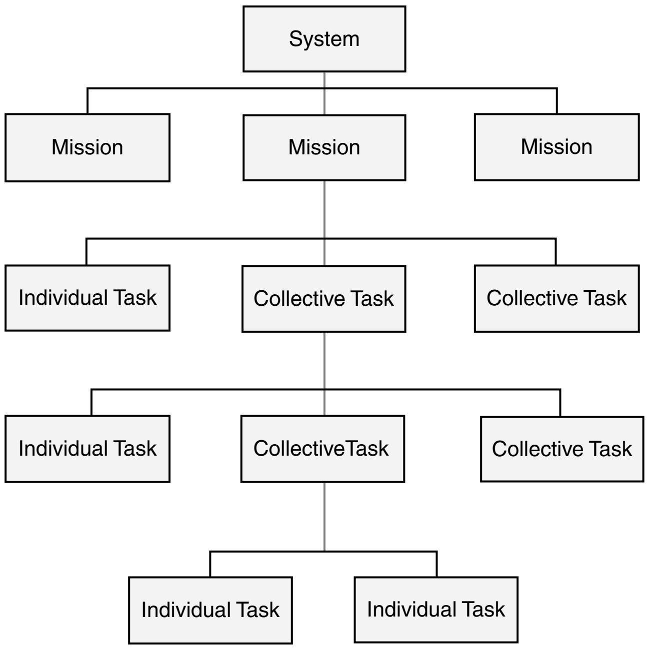


Figure 6.2 Collective and Individual Tasks Are Part of Most Complex Systems.
Adapted from TRADOC, 1999

tasks so they can be further analyzed for purposes of doctrine development, equipment design, or training.

A behavioral task analysis involves defining and describing the tasks that individuals must (learn to) perform. Although this chapter focuses on describing how task analysis is used to identify training requirements, the basic concepts are relevant to the analysis of any human-centered system. For example, a related set of task-based methods has come from the trend toward ''user-centered'' system design, where the activities a user (for example, a computer user) tries to accomplish are analyzed to design better user interfaces and processes (Osga, 2003).

BEHAVIORAL TASK ANALYSIS METHODS

Unfortunately, behavioral task analysis is not an exact science. It involves the application of both accepted procedures and analytic judgment to describe how people perform their work. As the work being analyzed becomes more complex, the analyst's experience and theoretical orientation play increasingly important roles in determining the final product. However, most behavioral task analyses share a number of common features and techniques that are described below.

When performing a behavioral task analysis, the analyst typically performs a number of activities that include:

- Reviewing system and mission documentation;
- Interviewing subject-matter experts;
- Observing people performing tasks;
- Recording information;
- Organizing information and observations; and
- Documenting and validating the analysis.

Except for very simple analyses, these activities are usually performed in an iterative manner. The goal is to develop an in-depth understanding of what is done and how it is done. Typically, one begins by reviewing documents that describe the system and its mission. These documents may include mission need statements, business plans, system specifications, scenario descriptions, use cases, and technical drawings. These documents provide critical information about the displays, controls, interfaces, and task conditions. They also allow the analyst to identify potential questions to ask subject-matter experts and highlight areas of special concern to watch as people perform their tasks.

The next step typically involves structured interviews with subject-matter experts, either individually or in small groups. These knowledgeable individuals know how to perform the task or similar tasks. There are numerous ways to conduct these interviews. One approach is to simply have the subject-matter expert list the tasks in sequence as they are performed within a mission context. Another approach is to have the subject-matter expert provide a highly detailed verbal protocol. In addition to describing the actual behaviors that must be performed, it is also desirable to have the subject-matter expert describe the stimuli that cue the start and completion of the task as well as those stimuli that allow her to monitor successful task progress. There are obviously a number of variations on these approaches. The analyst may rely solely on verbal description or combine verbal descriptions with actual task performance using simulated or actual equipment. The analyst may also obtain detailed recordings of what the subject-matter experts actually did and compare those to what the subject-matter experts verbally described. Selection of a specific approach depends upon the complexity of the mission and the associated behavioral tasks, the purpose of the analysis, and the experience of the analyst.

Whichever approach the analyst chooses, selecting subject-matter experts is often a challenge. It is difficult to define expertise and it is often difficult to gain access to the true experts in a field. Some subject-matter experts also have difficulty generalizing to new systems that are under development and frequently will remain focused on describing tasks as they are performed using the current system.

If possible, observing typical operators performing the task as individuals or crewmembers is extremely valuable. If the task is a collective task, team members should perform their normal roles while carrying out the task. This allows identifying discrepancies between the expert's view of the task and the way the task is routinely performed by ''average'' operators. It also allows identifying where there is the potential for performance problems based on workload, equipment layout, or underlying skills.

After tasks to be trained are identified, a more detailed analysis of each task is performed. Task analysis is the process of breaking a task down to identify the:

- Component steps of a task;
- Sequence of those steps;
- Conditions under which the task will be performed (for example, at night, in the field);
- Task cues; and
- Standards of performance that must be achieved, expressed in terms of accuracy completeness, sequence, or speed.

There are many other methods that have been developed over the past fifty years. Carlisle (1986, 1989) gives many observational and analytic methods. Table 6.4 lists methods and brief descriptions of techniques given by Carlisle and others.

Obviously, this entire process must be documented and analyzed. It is not at all uncommon for a complex system to involve hundreds of tasks. Therefore, the analyst needs to have some systematic means of identifying the task. There

Table 6.4 Behavioral Task Analysis Methods and Descriptions

Method	*Description*
Interview Analysis	Interview job performers to elicit task descriptions. (Carlisle, p. 24)
Card-Sort Analysis	Have job performers sort cards containing actions and objects. (Carlisle, p. 28)
Task-Matrix Analysis	List objects in left column, provide actions across top row. Each cell is an action/object pair. (Carlisle, p. 32)
List-Expansion Analysis	Decompose actions into sub-steps. (Carlisle, p. 36)
Daily Log Analysis	Job performers keep a log of everything they do. (Carlisle, p. 39)

(*Continued*)

Table 6.4 (*Continued*)

Method	*Description*
Walk and Talk Analysis	The analyst ''shadows'' the job performer and elicits narrations about what is being done. (Carlisle, p. 42)
Job Function Analysis	Provides standardized categories that can be used to identify and organize specific tasks. Analyst and master performer review possible job functions to determine task statements. (Carlisle, p. 45)
Performance Deficiency Analysis	Analyst prepares ''what is'' and ''what should be'' chart for job performance based on deficiencies in process and/or product. (authors)
Risk Assessment	Done after the task inventory is compiled to determine the importance and difficulty of each task. From this assessment the analyst can target tasks for further analysis and training. (Carlisle, p. 50)
Performance Probe Analysis	Assess the information, resource, and motivation requirements of the job and the worker in order to suggest needed improvements. (Carlisle, p. 128)
Ergonomic Analysis	Define the cognitive and physical ''fit'' of the person to the job tasks. (Carlisle; authors)
Problem Analysis	Use description and analysis to determine the underlying reasons for faulty performance. (Carlisle, p. 123)
Job Satisfaction Analysis	Determine how meaningful the job is to the workers. The job can be redesigned based on the analysis, to make it more satisfying. (Carlisle, p. 133)
Paradigm Analysis	Dividing the entire job into component parts. (authors)
Process Charting	Recording and categorizing the steps in a task. (Carlisle, p. 71)
Flow Charting	Shows the sequential actions and decisions in a complex process. It reduces complexity by showing a likely set of actions and simple decisions. (Carlisle, p. 84)
Operation Charting	Used to record, categorize, and improve the detailed motions and senses involved in skilled jobs. (Carlisle, p. 77)
Decision Technique	The decision technique is used when a task is essentially non-sequential or when various decisions must be made, based on the symptoms of a particular situation, in order to select the correct procedure. This

	technique is ideal for troubleshooting and diagnostic tasks. (Carlisle, p. 96)
Stimulus-Response Charting	Used to describe task steps in great detail. It is important when tasks are very complex, involving numerous people, data inputs, or decisions. (Carlisle, p. 65)
Picture Technique	Used when a drawing or photograph of the task makes it easier to analyze and understand how the task is done. (Carlisle, p. 92)
Critical Incident Technique	Used to identify the critical job requirements that are the difference between doing the job correctly and doing it incorrectly. The worker's actual performance is reported, compared, and classified as effective or ineffective. (Carlisle, p. 119)
Learning Hierarchy Technique	Used to order and sequence tasks according to logical relationships. This ensures a correct learning sequence. (Carlisle, p. 178)
Operator Function Modeling	Modeling tasks in complex and dynamic systems. Result is a network diagram showing how human operators accomplish simultaneous activities. (authors)
Operational Sequence Diagramming	Provides a graphical method of task analysis aimed at ''describ[ing] clearly the functions of the system integrating all potential hardware requirements.'' (Walley & Shepherd, 1992, p. 18)
Time Line Analysis	Time line analysis identifies how much time a task will take in order to determine if the task can be completed within the available time. (authors)
Algorithmic Analysis	Analyzes the procedure used to perform the task as if it were a computer program. (authors)
Equipment Analysis	Determine what equipment needs to have done to it for maintenance or fault prevention. (authors)
Interface Analysis	Systematically explore all possible inputs/actions to a computerized task interface. (authors)
Design Analysis	Analyze the design of a system to determine its operational requirement. (authors)
Design Interview	Interview designers of a system to elicit their description of intended operation. (authors)

Definitions adapted from Kenneth E. Carlisle (1986) with page numbers. In some cases, where noted, the definitions were developed by other authors, including the authors of this chapter.

are a number of templates available for recording the information collected during a task analysis. Table 6.5 illustrates one such template. In addition to capturing a written description, additional products such as sequence diagrams, flow charts, and task hierarchies are typically produced as part of the analysis.

Table 6.5 Routine Automotive Maintenance Task

Task Number	*Task*
1	Perform routine automotive maintenance
1.1	Maintain fluid levels
1.1.1	Maintain transmission fluid level
1.1.1.1	Check level
1.1.1.1.1	Start engine and run until normal operating temperature is reached
1.1.1.1.2	Locate transmission dipstick
1.1.1.1.2.1	Visually search engine compartment
1.1.1.1.2.2	If not located, consult maintenance manual
1.1.1.1.2.2.1	Locate engine compartment diagram in manual
1.1.1.1.2.2.1.1	Search list of figures for ''engine compartment''
1.1.1.1.2.2.1.2	If not located, search contents for ''maintenance''
1.1.1.1.2.3	Search diagram for label ''transmission dipstick''
1.1.1.1.2.4	Match diagram to actual engine compartment
1.1.1.1.3	Remove dipstick
1.1.1.1.4	Remove fluid from dipstick by wiping with rag
1.1.1.1.5	Re-insert dipstick until it is inserted as far as possible
1.1.1.1.6	Remove dipstick
1.1.1.1.7	Check level of fluid against dipstick gradations
1.1.1.1.8	Wipe dipstick, reinsert
1.1.1.2	Add fluid as necessary
1.1.1.3	Recheck level
1.1.2	Maintain oil level
1.1.3	Maintain coolant level
1.1.4	Maintain brake fluid level
1.2	Maintain proper tire pressure
etc.	

DESCRIBING BEHAVIORAL TASKS

Consider a vehicle mechanic whose job includes performing ''routine automotive maintenance.'' The mechanic's goal is to maintain the vehicle in accordance with published standards. In order to achieve that goal, the mechanic must perform a number of tasks and associated subtasks. Table 6.5 shows a hierarchical breakout of some of those tasks and subtasks. It also illustrates a typical numbering convention for hierarchical task decomposition.

Two high-level tasks associated with routine automotive maintenance are maintaining proper fluid levels and tire pressure. Table 6.6 lists these high-level tasks as Tasks 1.1 and 1.2, respectively. Each task is recorded using an action verb (explaining what the behavior is), an object (to which the action is applied or performed), and qualifiers or additional information clarifying the intent of the task. Good task statements are clear, complete, and concise. Often, the most difficult part of this process is finding the right action verb that best indicates the behavior involved. Many task analysis guidebooks provide lists of action verbs to help precisely specify behavior. Table 13 of Mil-HDBK-29612-2A (Department of Defense, 2001b) provides a list of action verbs as well as typical learning objectives. Jonassen, Tessmer, and Hannum (1999) illustrate the use of syntactic analysis of these verbs, objects, or qualifiers to identify clusters of tasks that share common characteristics in order to structure a training curriculum.

Table 6.5 also shows how each higher-level task can be decomposed into separable subtasks. For example, maintaining proper levels of transmission fluid (1.1.1), oil (1.1.2), coolant (1.1.3), and brake fluid (1.1.4) are primary subtasks associated with the higher-order task of maintaining fluid levels (1.1).

These subtasks can be broken down into procedural steps such as checking the level of transmission fluid (1.1.1.1) and the specific actions required to check the level of transmission fluid such as running the engine (1.1.1.1.1) and locating the transmission dipstick (1.1.1.1.2). If necessary, tasks such as locating the transmission dipstick can be further decomposed into more detailed procedural steps that include alternative tasks, for example, what to do if the mechanic cannot locate the dipstick (1.1.1.1.2.2).

After the tasks and procedural steps are identified, the following are specified:

- Conditions under which the task will be performed. In this case, the task would be performed in a typical automotive service facility;
- Task cues. In this case, cues might include assignment of the task by the service manager, observation of a fluid leak, or complaint by vehicle owner; or

- Standard of performance that must be achieved. In this case, the accuracy standard would be that the fluid is filled to the correct level. A speed standard, such as ''within ten minutes,'' might also be used.

Pitfalls and Problems

Over-Analysis. In the example above, there are already eight levels of analysis. As Table 6.5 shows, tasks can be decomposed into increasingly smaller subtasks much like a Russian matryoshka nesting or stacking doll until the analyst reaches a ''logical stopping point.'' Obviously, this process can continue to ridiculously fine levels of detail, depending on the requirements of the task and the assumed preparation level of the trainees. One rule for stopping the decomposition process is to continue until actions are reached which the trainee can perform without specific instruction. Annett and Duncan (1967) proposed what is known as the ''P $\times$ C rule'' as a means of defining this stopping point. According to this rule, each task is evaluated based on the probability of inadequate performance and the cost to the system of inadequate performance. If the resulting product is unacceptable, the task is decomposed into subtasks and the P $\times$ C rule is applied to each subtask. This process continues until the probability of inadequate performance or the consequences of inadequate performance are acceptable. While this rule suggests a rigorous criterion, it should be remembered that in most cases both the probability of inadequate performance and its consequences are based on expert opinion and rarely involve precise measures. Although the P $\times$ C rule is not as precise as one might hope, it does provide the analyst with a good rule of thumb—if a task is highly unlikely to be performed incorrectly, then a more detailed analysis of that task will not significantly increase our understanding of what the operator must do to successfully accomplish that portion of the job.

Paradoxically, though, continued finer-grained analysis can overcomplicate the whole process: subordinate steps often seem more complex than higher-level tasks and they also become more ''cognitive.'' For example, ''checking fluid level with a dipstick'' is quite a simple task, while locating information in a technical manual can be much more difficult, and visual search, when analyzed in terms of eye movements, is surprisingly intense.

Completeness and Accuracy. In the example above, a good maintainer will also inspect the fluid for color (evidence of oxidation) and contamination in Step 1.1.1.1.7. How does the analyst know that a critical step has been left out? Similarly, how does the analyst know that the process or steps of procedure are actually correct?

Both completeness and accuracy can present serious problems when task analysis is performed by so-called human-performance specialists who are not content specialists, and this is true for both behavioral and cognitive task analytic approaches. The standard approach to this problem is to use subject-matter experts to assist with and review the analysis. However, this often just shifts the problem because the putative expert is often a person who can be spared from actual operations, rather than the most capable, and there is generally no independent way to verify the expertise of the putative expert. A better way is for a verified content expert to do the analysis since it is often easier to teach a content expert task analysis than to teach an education specialist highly technical job content. It is important that content experts be not only expert performers, but that they also understand the theory and/or science underlying the tasks and the systems that are implemented in the real world. For example, in analyzing tasks involved in operation of a radar, the content expert should understand the underlying physics of electromagnetic wave propagation and reflection, the reflective properties of the targets the radar is designed to detect, the real-world design and implementation of radar systems (because there are always design compromises from theoretical optimality), as well as the operational implications from physical and design constraints. In the end, accuracy and completeness depend on the perspicacity of the analyst, and anyone who has ever done a complete task analysis usually ends up knowing more about the tasks than most job experts know.

For maintenance/repair tasks, another approach to ensuring completeness is to analyze the equipment and its design. What does the equipment need to have done to it in order to maintain or repair it? These tasks should NOT be analyzed by asking performers what they do, unless there is some independent way to verify that they actually understand what they are doing. Rather, maintenance and repair tasks are best identified and analyzed by examining the design and implementation of the device to be maintained, and the quality statistics that are usually accumulated over time by competent organizations. For example, the majority of maintenance requirements of a pump are governed by the materials and design of the pump, especially the pump seals and bearings, and the pump's operational history. Therefore, the best course is to consult the designers and manufacturers of the pump to identify maintenance requirements, to confirm maintenance histories comply, and verify that there is no unexplained flaw that has skewed the data. Probably the best implementation of this method comes from the U.S. Navy's nuclear submarine program. The training analysis and design methods are documented in NAVEDTRA 131, Personnel Performance Profile Based Curriculum Development Manual.

A similar approach also can work for operator tasks. The analyst should first consult the tactical or operational requirements that a particular system was designed to meet, then consult the designers, who had operations in mind as they designed the system, to identify what operator controls are built into the system to support the operational tasks that were envisioned during design. Again, it may be risky to consult current job performers as ''experts'' unless there is independent verification of their understanding of the underlying theory, system concept, and system design, as well as operational employment. Sometimes ''expert'' misconceptions have led to poor task analyses. For example, a senior Army air defense radar operator told trainees not to use a control to correct for refractivity of the atmosphere because doing so would ''bend the radar beam'' (Larson, personal communication, 1995). (Actually, the control aimed the antenna at a slightly different elevation in anticipation of atmospheric refraction.)

Oversimplification. The emphasis on observable steps often leads to omission of decision making, reflection, deduction, integration, and other so-called ''mental'' or ''cognitive'' tasks. This can often lead to detailed specification of trivia and neglect of the ''hard parts'' of the task. For example, an analysis for the task ''Write a Great American Novel'' might be:

1. Obtain an American English dictionary.
2. Choose words from the dictionary.
3. Arrange words in proper order.
4. Repeat 2 and 3 until novel is complete (Note: Words may be used more than once).

There are several ways to handle such situations. One is to use the methods of cognitive task analysis, described by Villachica and Stone (Chapter Seven). Another is to analyze these tasks more schematically and procedurally, by specifying the behaviors involved at finer-grained levels of analysis. For example, novel writing may involve several higher-level but nonetheless behavioral subtasks:

1. Develop overall plot outline.
2. List major and minor characters.
3. For each major character, write actions consistent with plot outline that give insight into character motives.
   - 3.1. For villain, write scene involving premeditation of crime.
   - 3.2. Write description of earlier life events leading to antisocial outcome.
   - 3.3. Write scene showing gratification with nefarious result.

While this approach may lead to formulaic writing, it at least expresses the specifics of what needs to be written.

Commonalities. Behavioral task analysis may ignore the connections among related tasks. This can lead to instruction, especially for introductory material, that is a series of isolated topics. After several top-level tasks are completed, the analyst will likely notice commonalities across different task hierarchies. For instance, if the analysis in the automotive example above were pursued, there might be several different task-subtask decompositions that call for consulting the maintenance manual. When this occurs, the standard practice is to designate such common tasks as ''KSAs'' (knowledge, skills, and abilities) as in ''knowledge of maintenance manual'' or ''skill in locating information in a maintenance manual'' or ''ability to locate and interpret graphical and figural information in a technical publication.'' In general, this means that these supposedly more basic KSAs will not be further analyzed, and instead become enabling prerequisites for the to-be-designed training. Alternatively, similarities in the analysis can all be grouped together, analyzed once, and later taught as a common prerequisite to the otherwise-unrelated higher-level tasks.

Over-Emphasis on Procedural Skill. In the examples above, the task breakdowns resulted in specifying procedural steps in greater and greater behavioral detail. This is relatively easy when the job consists of highly proceduralized steps that involve observable behaviors. However, many jobs also require cognitive tasks such as problem solving, and the analyst often needs to identify behavioral objectives for these tasks as part of these tasks. For tasks such as problem solving, there are other ways to decide what behavioral objectives should be included in the training program. These include various forms of algorithmic analysis, and the use of model- or theory-based characterizations of tasks to provide a basis for determining what subtasks should be included in the analysis.

For example, problem-solving tasks can often be grouped by similarity of solution methods, for example, Hively, Patterson, and Page (1968). Here, the idea is to specify solution methods or processes in advance (by looking at rules or algorithms for solving problems), then use these as a basis for understanding the behavior. This approach was first described for mathematical or arithmetic tasks (Polya, 1957) and then extended to the diagnosis of incorrect performance during training (Brown, Burton, & Larkin, 1977; Scandura, 1983). Another approach is to use mathematical or qualitative models that represent parts of tasks (de Kleer & Brown, 1983; Forbus, 1981), and then use these as a basis for identifying what knowledge is needed to execute the task even though this makes the analysis more ''cognitive.''

BEHAVIORAL TASK ANALYSIS APPLICATION EXAMPLES

Interactive Multisensor Analysis Training. Wulfeck, Wetzel-Smith, and Dickieson(2002) provide an example of a task analysis and development of training objectives drawn from sonar training. The model-based scientific visualizations in the Interactive Multisensor Analysis Training (IMAT) project have also enabled a new approach to the specification of training tasks for acoustic, electromagnetic, and electro-optical systems.

As seen earlier, the traditional method for analyzing a task is to identify the components of a task by hierarchically decomposing it into subtasks, skills, and knowledge. Training is then based on these units, and they are tested mostly individually. This can often result in a focus on low-level detail in training, so that tasks are independent and serial, with limited cause and effect explanation as to how those topics interrelate. This approach often leads to instruction containing descriptions of complex phenomena and large amounts of factual data with little contextual reference. Feltovich, Spiro, and Coulsen (1991) point out that teaching isolated topics or ''compartmentalizing knowledge'' makes it more difficult for students to integrate their knowledge or to generalize knowledge in new applications.

Further, when task analysis results in introductory instruction for complex interrelated tasks as a series of isolated topics, there may be a detrimental effect on future learning because oversimplification early in training may result in later difficulty due to the need to unlearn the too simple explanations and replace them with more mature knowledge.

The Interactive Multisensor Analysis Training project has led to a process for conducting conceptual analyses, which involves the following general steps:

a. Determine the most complex performance problem for which a training solution is required.
b. Identify and refine the variables, and dimensions along which they vary, necessary to model the problem.
c. Obtain or develop mathematical and/or qualitative-process models that relate these variables/dimensions and specify how they interact.
d. Design an interface and display system that facilitates understanding of the variables and their relationships.
e. Identify problem scenarios (cases) using the resulting simulation.
f. Validate the problem scenarios by working through them with operators and tacticians.

In general, the process of constructing and validating model-based visualization systems identifies the underlying critical variables, their relationships, and their tactical implications. These then become the enabling concepts and tasks in the analysis. This analytic methodology has been applied to acoustic, electromagnetic, and electro-optical systems and has successfully identified training requirements for developmental systems still in test and evaluation.

Mission Essential Competencies^{SM}. The Mission Essential Competencies^{SM} work of the Air Force Research Laboratory represents a new approach to capturing job performance requirements (Alliger, Beard, Bennett, Colegrove, &

Garrity, 2007). Mission Essential Competencies^{SM} link knowledge, skills, and individual experiences in order to understand the performance requirements associated with a specific job. This approach combines elements of behavioral and cognitive task analyses in order to identify performance requirements at different levels of abstraction. At the highest level is a *Mission Essential Competency*, which describes a higher-order individual, team, or inter-team competency needed for successful mission completion. An example of a mission essential competency for an F-15 pilot is to *intercept and target enemy aircraft*. At the next level of abstraction are the *Supporting Competencies*, the generic competencies that enable completion of one or more mission essential competencies such as being able to *clearly, concisely, and correctly communicate information*. The lowest level of abstraction consists of the *specific knowledge* (information or facts) and *skills* (compiled sequence of actions) that are associated with a competency such as *knowing the rule of engagement*. Once this abstraction process is completed, the process also identifies the specific *experiences* that are important for learning, refining, or sustaining those competencies.

Mission Essential Competencies^{SM} provide a hierarchical scheme that captures the high-level competencies needed for a particular job and then systematically decomposes those competencies into the specific knowledge and skills that underlie those competencies. They focus on the competencies needed to accomplish a particular mission and are developed in facilitated workshops with subject-matter experts. Competency analysis have been conducted across a wide variety of missions and are currently being used to help identify options and requirements for training environments, training devices, and training frequency.

Driving. Behavioral task analysis is also valuable in fields such as computer science and robotics, as they attempt to develop autonomous and/or intelligent systems. Task-analytic methods help inform the development process by describing the functions such systems must perform and the range of conditions under which that performance occurs.

An example is the use of a behavioral task analysis to support development of more capable autonomous vehicles (National Institute of Standards and Technology, 2003). The program used an analysis of human driving behavior (McKnight & Adams, 1970) to help them develop a hierarchical taxonomy of driving tasks, identify stimulus events, and estimate complexity as part of creating the computer algorithms and data structures necessary to develop an autonomous vehicle. The McKnight and Adams task analysis was done to support development of driver education objectives and provided a detailed description of forty-five passenger car driver tasks and fifteen hundred driver behaviors. These tasks were broken down into two major categories: on- and off-road tasks. The on-road tasks and subtasks were classified as basic control tasks, general driving tasks, and situation specific tasks. Supporting material for these tasks included estimates of performance limitations, criticality, and underlying skill (perceptual, motor, or cognitive). Table 6.6 shows the major on- and off-road tasks identified by McKnight and Adams and provides a few examples of the subtasks and behaviors associated with these major tasks.

Table 6.6 Examples of Tasks and Subtasks

Task Categories	*Task Category*	*Subtasks*	*Goal*	*Actions*
On-Road Tasks				
	Basic Control Tasks			
		Steering		
		Skid Control		
			Takes Preventive Measures to Avoid Skids	
				Enters curves or turns at moderate speeds
				Attempts to avoid panic stops or hard braking if possible
			Attempts to Arrest Skid	
				Keeps foot off brake
	General Driving Tasks			
		Surveillance		
		Navigation		
		Urban Driving		

Task Categories	*Task Category*	*Subtasks*	*Goal*	*Actions*
	Tasks Related to Traffic Conditions			
		Passing		
		Lane Changing		
	Tasks Related to Roadway Characteristics			
		Lane Usage		
		Weather Conditions		
	Tasks Related to the Car			
		Hauling and Towing Loads		
		Pushing and Towing		
Off-Road Behaviors				
	Pre-Trip Planning			
		Planning		
		Loading		
	Maintenance Tasks			
		Routine Care and Servicing		
	Legal Responsibilities			

Aircraft Maintenance. In the late 1990s, Northwestern University conducted an aviation maintenance technician job/task analysis for the Federal Aviation Administration (Adam, Czepiel, Henry, Krulee, Murray, & Williamson, 1997). The goal of this analysis was to obtain data to update the core curriculum requirements for obtaining an Aviation Maintenance Technician Certificate.

One interesting aspect of this analysis was its magnitude. This analysis illustrates some of the complexities in analyzing the tasks associated with a complex job across an entire industry. Unlike most job task analyses, which focus on a particular job within a particular organization, this analysis obtained data on over three hundred tasks from 2,434 surveys administered to respondents at eighty-four different aviation facilities, ranging from major airlines to small general aviation shops.

The overall objectives of this analysis were to:

- Identify tasks that broadly define the job of an aviation maintenance technician;
- Survey a representative sample of aviation facilities to determine
  - Task relevance/importance,
  - Tasks that reflect technology change, and
  - Similarities and differences between different segments of the industry; and
- Facilitate revisions to aviation maintenance technician school curricula.

One of the challenges in conducting such a broad analysis is determining the number of tasks to include in the survey. The analysts had to find an appropriate balance between an exhaustive listing of all possible aviation maintenance technician tasks, which would be too long for a survey, and a shorter list that would sacrifice performance details. To achieve this balance, the analysts focused on three major task categories:

- Check, Test, Service, Inspect
- Repair, Replace, Modify, Calibrate
- Troubleshoot

These three major categories were then grouped into twenty Air Transportation Association subject categories such as landing gear, flight controls, or engines. Respondents rated each task along three dimensions:

- Frequency of task performance—less than once a quarter, quarterly, monthly, weekly, daily;
- Criticality to flight operations—negligible, low, average, high, extremely high; and
- Difficulty to learn—not difficult, somewhat difficult, moderately difficult, increasingly difficult, very difficult.

The surveys contained brief descriptions such as ''not critical to the continuation of flight'' or ''task is complex and involves multiple steps'' to provide common anchor points for the respondents.

In their discussion of the results of this task analysis, Adam and his associates (1997) highlight the differences between different industry segments in how frequently an aviation maintenance technician performs a particular task. Base maintenance facilities, such as those operated by large airlines, typically involve a high degree of specialization. As a result, technicians typically perform a narrower range of tasks than those working in a general aviation facility. Because of this specialization, technicians at major base facilities are likely to report performing a narrower range of tasks or performing some tasks less frequently than technicians at less specialized facilities. Task analysts need to be alert to such differences in developing their data collection protocols. Selecting an inappropriate segment of the industry or experience level can result in data that misrepresents how frequently certain tasks are performed. This frequency effect could also inadvertently influence the perceived occurrence or difficulty of those tasks.

TECHNIQUES AND TOOLS FOR BEHAVIORAL TASK ANALYSIS

Performing and documenting a task analysis to support system design, job/organizational analysis, or training is often a difficult and time-consuming process. Not only must the analysts identify goals, determine the appropriate level of decomposition, and describe the associated tasks and actions, but they must also document this information so that it can be used to enable effective human performance. As the system becomes increasingly complex, it becomes harder to grasp the interrelationships between various levels of decomposition and to provide the necessary documentation.

Analysts have traditionally relied on paper and pencil as the primary media for recording the results of this work. Even when software tools are used, their primary purpose has been to facilitate data entry, as opposed to assisting the analysis in the actual conduct of a behavioral task analysis. Once the data are captured, analysts must analyze, synthesize, format, and present the results. The synthesized results are typically presented in text format accompanied by either graphical or tabular material. The widespread availability of personal computers and relatively inexpensive graphical (for example, Visio™) and spreadsheet (for example, Excel™) software has greatly reduced the work involved in keeping the data organized and formatting it for presentation.

Computer Aids for Task Analysis

Comprehensive Occupational Data Analysis Program (CODAP). The Comprehensive Occupational Data Analysis Program represents an empirical approach to occupational analysis developed during the 1980s. The underlying assumption of CODAP is that jobs must be defined in terms of the tasks performed by the workers. Using task statements and background information, CODAP sought to provide a tool kit of computer programs, analysis guidelines, and a theory-based approach to job and occupational analysis. The goal was to provide a common foundation that would help organizations perform human resources management functions such as recruitment, selection, classification, training, and job design. Background material on CODAP is available from several online sources (http://www.codap.com/faq.htm# what; http://www.icodap.org/; and http://www.metricanet.com/groups/codap/index.html).

O*NET. The U.S. Department of Labor has developed an extensive job analysis database of occupational requirements, tasks, and job performer skills and abilities, called O*NET. It is the nation's primary source of occupational information (Department of Labor Employment and Training Administration, 2008). Publicly available online access to the O*NET database allows users to explore occupations, tasks, and knowledge and skill requirements at http://online.onetcenter.org. It includes the ability to relate occupations to other job classification systems such as those in the military. O*NET should be used at the start of any behavioral task analysis.

The Authoring Instructional Materials (AIM) System. The AIM system is a set of computer-based tools for curriculum design and instructional materials preparation. It was originally proposed as a developmental project in the late 1970s, and early versions were implemented and fielded through the 1980s (Wulfeck, Dickieson, Apple, & Vogt, 1992). In general, the idea was to use computer interviews to conduct a dialog with subject-matter specialists to identify and analyze training tasks, then to organize them and their subordinate and superordinate relatives in a relational database. This way, the links among tasks, subtasks, learning objectives, instruction, and technical documentation could be maintained much more efficiently than by traditional methods. AIM development continues to the present time, and AIM versions currently support the Navy's approach to instructional systems development documented in NAVEDTRA 130/131, and the Navy's Integrated Learning Environment. Over 300,000 hours of formal training courseware have been supported by the AIM system. Current AIM information is available at https://ile-help.nko.navy.mil/ile/content/supportapps/aim.aspx.

Automated Systems Approach to Training (ASAT). The U.S. Army's ASAT (Automated Systems Approach to Training) is a software application that is used for Army training and development, support, and management functions. It operates as a training information system, a tool for decision making, and a training development product production system. It has modules that support both collective and individual task development, and then collective and individual training publications, lesson plans, and other documentation. The system is described at www.asat.army.mil.

IMPRINT. The U.S. Army Research Laboratory has developed an Improved Performance Research Integration Tool (IMPRINT), a discrete event simulation tool for analyzing human performance in system design and acquisition. IMPRINT's website states, ''Task-level information is used to construct networks representing the flow and the performance time and accuracy for operational and maintenance missions. IMPRINT is used to model both crew and individual soldier performance. For some analyses, workload profiles are generated so that crew-workload distribution and soldier-system task allocation can be examined. In other cases, maintainer workload is assessed along with the resulting system availability. Also, using embedded algorithms, IMPRINT models the effects of personnel characteristics, training frequency, and environmental stressors on the overall system performance. Manpower requirements estimates can be generated for a single system, a unit, or Army-wide. IMPRINT outputs can be used as the basis for estimating manpower lifecycle costs.'' It is described at: http://www.arl.army.mil/ARL-Directorates/HRED/imb/imprint/Imprint7.htm.

IMPRINT uses a discrete event simulation program called ''Micro Saint Sharp'' from Macro Analysis & Design (MAAD). MAAD also has developed the Integrated Performance Modeling Environment (IPME), a simulation environment for examining human performance in complex task situations. It contains a collection of tools for describing, simulating, and analyzing operator tasks. IPME is described at: http://www.maad.com/index.pl/ipme.

Multimedia Video Task Analysis™. Multimedia Video Task Analysis™ (MVTA™) was developed by Professor Robert G. Radwin and Dr. Thomas Y. Yen in the Ergonomics Analysis and Design Consortium at the University of Wisconsin-Madison to help automate time studies of observable behaviors (see http://mvta.engr.wisc.edu/). MVTA allows interactive study of activities recorded on a computer-based video system.

TaskArchitect™. TaskArchitect is a computer-based tool for task analysis for complex system design or to create documentation or training materials (see http://www.taskarchitect.com/products.html). TaskArchitect provides graphical and textual tools that support entering and describing tasks and the hierarchical relationships among them. TaskArchitect captures the relationships between tasks and can redraw the analysis automatically after every edit. Task tables and task diagrams are linked together to allow display of either format. It supports both the creation of indented lists of tasks and task diagrams as well as the dynamic reordering of tasks and their relationships. Parent and sibling selection, cut and paste, drag and drop, and task references are supported in all analysis diagrams in order to allow the user to reshape and duplicate areas of the analysis quickly and easily.

The company claims that TaskArchitect allows on-the-fly interaction with subject-matter experts to very quickly produce finished analyses. The system also provides for export of data to other analysis tools like IPME, or to graphical tools like Visio™.

Mindmapping and Concept Mapping. The idea of organizing relationships among concepts, or words, or object-definitions, or familial relationships, or taxons into graphical (usually two-dimensional) ''maps'' is many centuries old. More modern conceptions of such ''linked node'' relationships were formalized by Collins and Quillian (1969) originally as a cognitive theory of memory, called ''Semantic Networks.'' Since then much work in the fields of artificial intelligence and cognitive science has explored and developed these ideas. In addition, these techniques have been popularized and in some cases commercialized as so-called ''concept maps'' or Mind Maps™. Today, computerized tools are available to aid in the construction of linked-node diagrams. While these are most often used for cognitive analysis, they can be useful for depicting hierarchical behavioral task relationships as well. An extensive list of such tools is given in Wikipedia articles at http://en.wikipedia.org/wiki/List_of_Mind_Mapping_software and http://en.wikipedia.org/wiki/List_of_concept_mapping_software.

SUMMARY AND CONCLUSIONS

Criticisms and Limitations of Behavioral Task Analysis

Despite its widespread use for a variety of functions, behavioral task analysis is not without its weaknesses and critics. Some of those weaknesses and criticisms are briefly addressed here.

A major challenge is defining the task and then determining the level of analysis required for a particular application. There is usually pressure on the analyst to finish the analysis as quickly as possible so that the rest of the instructional design process can proceed. In those cases, the analysts may begin the analysis without giving proper analytical consideration to the job or function or sub-function. That approach can lead to products that lack the depth and rigor that will be required later on in the instructional design process. The feeling might be, ''We already know the top-level information like the job and function, so let's not waste time at those higher levels. Let's go right to the 'meat' of the analysis of the tasks.'' Such an approach is suboptimal and the analyst will not have the proper context to define tasks for the entire job. The main danger is that there will be a complete low-level analysis, and good training, for the wrong tasks and functions.

Another criticism of behavioral task analysis is its heavy reliance on the use of subject-matter experts. It is not often possible to do a complete task analysis based on observations and interviews in the field alone. In many cases, analysts must rely on subject-matter experts to give them a detailed understanding of what tasks must be performed and why. Subject-matter experts can be an excellent source of task information, but it may be difficult to obtain the required number of subject-matter experts because they are usually in high demand performing the job. In an effort to achieve reliability of information, analysts usually should interview at least three to five subject-matter experts. In addition, there is a criticism that the information obtained from subject-matter experts may be somewhat biased because they have learned ''shortcuts'' through the years in doing the tasks that require extraordinary knowledge or ability. In those cases, it would not be appropriate to teach the shortcuts to inexperienced trainees because they do not have that extra knowledge or skill yet. Another problem is that purported subject-matter experts may not really be expert; while sufficiently experienced, they may have little in-depth technical understanding of the specific task. Finally, a practical difficulty in working with some subject-matter experts is whether they are able to articulate what is required to perform a task. They may be experts in their field, but that does not necessarily mean they have the communication skills necessary to explain what tasks must be performed, or when, or how, or why. Alternatively, they may be quite inexpert, but have good persuasion skills.

Are these criticisms of behavioral task analysis discussed above justified? Although there is merit to these critiques, these problems are relatively minor when compared to the benefit that behavioral task analysis brings to the performance improvement process. The problems cited with using subject-matter experts can be largely mitigated as long as the analyst anticipates the difficulties. For example, the analyst can explain to managers and colleagues on the performance improvement team that the analysis should not be rushed or curtailed simply to meet a timeline. Another consideration is that an analyst might ask for twice as many subject-matter experts as really needed just to make sure they have the desired number of experts. The analyst can continually remind the subject-matter experts that they need to stick to the formal method for performing the tasks and not implement shortcuts because novices will be the primary users of the resulting instruction or job performance aid.

Future of Behavioral Task Analysis

As long as jobs change, new systems are developed, or there is a desire to improve job performance, there will be a continuing need for quality behavioral task analysis. And, as such, task analysis is a dynamic rather than static field that continues to evolve in response to both the demands of the workplace and the increasing understanding of human performance.

Currently, behavioral task analysis provides the tools necessary to characterize well-defined procedural work. It will become easier to perform such analysis in the future as information technology automates the routine, mechanical aspects of documenting tasks and describing their relationships.

The challenge for future task analytic methods will be the continuing development of hybrid procedures that include both behavioral and cognitive components. The need to link behavior and cognition in order to provide a unified description of the work that people perform is critical as we continue to move from routine procedural work to work that is performed within increasingly complex socio-technological systems. These systems involve numerous individuals, teams, and technologies that respond dynamically to their changing environment. As a result, human performance professionals need to develop greater understanding of how the relationship between behavioral, cognitive, social, and technological factors shapes the behavioral demands for the next generation of work.

References

Adam, L. K., Czepiel, E. J., Henry, D. J., Krulee, G. K., Murray, G. C., & Williamson, B. M. (1997, March). *Job task analysis for the aviation maintenance technician—Phase II final report*. Evanston, IL: The Transportation Center, Northwestern University.

Alliger, G. M., Beard, R., Bennett, W., Jr., Colegrove, C. M., & Garrity, M. (2007). *Understanding mission essential competencies as a work analysis method* (AFRL-HE-AZ-TR-2007-0034). Mesa, AZ: Air Force Research Laboratory, Human Effectiveness Directorate, Warfighter Readiness Research Division.

Andrews, D. H., & Goodson, L. A. (1980). A comparative analysis of models of instructional design. *Journal of Instructional Development, 3*(4), 2–15.

Annett, J. (2004). Hierarchical task analysis. In D. Diaper & N. A. Stanton (Eds.), *The handbook of task analysis for human-computer interaction* (pp. 67–82). Mahwah, NJ: Lawrence Erlbaum Associates.

Annett, J., & Duncan, K. D. (1967). Task analysis and training design. *Occupational Psychology*, *41*, 211–221.

Annett, J., & Stanton, N. A. (Eds.). (2000). *Task analysis*. London: Taylor & Francis.

Benyon, D. (1992). The role of task analysis in systems design. *Interacting with Computers*, *4*(1), 102–123.

Bower, G. H., & Hilgard, E. R. (1981). *Theories of learning*. New York: Appleton-Century-Crofts.

Branson, R. K., Rayner, G. T., Cox, J. G., Furman, J. P., King, F. J., & Hannum, W. H. (1975). *Interservice procedures for instructional systems development* (NAVEDTRA 106A, 5 Vols.). Pensacola, FL: Naval Education and Training Command.

Brown, J. S., Burton, R. R., & Larkin, K. M. (1977). *Representing and using procedural bugs for educational purposes*. Paper presented at the Association for Computing Machinery 1977 Annual Conference. Retrieved April 9, 2008, from http://portal.acm.org/citation.cfm?id = 810211.

Carlisle, K. E. (1986). *Analyzing jobs and tasks*. Englewood Cliffs, NJ: Educational Technology Publications.

Carlisle, K. E. (1989, March 27–28). Analyzing jobs and tasks. Workshop presented at the Annual Conference of the National Society for Performance and Instruction.

Casner, S. M., Encinas, C. M., & Puentes, A. (2004). *Computer and broadband technology in aircraft line maintenance: A task analysis and questionnaire*. Sunnyvale, CA: NASA Ames Research Center. Retrieved September 16, 2008, from www.hf.faa.gov/docs/508/docs/maintFY04Broadrpt.pdf.

Chairman of the Joint Chiefs of Staff Manual. (2008, August 25). *Universal joint task manual* (CJCSM 3500.4E). Washington, DC: Author.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of verbal learning and verbal behavior*, *8*(2), 240–248.

de Kleer, J., & Brown, J. S. (1983). Assumptions and ambiguities in mechanistic mental models. In D. Gentner & A. Stevens (Eds.), *Mental models* (pp. 155–190). Hillsdale, NJ: Lawrence Erlbaum Associates.

Department of Defense (1999, August 16). *Department of Defense: Definitions of human factors terms* (MIL-HDBK-1908B). Redstone Arsenal, AL: U.S. Army Aviation and Missile Command.

Department of Defense (2001a, August 31). *Performance specification: Training data products* (MIL-PRF-29612B). Lakehurst, NJ: Naval Air Warfare Center Aircraft Division.

Department of Defense (2001b, August 31). *Instructional systems development/systems approach to training and education* (MIL-HDBK-29612-2A). Lakehurst, NJ: Naval Air Warfare Center Aircraft Division.

Department of Labor Employment and Training Administration. (2008). O*NET Resource Center. Retrieved December 3, 2008, from http://www.onetcenter.org/.

Federal Aviation Administration. (n.d.). FAA Human Factors Workbench. Retrieved January 3, 2008, from http://www2.hf.faa.gov/workbenchtools/default.aspx?rPage = subCatDetails&CatID = 7.

Feltovich, P. J., Spiro, R. J., & Coulson, R. L. (1991). *Learning, teaching, and testing for complex conceptual understanding*. (ADA248728). Springfield, IL: Southern Illinois University School of Medicine. Also in N. Frederiksen, R. J. Mislevy, & I. I. Bejar (Eds.), *Test theory for a new generation of tests* (pp. 181–217). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ferguson, D. (2000). *Therbligs: The keys to simplifying work*. Retrieved on April 1, 2008, from http://gilbrethnetwork.tripod.com/therbligs.html.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, *51*(4), 327–358.

Forbus, K. D. (1981). Qualitative reasoning about physical processes. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence* (pp. 326–330). Vancouver, BC, Canada

Gilbert, T. F. (1962). Mathetics: The technology of education. *Journal of Mathetics*, *1*, 7–73.

Gordon, S. E. (1994). *Systematic training program design: Maximizing effectiveness and minimizing liability*. Englewood Cliffs, NJ: Prentice-Hall.

Hackos, J. T., & Redish, J. C. (1998). *User and task analysis for interface design*. Hoboken, NJ: John Wiley & Sons.

Hively, W. H., Patterson, H. L., & Page, S. (1968). A ''universe-defined'' system of arithmetic achievement tests. *Journal of Educational Measurement*, *5*, 275–290.

Howell, W. C., & Cooke, N. J. (1989). Training the human information processor: A review of cognitive models. In I. Goldstein (Ed.), *Training and development in work organizations: Frontiers of industrial and organizational psychology* (pp. 121–182). San Francisco: Jossey-Bass.

Jonassen, D. H., Hannum, W. H., & Tessmer, M. (1989). *Handbook of task analysis procedures*. New York: Praeger.

Jonassen, D. H., Tessmer, M., & Hannum, W. H. (1999). *Task analysis methods for instructional design*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kelly, C. R. (1968). *Manual and automatic control*. Hoboken, NJ: John Wiley & Sons.

Kirwan, B., & Ainsworth, L. K. (1992). *A guide to task analysis*. London: Taylor & Francis.

Koppes, L. L. (Ed.). (2006). *Historical perspectives in industrial and organizational psychology*. Mahwah, NJ: Lawrence Erlbaum Associates.

Larson, O. (1995). Personal communication.

Mager, R. F. (1988). *Making instruction work, or, skillbloomers*. Belmont, CA: Lake.

McKnight, A. J., & Adams, B. B. (1970). *Driver education task analysis: Volume 1. Task descriptions final report* (Report No. DOT-HS-800-367; HumRRO-70-103). Washington, DC: National Highway Traffic Safety Administration.

Meister, D. (1985). *Behavioral analysis & measurement methods*. Hoboken, NJ: John Wiley & Sons.

Meister, D. (1989). *Conceptual aspects of human factors*. Baltimore: The Johns Hopkins University Press.

Miller, R. B. (1953). *A method for man-machine task analysis*. (WADC Technical Report 53–137, AD015921). Wright-Patterson AFB, OH: Wright Air Development Center.

Miller, R. B. (1962). Task description and analysis. In R. M. Gagné (Ed.), *Psychological principles in system development* (pp. 187–228). New York: Holt.

Montemerlo, M. D., & Tennyson, M. E. (1975). *Instructional systems development: Conceptual analysis and comprehensive bibliography* (NAVTRAEQUIPCEN Tech. Rep. IH-257). Orlando, FL: Naval Training Equipment Center.

National Institute of Standards and Technology (2003). *Driver task analysis*. Retrieved September 16, 2008, from http://www.isd.mel.nist.gov/projects/darpa_mars/Task_Analysis.html.

National Research Council. (2008). *Emerging cognitive neurosciences and related technologies*. Washington, DC: National Academies Press.

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.

Nemeth, C. P. (2004). *Human factors methods for design: Making systems human-centered*. New York: CRC Press.

Norman, D. A. (1988). *The design of everyday things*. New York: Basic Books.

Osga, G. A. (2003). *Human-centered shipboard systems and operations*. In H. R. Booher (Ed.), *Handbook of human systems integration* (pp. 743–793). Hoboken, NJ: John Wiley & Sons.

Polya, G. (1957). *How to solve it: A new aspect of mathematical method* (2nd ed.). Princeton, NJ: Princeton University Press.

Prekeges, J. L. (2003). Nuclear Medicine Technology Certification Board (NMTCB) 2003 Task Analysis Report. *Journal of Nuclear Medicine Technology*, *31*(2), 86–91. Retrieved March 25, 2008, from http://tech.snmjournals.org/cgi/content/abstract/31/2/86.

Scandura, J. M. (1983). Instructional strategies based on the Structural Learning theory. In C. M. Reigeluth (Ed.), *Instructional design theories and models* (pp. 213–248). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schraagen, J. M., Chipman, S. F., & Shalin, V. L. (2000). Introduction to cognitive task analysis. In J. M. Schraagen, S. F. Chipman, & V. L. Shalin (Eds.), *Cognitive task analysis* (pp. 3–23). Mahwah, NJ: Lawrence Erlbaum Associates.

Stanton, N. A. (2006). Hierarchical task analysis: Developments, applications, and extensions. *Applied Ergonomics*, *37*, 55–79.

Task Architect Inc. (2008). http://www.taskarchitect.com/products.html.

Texas Commission on Law Enforcement Officer Standards and Education. (1997). *Texas peace officer job task analysis report*. Retrieved April 1, 2008, from http://www.tcleose.state.tx.us/commreports/Statewide%20Job%20Analysis%20for%20Texas%20Peace%20Officers.htm.

TRADOC (1999, March 9). *Systems approach to training management, process, and products* (TRADOC Regulation 350-70). Ft. Monroe, VA: TRADOC. Available at http://www.tradoc.army.mil/tpubs/regs/r350-70/index.html.

Vincente, K. J. (1999). *Cognitive work analysis: Toward safe, productive, and healthy computer-based work*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Vocational-Technical Education Consortium of States. (2000). *A consortium for innovative career and workforce development resources*. Retrieved on April 1, 2008, from http://www.v-tecs.org/.

Walley, S. P., & Shepherd, A. (1992). The task analysis process. In B. Kirwan & L. K. Ainsworth (Eds.), *A guide to task analysis* (pp. 15–34). London: Taylor & Francis, Inc.

Welford, A. T. (1968). *Fundamentals of skill*. London: Methuen & Co Ltd.

Wulfeck, W. H., Dickieson, J. L., Apple, J., & Vogt, J. L. (1992). The automation of curriculum development using the Authoring Instructional Materials (AIM) system. *Instructional Science*, *21*(4), 255–267.

Wulfeck, W. H., Wetzel-Smith, S. K., & Dickieson, J. L. (2004). Interactive multi-sensor analysis training. *Proceedings of the NATO RTO Human Factors and Medicine Panel (HFM) Symposium on Advanced Technologies for Military Training*, October 2003, Genoa, Italy. Neuilly-sur-Seine Cedex, France: North Atlantic Treaty Organization Research and Technology Agency. Retrieved October 8, 2008, from ftp://ftp.rta.nato.int//PubFullText/RTO/MP/RTO-MP-HFM-101/MP-HFM-101-04.pdf.

APPENDIX A: U.S. GOVERNMENT DOCUMENTS DETAILING PROCEDURES FOR BEHAVIORAL TASK ANALYSIS

As an instructional systems development benchmark, the 1975 Interservice Training Review Organization (ITRO) Instructional Systems Development methodology (Branson, Rayner, Cox, Furman, King, & Hannum, 1975) (described in the rescinded NAVEDTRA-106A and TRADOC Pamphlet 350-30) provided the initial framework for the Joint Service process model. The interservice procedures were amended around 1980 (e.g., NAVEDTRA 110A), and later were replaced in the mid-1980s with a Military Standard for Instructional Systems Development (MIL STD 1379D) supported by MIL HDBK 1379 (four volumes) and MIL HDBK 292 (two volumes), which itself was replaced in the 1990s with a ''Performance Specification for Training Data Products'' (MIL PRF 29612), supported by a new Department of Defense Handbook for Instructional Systems Development/Systems Approach to Training and Education (MIL-HDBK-29612-2A). (This is Volume 2 of a five-volume series on military training.) The others are MIL-HDBK-29612-1, Department of Defense Handbook, Guidance for Acquisition of Training Data Products and Services, which contains guidance to be used by all services for the preparation of solicitations and evaluation of solicitation responses for training. MIL-HDBK-29612-3, Department of Defense Handbook, Development of Interactive Multimedia Instruction (IMI), which contains guidance on the application of the multimedia training courseware development process. MIL-HDBK- 29612-4, Department of Defense Handbook, Glossary for Training, which contains a listing of training terms and definitions. MIL-HDBK-29612-5, Department of Defense Handbook, Advanced Distributed Learning (ADL) Products and Systems describes methods and procedures for developing distance- and distributed-learning services and curricula.

The current Military Handbook MIL-HDBK-29612-2A contains an extended discussion of behavioral task analysis methods.

Each of the United States Armed Forces and the Department of Defense has its own amplifying information and guidance. These are listed in Table 6.7.

Table 6.7 Department of Defense Guidelines

Department of Defense	
DoDISS	Department of Defense Index of Specifications and Standards
DI-SESS-81518B	Instructional Performance Requirements Document
CJCSM 3500.04	Universal Joint Task List
Department of the Army	
TRADOC Regulation 350-70	Training Development Management, Processes, and Products
TRADOC PAM 350-70-1	A Guide for Producing Collective Training Products
TRADOC PAM 350-70-2	Multimedia Courseware Development Guide
Department of the Navy	
NAVEDTRA 130	Task Based Curriculum Development Manual
NAVEDTRA 131	Personnel Performance Profile Based Curriculum Development Manual
NAVEDTRA 134	Navy Instructor Manual
NAVEDTRA 135	Navy School Management Manual
United States Marine Corps	Systems Approach to Training Manual
Department of the Air Force	
AFPD 33-22	Military Training
AFMAN 36-2234	Instructional Systems Development
AFH 36-2235	Information for Designers of Instructional Systems (in twelve volumes) Volume 1–Executive Summary Volume 2–ISD/SAT Automated Tools/What Works Volume 3–Application to Acquisition

(*Continued*)

Table 6.7 (*Continued*)

Department of the Air Force	
	Volume 4–Manager's Guide to New Education and Training Technologies Volume 5–Interactive Courseware (ICW) Design, Development and Management Guide Volume 6–Guide to Needs Assessment Volume 7–Design Guide for Device-Based Aircrew Training Volume 8–Application to Aircrew Training Volume 9–Application to Technical Training Volume 10–Application to Education Volume 11–Application to Unit Training Volume 12–Information for Designers of Instructional Systems
Coast Guard	
Coast Guard Commandant Instruction (COMDTINST) 1550.9	Management of the Coast Guard's Training System
COMDTINST M1414.8C	Enlisted Performance Qualifications Manual
Other Government Agencies	
Department of Energy	10 CFR 712 Human Reliability Program 10 CFR 1046 Physical Protection of Security Interests
Federal Railroad Administration, Department of Transportation	49 CFR 236 Rules, Standards, and Instructions Governing the Installation, Inspection, Maintenance, and Repair of Signal and Train Control Systems, Devices, and Appliances

CHAPTER SEVEN

Cognitive Task Analysis

Research and Experience

Steven W. Villachica
Deborah L. Stone

INTRODUCTION

Most experienced instructional designers and HPT practitioners are already familiar with task analysis. They know to conduct one after a needs assessment has identified one or more gaps in performance that are worth closing.

In instructional settings, task analysis specifies what should be taught to close performance gaps arising from skill/knowledge deficits. Jonassen, Tessmer, and Hannum (1999) note that instructional designers conduct procedural task analyses (also called a behavioral task analysis) to decompose jobs into smaller parts, ending with discrete steps. Figure 7.1 illustrates a typical hierarchy of job tasks resulting from a procedural task analysis. Procedural task analyses work best when the tasks themselves are observable. Klein and Associates Inc. (1997) note that this approach is flexible, general, and logical. However, it can also overlook the hidden components of a task, especially knowledge work involving decision making and problem solving. A procedural approach to task analysis may also oversimplify the cognitive activity required to perform a complex task. So what should instructional designers and HPT practitioners do when they need to conduct a task analysis and:

- There are few observable behaviors?
- The behaviors are characterized by verbs such as ''evaluate,'' ''determine,'' ''assess,'' ''judge,'' ''notice,'' ''interpret,'' ''prioritize,'' ''anticipate,'' ''design,'' or ''plan''?

Job

1. Responsibilities
   1.1. Major tasks
      1.1.1. Sub-tasks
         1.1.1.1. Steps

Figure 7.1 A Procedural (Behavioral) Task Analysis Decomposes a Job into Discrete Steps.

- Exemplar performance varies in unpredictable ways in response to novel situations?
- The skills and knowledge producing the behaviors are otherwise invisible?

Jonassen, Tessmer, and Hannum (1999) suggest that instructional designers who wish to create instruction to improve problem-solving performance should use a *cognitive task analysis* (CTA), which models the actions, knowledge, and thinking that people engage in when performing a task.

As depicted in Figure 7.2, CTA is the extension of traditional job and task analysis techniques to yield information about the knowledge, thought processes, and goal structures that underlie observable task performance (Jonassen, Tessmer, & Hannum, 1999, p. 3). CTA is a broad area consisting of tools and techniques for describing the knowledge and strategies required for task performance (Schraagen, Chipman, & Shalin, 2000). The output of CTA is a mental model that depicts the knowledge and strategies that guide problem-solving performance. Practitioners often represent mental models graphically in the form of concept maps, diagrams, and user interfaces for software systems. For the purposes of this chapter, mental models are the ''causal understanding people develop about how to make things happen'' (Klein & Militello, 2005, pp. 335–336). Stated simply, ''CTA investigates what people know and how they think'' (Crandall, Klein, & Hoffmann, 2006, p. 7). In doing so, CTA specifies the mental models that support novice and expert thinking (cognition) and performance, with goals of helping:

- Novices and others in the organization act more like experts by making their knowledge and thinking visible.

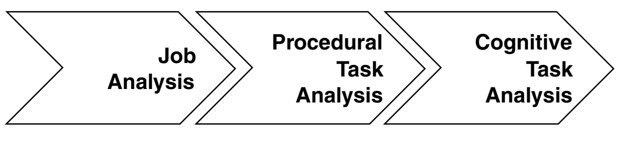


Figure 7.2 A Continuum of Job and Task Analytic Approaches. CTA Is a Natural Extension of Job and Procedural Task Analysis.

- Experts by making it easy for them to access knowledge, information, and tools they need to perform their jobs.

This chapter will introduce CTA by broadly describing the research and experience base supporting this practice. The chapter will begin by describing why practitioners may need to conduct a CTA and the different ways people use CTA to create a variety of performance improvement solutions. The chapter will then describe several approaches blended to form a hybrid CTA technique called Facilitated Case-Based Reasoning (F-CBR). Next, the chapter will describe how to conduct an F-CBR and its use in a recent effort. The chapter will conclude by summarizing the benefits of this CTA technique.

WHY CTA?

Unfortunately, experts cannot always tell instructional designers what they know or are thinking about. Experts literally do not know what makes them experts, and expertise does not lend itself to articulation. This situation is similar to trying to get recipes from good cooks. They may be willing to help, but what they articulate about how they do things is often untrustworthy.

The information experts articulate about their expertise is typically incomplete and inaccurate for several reasons. First, there is the problem of automaticity (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977), where frequent practice and feedback transform controlled cognitive processes that once required much of an individual's conscious attention into processes that can be performed effortlessly. Experts cannot describe what they do because it happens automatically, without conscious attention or thought. Second, experts have learned much of what they know implicitly. Berry and Dienes (1993) note that people can learn ''about the structure of a fairly complex stimulus environment, without necessarily intending to do so, and in such a way that the resulting knowledge is difficult to express'' (Berry & Dienes, 1993, p. 2). This is ''learning by mucking about.'' Children can learn how to play video games by experimentation. Broadbent (1977) found that people could learn to solve complex mathematical equations via trial and error. Furthermore, after reaching targeted levels of mastery, these individuals could not state how they did it. Automaticity and implicit learning conspire to render experts' articulations of their own expertise very unreliable.

Third, expertise is highly idiosyncratic. Every expert is a product of his or her unique experience, problem solving, and decision making. This is why, for every expert, there is often an equal and opposite expert. Instructional designers need ways to cut across multiple experts to derive consensus models of expertise that can be used in instruction, information systems, job tools, and other performance improvement solutions. CTA lets instructional designers see inside the heads of experts to look at the thinking, knowledge, and skills they cannot talk about. It also provides a way to cut across the idiosyncrasies of individual experts.

A CTA OVERVIEW

Just as there is no one way to conduct a procedural task analysis, there is no one way to conduct CTA. The www.ctaresource.com/ and http://mentalmodels.mitre.org/ websites list numerous approaches to conducting CTA. Practitioners use different approaches based on their experiences, the cause(s) of the performance gap they are trying to close, and project requirements. Table 7.1 describes different types of performance gaps and the uses of CTA.

Table 7.1 Performance Gaps and Corresponding Uses of CTA

Given a gap in performance arising from:	*HPT practitioners may opt to perform a CTA to:*
A lack of adequate instrumentation	Determine the requirements of a software system's graphical user interface. Kieras and Santoro (2004) used a CTA technique called GOMS (goals, operators, methods, and selection rules) to evaluate user interface concepts and team structure designs for a new class of military shipboard workstations. The new software automates some of the functions performed by humans and brings the information from multiple systems together to fewer operators. Benefits include fewer required crew members, better information integration, and improved decision making.
A lack of adequate instrumentation	Improve the usability of a software program. Ockerman and Mitchell (1999) employed three CTA techniques to improve the usability of a software tool called the Design Browser in creating a NASA satellite command management system (CMS). The researchers employed naturalistic decision making (Klein, Orasanu, Calderwood, & Zsambok, 1993), Klein's (1989) recognition-primed decision (RPD) model, and Kolodner's (1993) case-based reasoning. A formative evaluation of a software design team, users, and managers using this proof-of-concept prototype indicated the participants could extract useful information from the CMS Design Browser. Survey responses indicated that they perceived the CMS Design Browser as very useful.

A lack of accessible, contextualized information	Create a detailed interactive concept map depicting the organization of expert knowledge. Hoffman, Coffey, Ford, and Novak (2006) produced an online interactive concept map called STORM-LK for weather forecasting based on the results of a CTA using multiple techniques involving some twenty-two senior civilian forecasters. Using STORM-LK (see Figure 7.3), experienced and geographically distributed forecasters could build their own understanding of the dynamics of weather in a region with which they had little or no experience.
A lack of skill and knowledge	Build an online immersive learning simulation for Joint Forces Air Component Commanders (JFACCs), who must often juggle multiple roles while performing in high-stakes, rapidly evolving operational environments. Zimmerman, Burns, and Sestokas (2007) employed a CTA technique called the Knowledge Audit to elicit data describing expert/novice differences in contextualized decision making. Using this approach, they indentified eight critical tasks that the simulation would address and corresponding scenarios. JFACCs completing the simulation must establish a situational awareness of the scenario and then answer detailed questions about inherent potential risks and how to mitigate them. As JFACCs make decisions and take actions, simulated staff members provide feedback and input to decisions.
A lack of skill and knowledge	Create training to help novices acquire expertise in a shorter time. Gott and Lesgold (2000) created Sherlock, an online tutor based on the results of a CTA technique called PARI—Precursor, Action, Response, Interpretation (Hall, Gott, & Pokorny, 1990). This tutor taught Air Force F-15 technicians how to troubleshoot the airborne electronics involved in communication and navigation, along with the equipment used to test these systems. Novice avionics technicians with 2.75 years' experience using this online tutor outperformed master technicians possessing over ten years' experience.
Multiple performance gaps	Create a performance support system (PSS) comprised of integrated performance improvement solutions. Villachica and Stone (1998a) employed CTA to create a large-scale PSS called CornerStone, which reduced the time examiners required to audit NASDAQ firms by 20 percent, decreased administration costs by $2M, and produced an ROI of 229 percent with a five-year payoff.

Regardless of the specific technique, CTA efforts broadly consist of three phases:

1. Knowledge elicitation,
2. Data analysis, and
3. Knowledge representation.

Clark, Feldon, van Merriënböer, Yates, and Early (2007) note that CTA uses a variety of interview and observation strategies to represent the knowledge experts use to solve complex tasks. Olson and Biolsi (1991) provide a wider array of CTA options, contending that practitioners use either direct or indirect elicitation methods. Direct methods include structured interviews and focus groups, think-aloud protocols (Ericsson & Simon, 1993), observation, interruption analysis, commentary, drawing, card sorting, video debriefing, and concept mapping (Crandall, Klein, & Hoffman, 2006). Practitioners may also draw on indirect methods involving quantitative processes including multidimensional scaling, hierarchical cluster analysis, and the Pathfinder Scaling Algorithm (Schvaneveldt & Durso, 1981). Crandall, Klein, and Hoffman (2006) note

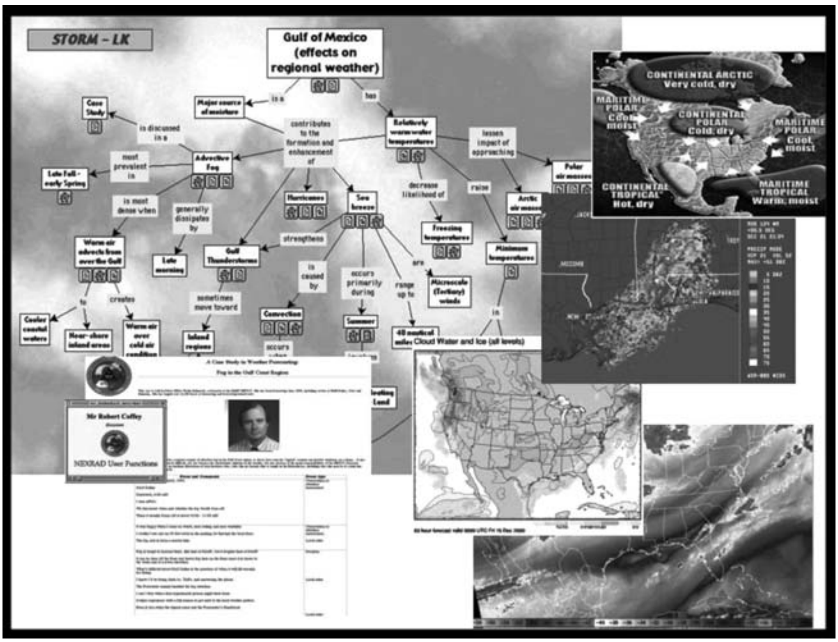


Figure 7.3 A Screen Shot Showing a Concept Map from STORM-LK Reflecting an Expert's Organization of a Local Weather System and Some Opened Resources (Radar, Satellite, Digital Video, Graphics).

Courtesy of R. R. Hoffman, Institute for Human and Machine Cognition that the professional literature on CTA contains far more information about knowledge elicitation than data analysis and representation.

Regardless of the specific technique and all other things being equal, instructional designers who appropriately use properly conducted CTA can improve instruction in problem solving in ways that produce measurable results. Velmahos and his associates (2004) reported the results of using a cognitive task analysis to teach resident interns to perform a central venous catheterization (CVC) procedure. Divided into two groups, one group learned the procedure using the traditional resident-to-resident ''see one, do one, teach one'' approach. The other group used a surgical skills laboratory and a checklist based on the results of a cognitive task analysis. Upon the completion of their instruction, both groups earned similar scores on a multiple-choice test. However, the CTA interns scored significantly higher in a repeat test. They also required fewer attempts to find a vein and required less time to complete the procedure. These results are not surprising. In a recent meta-analysis of seven studies, Lee (2004) noted that training interventions based on CTA results typically improve performance by 35 percent. This large overall effect size (+1.70) held across specific CTA techniques, workplace settings, and levels of worker experience.

THE CREATION OF FACILITATED CASE-BASED REASONING CTA

In projects to support knowledge workers' performance in decision making and problem solving, DLS Group, Inc. (DLS) has integrated two CTA techniques and a software development approach to create what we call ''Facilitated Case-Based Reasoning'' (F-CBR) (Stone, Clark, Foshay, & Villachica, 1999). DLS first employed this technique to create a large-scale performance support system for auditors who examine securities firms (Villachica & Stone, 1998a). DLS created F-CBR by merging the two CTA techniques and software development approach that follow:

- Precursors, Actions, Results, Interpretations (PARI) (Hall, Gott, & Pokorny, 1990),
- Case-Based Reasoning (Schank, 1990), and
- Rapid Application Development (Martin, 1991).

Precursors, Actions, Results, Interpretations (PARI)

To create the previously mentioned online tutor for avionics technicians, Gott and Lesgold (2000) used the results of a CTA technique called PARI. Created as part of the Basic Job Skills Research Program carried out at the U.S. Air Force's

Table 7.2 A PARI Structure Depicting a Recorded Action and Three Corresponding Probing Questions for the Action as Applied to an Avionics Troubleshooting Example

PARI Probe	*Avionics Troubleshooting Example*
Action: Recording of the discrete solution steps the expert took (used to ask subsequent probing questions about precursors, results, and interpretations).	''Remove the cable from J12 of the LRU ID resistor and ohm out the path through the LRU from pin 68 to pin 128.''
Precursor: What were the reasons for the actions you took (goals and intentions)?	''I want to see if the LRU ID resistor is good.''
Results: What were the results you obtained by completing the action?	''The reading is 1.55 ohms.''
Interpretation: What is your interpretation of the results you obtained?	''The problem isn't in the LRU, it's in the test station of the package.''

Examples are from Hall, Gott, and Pokorny, 1990, p. 25

Armstrong Laboratory, PARI was developed to detail the cognitive processes structures that underlie the performance of complex troubleshooting tasks. Over two hundred Air Force maintenance technicians whose primary job lay in troubleshooting participated in PARI CTA efforts.

The PARI technique employs authentic problem-solving sessions wherein two experts are given representative, real-world problem-solving tasks to troubleshoot. One technician is responsible for solving the problem. The other is responsible for recording the actions the problem solver took and then asking three additional probing questions for each action, as illustrated in Table 7.2. A PARI structure contains several related PARI sets required to solve a troubleshooting problem. An independent panel of experts reviews the PARI structures, which are subsequently compared to those of intermediate and novice performers. The PARI CTA team then reviews all precursors and actions to group them into subsequent cognitive skill categories for training. PARI's use of a contextualized setting to make knowledge and thinking processes visible ensures that subsequent training addresses both the observable and cognitive cues used in problem solving. Cues in a PARI-derived learning environment match those of the job environment, thereby maximizing the potential transfer of skills and knowledge.

Case-Based Reasoning (CBR)

Kolodner (1992) suggests that people reason using old experiences to understand and solve new problems they encounter. They remember a previous similar situation (a ''case'') and use it to solve the new problem. Jonassen, Tessmer, and Hannum (1999) suggest that these cases are similar to stories. Experts possess numerous stories, each of which is a case based on an experience solving a problem or interpreting a situation. Jonassen and Hernandez-Serrano (2002) contend that these stories are a primary medium for problem solving. Experts have indexed these stories in ways that facilitate their recall in similar situations. According to Jonassen, Tessmer, and Hannum (1999), such stories shed light on solving new problems by helping recall similar cases and their solutions. CBR is a mechanism for analyzing stories that embody such expertise. There are four major phases in this CTA technique (Jonassen, Tessmer, & Hannum, 1999, p. 150):

1. Collect a set of cases representative of domain problems.
2. Identify the points that each case can make (its ''lessons learned'').
3. Characterize the situations in which the case can make its points.
4. For each case, identify relevant cues (indexes) that would allow cases to be recalled in each situation. Jonassen and Hernandez-Serrano (2002) suggest the indexes shown in Table 7.3.

These authors also note that instructional designers can use the resulting cases three different ways in creating instruction. First, the cases can act as exemplars for teaching concepts, principles, or theories. Second, they can act as problem cases that students solve during their instruction. Third, they can act as a repository of advice and other best practices that students can use during their instruction and later as an on-the-job reference.

Rapid Application Development (RAD)

Realizing that the cost of software revisions increased during later phases of a development effort, Martin (1991) recognized that inaccurate user requirements and inappropriate design approaches were largely at fault. RAD consists of five different strategies that act as a set of checks and balances for speeding up development efforts, improving quality, and lowering costs. Table 7.4 describes these strategies and their corresponding benefits (Stone & Villachica, 2005). As the integration of these strategies within ISD efforts has been described elsewhere (see Villachica & Stone, 1998b), they will not be described here.

By hybridizing PARI, CBR, and RAD, DLS created a new CTA technique: F-CBR, which lets instructional designers and HPT practitioners:

- *Conduct a CTA at the level of the organization, rather than individual experts or expert dyads.* F-CBR enables the CTA team to produce consensus expert mental models applicable to entire organizations. These mental models are free of individual idiosyncrasies that cannot be leveraged over the entire organization. As opposed to CTA techniques

Table 7.3 Indexes

Problem-Situation-Topic Indexes	*Appropriate Solution Indexes*	*Appropriate Outcomes Indexes*
What were the goals-subgoals-intentions to be achieved in solving the problem or explaining the situation?	What solution was used?	Was the outcome fulfilled?
What constraints affected those goals?	What activities were involved in accomplishing the solution?	Were expectations violated?
Which features of the problem situation were most important, and what was the relationship between its parts?	What were the reasoning steps used to derive the solution?	Was the solution a success or failure?
What plans were developed for accomplishing the goal?	How did you justify the solution?	Can you explain why any failures occurred?
	What expectations did you have about results?	What repair strategies could have been used?
	What acceptable, alternative solutions were suggested but not chosen?	What could have been done to avoid the problem?
	What unacceptable, alternative solutions were not chosen?	

Source: Jonassen and Hernandez-Serrano, 2002, 13, p. 72

that employ less than a handful of experts at a time, F-CBR lets us conduct CTA in workshops that involve working with three to twelve experts at once and multiple workshops iteratively refining the mental model. Using multiple SMEs also enables us to resolve SME-related disputes as they arise in the collaborative analysis and design workshops. Over the course of multiple workshops, different groups of experts could validate the work of experts in the preceding workshop, thereby enabling F-CBR practitioners to obtain an adequate sample of organizational expertise for use in the CTA and the creation of subsequent performance improvement solutions.

Table 7.4 RAD Strategies, Descriptions, and Benefits

RAD Strategy	*Description*	*Benefit*
Skilled with Advanced Tools (SWAT) Teams	A small team of highly trained developers (3 or 4 people) who work closely together at high speed over multiple projects using a tool set selected to facilitate the fastest possible iterative development	Improved communications, empowerment, resource allocation, accountability, and productivity
Collaborative Analysis and Design	Stakeholders, end-users, and developers work collaboratively as a team in a series of structured, iterative workshops to specify performance requirements, identify content, and approve designs	Elicits consensus regarding performance requirements; resolves subject-matter expert disputes within the workshop; generates buy-in and a smoother implementation; helps avoid end-user surprises; aligned with principles of performance-based consulting
Prototyping	An iterative design approach that produces successive approximations of the end product	Illustrates approaches for meeting the performance requirements; provides accurate metrics for the development effort; sets up a finely tuned assembly line for the development effort; generates buy-in; minimizes changes later in the project when they are more expensive; delivers results quickly and inexpensively; eliminates the need for extensive design documents
Usability Tests	A ''pre-pilot'' test occurring toward the end of the design phase that employs a small, but representative, sample of end users to ensure that end users can intuitively and self-sufficiently use materials to meet business needs and objectives. There are no hidden points of confusion. Any training time is minimized.	• Ensures the accuracy and completeness of performance requirements. • Ensures user acceptance of the approaches to meet the performance requirements. • Identifies otherwise hidden errors or points of confusion.

(*Continued*)

Table 7.4 (*Continued*)

RAD Strategy	*Description*	*Benefit*
		• Minimizes the number of substantive changes from later pilot tests. • Improves overall quality. • Can be low cost, as dictated by the intervention, end users, and stakeholders. • Generates buy-in.
Timeboxing	An evergreen scheduling and prioritizing technique for meeting immovable deadlines while maximizing end-user input and buy-in. Within the timebox, features are negotiable. The schedule is not.	Controls feature, scope, and schedule creep. Facilitates rapid revision to completed materials—if the business and performance needs still exist. Focuses on client priorities.

- *Collect CTA information in a manner that lends itself to the development effort.* By asking SMEs to bring prototypical cases, including easy and difficult ones, to the collaborative analysis and design workshops, DLS was able to collect rich cases for both the CTA as well as subsequent case- and simulation-based training. As practitioners collect requirements and CTA-related information, they can quickly prototype potential training, information systems, software tools, and other performance improvement solutions. These prototypes further verify the results of the CTA and facilitate organizational buy-in.
- *Ensure that the expertise represented in the CTA is acceptable to the sponsoring organization.* As expertise is a function of individual experience, expertise tends to be idiosyncratic. Each expert has his or her own way of solving problems and making decisions. Sometimes these personal approaches stand at odds with the strategic direction of the organization. As workshops also include a client-side project manager, line managers, supervisors working under the direction of a steering committee, the organization can ensure that the tacit knowledge made visible is something the organization could truly leverage.

F-CBR Method

DLS created F-CBR as a scalable, flexible technique for CTA. It consists of three iterative phases, as depicted in Figure 7.4, and described below.

1. Set Up the CTA Workshop. The F-CBR technique begins by setting up a series of iterative CTA workshops. The number of workshops will vary, depending on the scope and complexity of the project and the extent to which expert cognition is known. Smaller, less complex efforts addressing areas where the CTA is merely filling in what is already known about expert cognition require fewer workshops than larger, complex projects where little is known about how experts interpret situations, make decisions, and solve problems.

To identify the participants for a workshop, begin by specifying required levels of expertise. For smaller efforts, we typically work with three or four experts and a novice. For larger efforts, we may meet with experts, journeyperson performers, and novices. In one of these larger workshops, we may have seven experts, two journeypersons, two novices, two supervisors, and a manager. Or we may choose to meet separately with groups of each. While the identification of novices and journeypersons is typically straightforward, practitioners want to ensure they pick experts carefully. We look for experts who:

- *Have paid their dues.* It takes time and dedication to become an expert. Ericsson and Charness (1994) note that it typically takes some 10 years of

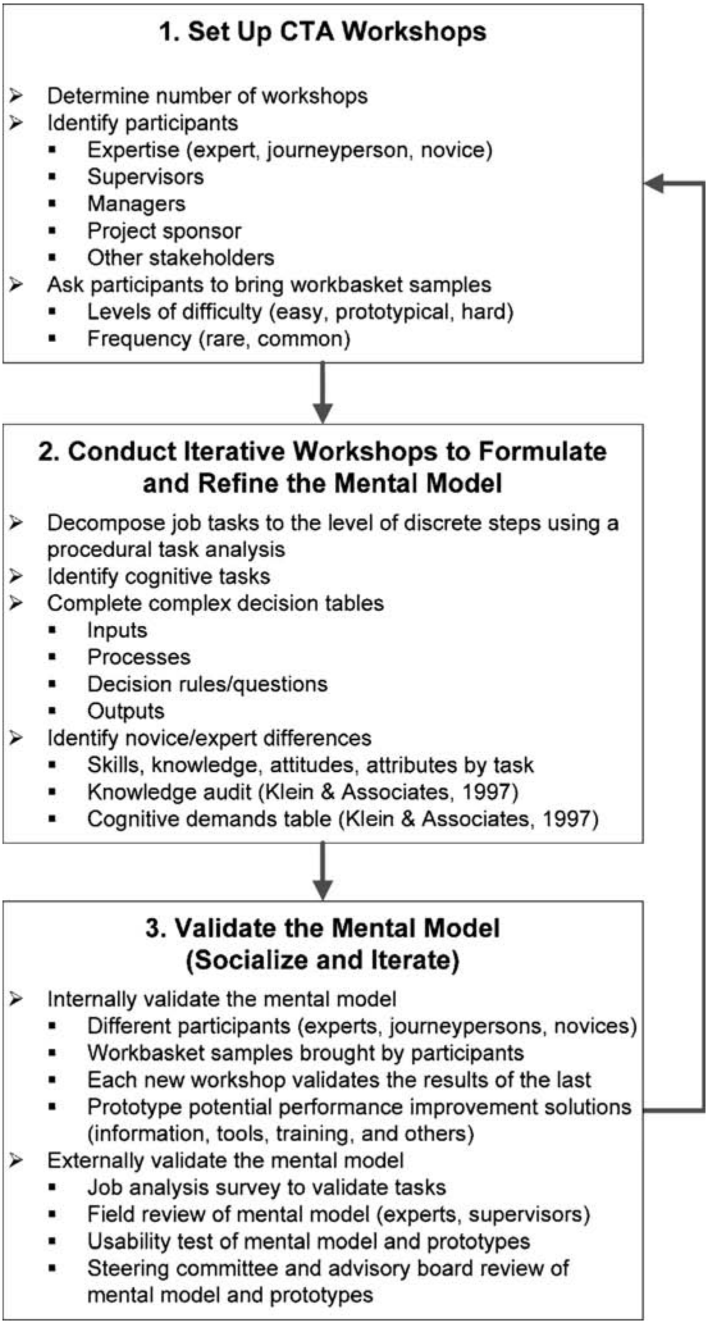


Figure 7.4 F-CBR Phases and Activities.

continuous practice solving progressively difficult problems within a domain to become an expert. This seems to be a common ''rule of thumb'' among CTA practitioners. These authors also wisely note that one of the first things to do is to ask several high-level, experienced people within the organization how long it typically takes to become an expert.

- *Are recognized as ''water-walkers'' within their organizations.* They are people who have spent the time required to acquire expertise. They are the people everyone goes to with their questions about ''Why are we doing this?'' ''How do we get this done?'' and ''How do we know when we do a good job?'' Obtaining release time for these experts reflects a significant short-term pain for the organization in the hopes of making longer-term performance gains when everyone in the organization acts more like the experts. Beware of any so-called experts who have lots of available time to work with you as they rarely provide trustworthy data. Likewise, be aware of trainers who fancy themselves experts—especially when no one else in the organization does.
- *Possess an observable track record of exceptional performance.* Typical indicators include superior accuracy and efficiency of work completed; marked absence of re-work; awards and other recognition; and authorship of internal reports, white papers, and publications in the professional literature. The goal is to pick experts on the basis of outstanding performance, rather than popularity or success at playing organizational politics.
- *Have developed twin areas of expertise: the content domain and teaching people about how to perform their jobs.* Within just about any community of practice, we have found a small population in the expert community who have spent their professional careers developing domain expertise and teaching it to others. They are experts who can provide tested insights about the organization of their knowledge and thought processes. In our experience, people with twin domain and communication/mentoring/instruction expertise comprise somewhere between 10 and 15 percent of an organization's experts.
- *Are willing to share their expertise.* Ultimately, participation in a CTA is voluntary. Some individuals may not want to give away what they have spent a professional career accumulating. Some organizations actually disincent CTA participation. For example, a sales person asked to share expertise in planning sales calls may actually be punished by lower sales figures during his or her participation in the CTA and later by the improved sales rates of colleagues relative to his or her own.

We ask the experts and journeypersons to bring workbasket samples to the workshop. These are the interim and deliverable materials they produce on the

Tasks

+ Sub-tasks

+ Steps

+ Complex decision tables for appropriate steps

Mental Model

Figure 7.5 F-CBR.

job. These often include case files, which typically consist of evidence, chronological logs, analyses, reports, and the like. We always ask that participants bring a prototypical sample with them to the meeting—work that the participant commonly performs. This request prevents participants from bringing their ''Magnum Opus'' to the workshop, which is typically unrepresentative of anything else they have done. Depending on the size and complexity of the project, we may also ask that they bring easy and hard samples as well.

2. Conduct Iterative Workshops to Formulate and Refine the Mental Model. As depicted in Figure 7.5, F-CBR produces a mental model comprised of observable, decomposed job tasks and tables representing the domain knowledge and complex decisions associated with specific cognitive steps. The mental model represents a network of knowledge, not just a list of isolated pieces of knowledge. The procedural task analysis combined with the decision tables forms a mental model capable of producing expert-like performance.

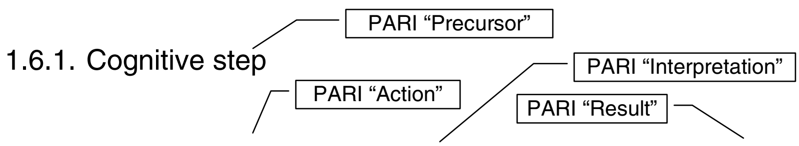
Mental Model. The workshops that produce such mental models are iterative in nature. Regardless of the number of workshops comprising the effort, we begin by establishing the tasks that comprise the job and then specify additional levels of detail until discrete steps emerge. We then create complex decision tables for those steps requiring complex cognition. Each refinement to the decomposed tasks or complex decision tables may require additional revision to reconcile.

This iterative approach is important to the overall success of the F-CBR effort. Letting SMEs talk off the top of their heads is a recipe for obtaining a lot of useless data. Rather, the responses that SMEs provide need to be framed, so they focus only on one particular step and its specific inputs, processes, decision rules, and outputs. The Applied Cognitive Task Analysis (ACTA) technique uses a similar iterative strategy to frame SME responses. An initial task diagram provides a high-level depiction of major tasks and subtasks. A subsequent knowledge audit uses a set of probing questions to describe different facets of expertise within the domain and to elicit appropriate examples. A knowledge audit table correlates aspects of expertise to cues/strategies and areas of difficulty. A subsequent simulation interview table correlates discrete events comprising a larger occurrence to actions, assessments, critical cues, and errors (Militello & Hutton, 1998).

Having conducted a procedural task analysis to decompose the behavioral components of the job, we use several convergent techniques to identify those steps that are cognitive in nature and the future focus of complex decision tables.

- We identify verbs associated with problem solving that appear in the steps (''analyze,'' ''evaluate,'' ''determine,'' ''assess,'' ''judge,'' ''notice,'' ''interpret,'' ''prioritize,'' ''anticipate,'' or ''plan'').
- We ask workshop participants to identify those steps they consider difficult because of the thought process and knowledge associated with them.
- For each step identified using the first two methods, we ask whether novices already know how to perform it. If they do, then no additional detail is needed.

Having identified the cognitive steps within the task, we create complex decision tables that represent the thinking process experts use to complete each step. A sample complex decision table appears in Figure 7.6 and is based on Foshay's adaptation of the PARI approach (personal correspondence, June 6,



Inputs (What Prompts You)	Process (Activities)	Decision Rule (Questions That Help You Make a Decision)	Outputs (Results)

Boldfaced text indicates high-priority questions that take precedence over others

Figure 7.6 F-CBR Complex Decision Table.

Callouts highlight relationships to the PARI (''Precursor,'' ''Action,'' ''Results,'' ''Interpretation'') CTA approach.

2007). The step in the task list corresponds to the ''Precursor'' in the PARI CTA technique. The step represents the goal, or the state the performer wants to reach, by making a complex decision. Inputs specify stimuli on hand when experts make a decision. Stimuli describe cues and resources available to experts. Processes (''Actions'' in PARI) represent any sub-steps expert perform. Decision Rules (''Interpretations'' in PARI) represent what experts focus on to complete the step. In settings where the problem is well-defined, such as troubleshooting, formal decision rules (''if/then/else'' statements) can represent the interpretation. In settings where the problem is ill-defined, such as criminal investigations, specifying a detailed set of decision rules would be cost-prohibitive. In these cases, the Decision Rules column specifies the questions that experts typically ask as they perform this step. Some questions take precedence over others, and we represent these ''trump'' questions by boldfacing them. The Outputs column (''Results'' in PARI) specifies the results experts obtain by completing the step. Table 7.5 summarizes common prompts used to elicit each type of information from SMEs.

Having created the expert's mental model, we typically turn our attention to specifying differences in expert and novice cognition and performance. Specifying these differences allows us to formulate performance improvement solutions to bridge these gaps. There are a number of ways to generate this information within an F-CBR workshop.

- *Specify expert/novice differences in skills, knowledge, attitudes, and attributes for each step in the task list associated decision table.* We obtain this information by conducting a dedicated F-CBR workshop where we split experts and novices into two groups, each of which specifies skills, knowledge, and attitudes related to a specific decision table. When that is not practical, we have obtained the novice information by working with supervisors or journeyperson performers with several years' experience performing the job.
- *Create a cognitive demands table (Klein & Associates Inc., 1997).* A dedicated F-CBR workshop reviewing and refining the decision tables can identify thematic components that separate novice and expert performance. Such workshops can also provide additional information about what is difficult about a given decision and potential sources of error. After the workshop, the F-CBR team can characterize each theme based on its cognitive demand, what is difficult about the demand, relevant cues, strategies in the decision rules, and sources of potential errors.

3. Validate the Mental Model (Socialize and Iterate). With F-CBR, validation, socialization, and iteration travel hand-in-hand. The mental models arising from cognitive task analyses are nothing more than representations of organized

Table 7.5 Prompts for Completing a Complex Decision Table

	Inputs (What Prompts You)	*Process (Activities)*	*Decision Rules (Questions That Help You Make Decisions)*	*Outputs (Results)*
Given the decision that appears in the preceding cognitive step, consider posing the following prompts:	• What causes you to make this decision? • What do you typically have on hand when you make this decision? • What resources do you draw on when you make this decision?	• What activities do you make in making this decision? • Which activities do you typically perform before others?	• What questions do you ask yourself to make this decision? • Which questions do you typically ask before others? • Are there any questions that are more important than others?	• What are the results of making this decision? • What do you have that you didn't have before? • Are there any standards associated with the decision you reach?

knowledge. The fact that the information they contain came from the mouths or actions of SMEs does not ensure their completeness, accuracy, or utility. We meet these goals when we create mental models that others can use to produce desired performances, nothing more and nothing less. As the first iteration of the mental model rarely produces desired results, CTA practitioners typically use additional iterations to refine the knowledge they elicit, analyze, and ultimately represent. With each iteration comes the need to socialize results. This social aspect of CTA is very important in F-CBR cases, where the goal is to represent organizational knowledge, rather than the knowledge of a handful of SMEs. Socialization means that larger numbers of SMEs, novices, managers, supervisors, and stakeholders must be involved in the process in order to buy into its results. The cast of characters involved in conducting an F-CBR and creating subsequent training and performance improvement solutions often reads like the cast of characters in a Hollywood movie. This widespread involvement creates a demand for the effort and its results. Such a ''pull'' for the F-CBR and subsequent solutions jump-starts change management and implementation efforts, which begin on Day One of these projects (Stone & Villachica, 2004). F-CBR employs a variety of internal and external strategies to continuously validate and socialize the mental model.

Internal Validation Strategies. To continuously validate, socialize, and iterate the mental model within the F-CBR team, we employ the following strategies.

- *Ask participants to bring workbasket samples that represent the work they do on the job.* F-CBR employs this strategy because experts' recall of their own performance is typically inaccurate and incomplete. F-CBR controls for this tendency in two ways. First, F-CBR frames each complex decision table within a specific step, which focuses SMEs on a manageable ''chunk'' of contextualized expertise. Second, F-CBR checks what SMEs are saying in structured interviews against the workbasket samples they bring to the workshop. If the tasks and decision tables produced at the workshop reflect SME consensus and describe the workbasket samples, then we assume they are internally validated.
- *Ask participants of each new workshop to review and validate the results of the previous workshop.* Having each new group act as a control for the last validates the work of each. In each new workshop, there may be several SMEs who were involved in the previous workshop and several new SMEs and others who were not. Over the course of multiple workshops, the new contributions of each new set of workshop participants begin to stabilize. At some point, the participants will say, ''What we have here is good enough.'' At that point, the job of the workshops is coming to a close.
- *Prototype the mental model as well as potential performance improvement solutions, including information, tools, and training.* As the mental model

begins to settle down, the F-CBR team and workshops should begin shifting focus from fleshing out the mental model to determining how to use it to address the sources of performance gaps identified in the needs assessment. The project team should begin producing rough, conceptual prototypes of the mental model and other potential performance improvement solutions. These proof-of-concept prototypes should indicate how users would employ them in realistic job and training settings.

External Validation. In addition to internal validation techniques, the F-CBR team should also employ external techniques that rely on the collection and analysis of data from sources outside the team. In addition to providing additional rigor to the validation process, these techniques also facilitate organizational buy-in. Various techniques for external validation include conducting a:

- Job analysis survey to validate the tasks using data collected from a representative sample or an entire population.
- Field review of the mental model (tasks and decision tables) to collect the comments and revisions of experts, competent performers, novices, and supervisors who have not been involved in the project.
- Usability test of the mental model and corresponding prototypes, using a representative sample of performers who have not been associated with the F-CBR effort.
- Steering committee and advisory board review of the mental model and corresponding prototypes.

F-CBR AND THE CREATION OF A MENTAL MODEL FOR CRIMINAL INVESTIGATORS

DLS Group is currently partnering with the California Commission on Peace Officer Standards and Training (POST) to improve training and tools for criminal investigators statewide. To provide trustworthy data for use in revising the training investigators receive, POST wanted to know what tasks comprised the investigative process that detectives use on the job. They wanted to generate a set of tasks that cut across different types of crimes, their complexity, and the size of the law enforcement agencies that investigate them. Given this goal, DLS and POST formulated the following questions that would be answered in a job and cognitive task analysis:

1. What tasks comprise the investigative process?

2. What are the characteristics of each task?
   - How often do investigators perform it?
   - How important do investigators consider it?
   - When do investigators typically learn it?
   - Under what conditions do investigators typically perform it?
3. What skills, knowledge, and attitudes (SKAs) are associated with each task at the level of the novice and experienced investigator?
4. What is the mental model that guided investigator performance?

This section of the chapter describes the method we employed and subsequent performance improvement solutions we selected and prototyped.

Method

Answering the fourth question required that DLS and POST conduct a CTA. We opted to use the F-CBR technique for the following reasons:

- *No one had specified the tasks comprising the criminal investigative process before.* Answering Questions 1 through 3 already required that we conduct both a job and procedural task analysis. Further refining this effort to include complex decision tables for cognitive steps was a reasonable and cost-effective extension of the job analysis effort.
- *The scope of this effort required the mental model to apply across all investigations throughout the state.* Owing to its RAD roots, F-CBR is well suited to the large-scale representation of expertise. DLS had previously used F-CBR to specify a mental model of the investigative process that auditors throughout the USA used to audit securities firms (Villachica & Stone, 1998a), and it was reasonable to apply this approach to a state which has the sixth largest economy in the world (Legislative Analyst's Office, 2004).
- *POST and law enforcement agencies agreed to provide unprecedented access to expertise.* The mental model arising from any CTA is only as good as its participants. POST and California's different law enforcement agencies agreed to provide expert, journeyperson, and novice investigators and their supervisors. Obtaining their participation required four to six weeks advanced notification and POST's ongoing logistic efforts to help agencies apply instructions for selecting participants, transport them, and compensate their expenses.

To complete this effort, we would conduct a procedural task analysis to specify the tasks and steps comprising the investigative process. Then we would conduct a CTA using F-CBR to specify the inputs, processes, decision rules, and outputs of those steps in the investigative model that involved a lot of decision making. Based on the tasks, their characteristics, SKAs, and mental model, POST and DLS collaborated to create prototypes of the revised investigator training and a transition plan.

To create the mental model, DLS and POST employed the method depicted in Figure 7.7. We initially conducted a kickoff workshop with our steering committee. This body was comprised of eleven experienced investigators with over two hundred years' worth of hard-won expertise. Their primary duty lay in providing strategic direction for the project. They also initially helped

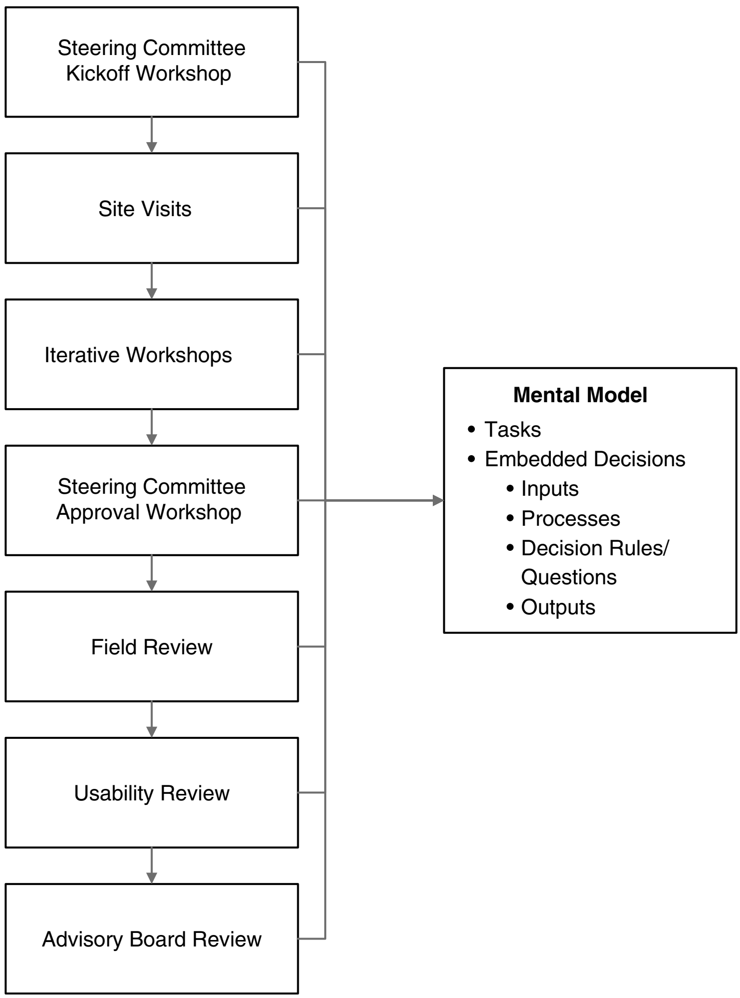


Figure 7.7 Participants Involved in the POST Job and Cognitive Task Analysis. Producing a Mental Model of the Investigative Process.

us determine the major tasks comprising the investigative process and gain a sense of the characteristics of each task. We also conducted two site visits and one teleconference to interview investigators, sergeants, and commanding officers about law enforcement, criminal investigations, and the investigative process. We then held a series of three iterative workshops, working with forty-six expert detectives. In the first workshop, journeyperson investigators who had been on the job for a couple of years worked with us to refine the list of tasks comprising the investigative process and complete an initial iteration of the mental model.

We then facilitated two workshops with expert investigators and supervisors to further extend and refine the mental model. One group consisted solely of experienced investigators who averaged between ten and fifteen years as detectives. One detective had more than fifty years of experience. His email address was ''Mr. Robbery.'' Another group consisted of former expert detectives who now worked as supervisors who coached new detectives. Each new group checked and refined the work of previous ones, until we started reaching consensus that we had adequately specified the mental model.

The mental model then went to the steering committee again for its final review and approval. We then conducted a field review, where supervisors and investigators reviewed and approved the mental model. We finally presented the mental model to POST's ICI Advisory Board for their final approval. This board consists of agency chiefs, executive level stakeholders, and experts.

Intervention Selection and Prototype Development

Combining data collected in the job and cognitive task analysis with a gap and cause analysis, DLS and POST worked together to create initial prototypes of performance improvement solutions to support novice and expert investigator performance. Our gap analysis revealed that we wanted more investigators to act more like experts. This meant supporting novices coming from patrol who had to learn many new investigative skills in their first year on the job, as well as journeypersons and experts who had to stay current in a rapidly evolving field. Using Gilbert's (2007) behavioral engineering model (BEM), we identified the causes of these performance gaps, which appear in Table 7.6. Using Gilbert's logic, recommended sequence, we would address the environmental causes of the performance gaps before causes arising from deficits in the personal repertory. In other words, we would first look at information and tools. Then we would look at the role these information and tools would play in training for new investigators.

As online instruments (information systems) now deliver data, including ''clear and relevant guides to adequate performance'' (Gilbert, 2007, p. 88), the causes of the performance gaps for the Data and Instruments cells are the same. This is a common finding in situations that lend themselves to performance support. Based on these findings, POST and DLS decided to prototype an online

Table 7.6 BEM Specifying Environmental and Personal Causes of Investigative Performance Gaps Identified in the POST Job Analysis and CTA Effort

	Information	*Instrumentation*	*Motivation*
Environmental Supports	*Data*	*Instruments*	*Incentives*
	Lack of access to a mental model	Lack of access to a mental model	Not applicable
	Lack of access to encyclopedic investigation information	Lack of access to encyclopedic investigation information	
	Lack of access to investigative stories	Lack of access to investigative stories	
Personal Repertory	*Knowledge*	*Capacity*	*Motives*
	Lack of investigative skill and knowledge	Not applicable	Not applicable

mental model, investigator reference system, best practices database, and elements of new investigator training based on a cognitive apprenticeship model.

Mental Model

The mental model existed in the heads of experienced investigators. We needed to make it manifest so that novices, journeyperson investigators, and experts alike could access it. To this end, DLS and POST created three iterative prototypes of the mental model. The first iteration of the mental model was an MS-Word prototype we iteratively refined in the workshops to specify all of the tasks comprising the investigative process and the decision tables associated with the cognitive steps. The second prototype was a hypertext model, rapidly created using MS-Visio, Adobe Dreamweaver, Adobe Acrobat, and MS-Paint, that ran over Internet Explorer. This prototype demonstrated the potential feasibility of putting the mental model online.

The third iteration of the mental model prototype is currently under refinement. Figure 7.8 depicts the components of this third-generation prototype. In addition to making the invisible investigative process visible, we expect the mental model to also act as a mechanism to provide timely access to contextualized resources. This feature implements a central tenant of performance support—providing what people need, when they need it, in the form they

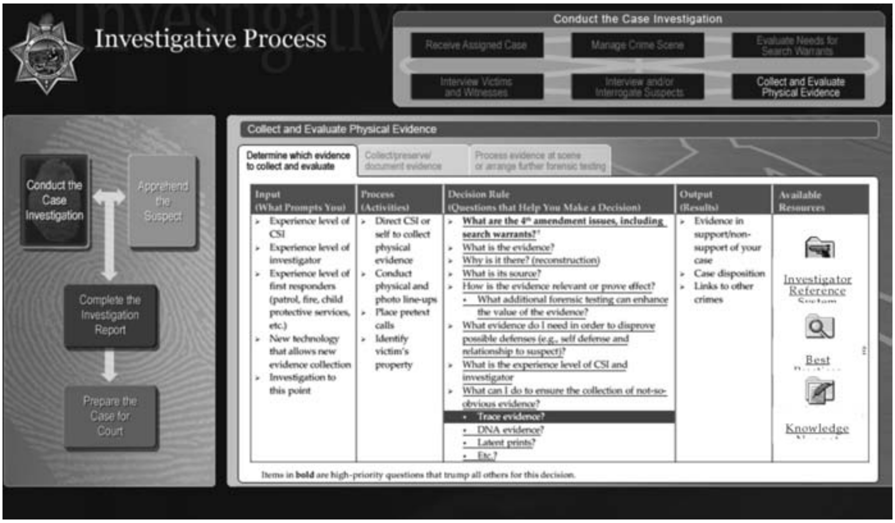


Figure 7.8 Third-Generation Prototype Depicting the Mental Model of the Investigative Process.

need in ways that let them perform their jobs to meet organizational objectives (Villachica, Stone, & Endicott, 2006).

Primary tasks appear on the left of Figure 7.8, with ''Conduct the Case Investigation'' selected. Tasks appear in the upper-right corner, with ''Collect and Evaluate Physical Evidence'' selected. Tabs depict subtasks, with ''Determine which evidence to collect and evaluate'' selected. Highlighting one of the decision rules displays relevant available resources contained in the Investigator Reference System, Best Practices Database, and Knowledge Nuggets.

Investigator Reference System

The knowledge and information that investigators employ can be highly technical and change rapidly. Investigative techniques and crimes co-evolve, where the actions of investigators and suspects continuously affect the other. New technical advancements offer the ability to obtain DNA matches from smaller samples. Suspects now plan, commit, and relive their crimes over cellular phones. Suspects now police their own crime scenes, picking up any spent cartridges so that the criminal investigators cannot find them. During our workshops, we noticed that all investigators continuously shared information with others about investigative techniques. Some of this information bordered on encyclopedic. Realizing that the need to access this type of information was largely unmet, DLS and POST prototyped elements of an investigator reference system, which would act as an evergreen, online encyclopedia of useful investigative techniques (see Figure 7.9).

Sworn officers (investigators, patrol, supervisors, captains, and chiefs) would use the online system shown in Figure 7.9 to find timely, complete, and relevant information about evidence principles and investigative techniques. In this example, information about trace evidence appears in a standard format for organizing any type of evidence. Each entry contains assembled materials from new and existing text, graphic, and video assets. Sworn officers would also be able to submit new information to the system, which would appear after a formal vetting process.

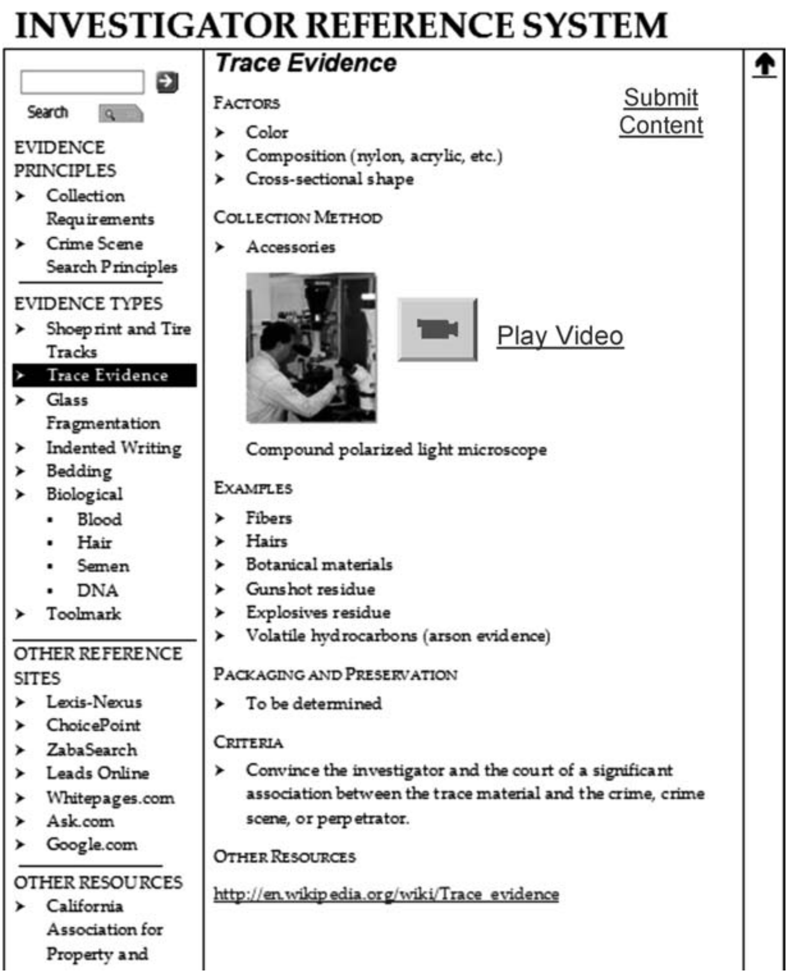


Figure 7.9 First-Generation Prototype Depicting Elements of an Online Investigative Reference System.

Best Practices Database. In addition to watching investigators share bits of encyclopedic knowledge throughout the F-CBR workshops, they also told each other stories about the criminal investigations they conducted. Often, these stories continued long after the workshop had ended and deep into the night. In our experience, SMEs often tell these stories to illustrate previously unknown features of the mental model, which are often associated with the ''punch lines'' of the stories. Sitting in on such stories can also provide an informal means to validate and tune the mental model.

Jonassen and Hernandez-Serrano (2002) have noted that such stories can be elicited, ''indexed for the lessons they have to teach, and made available to learners in the form of case libraries that can support a broader range of problem solving than any other strategy or tactic'' (Jonassen & Hernandez-Serrano, 2002, p. 65). With this in mind, DLS and POST prototyped a best practices database to act as the repository for these stories, which could be captured in text, audio, or video formats. Similar to the investigator reference system, stories appearing in the best practices database would be submitted by sworn officers and vetted through a formal process.

Cognitive Apprenticeship Training. Combining data obtained in the job analysis, the CTA, and the analysis of skills, knowledge, and attitudes, we identified skill and knowledge deficits related to interpretation, solving problems, and making decisions. Collins, Brown, and Holum (1991) describe a set of principles for designing cognitive apprenticeship environments that make invisible problem-solving skills visible for learning and performing. This instructional approach has a proven track record in improving performance in medical, legal, securities examination, military, and other areas requiring complex problem solving and decision making (cf., Czech, 1999; Lee & Jackson, 2007; Roesch, Gruber, Hawelka, Hamm, Arnold, Popal Segerer, Landthale, & Stolz, 2003; Williams, 1992). Initial training for novices would teach them to apply the mental model while using the Investigator Reference System and Best Practices Database and other online resources to manage and investigate cases. Capstone activities would include simulated investigations. This training would be highly interactive, involving problem-based group exercises, simulated investigations, and lab practica. In addition to promoting learning by doing, these learning activities foster the modeling, articulation, and reflection needed to help novices construct their own investigative skills and knowledge.

SUMMARY

In the continuum of analytical activities that instructional designers use to specify job performances, F-CBR offers a cognitive task analysis technique to identify, analyze, and represent knowledge at the enterprise or state level.

Practitioners can further refine mental models created using F-CBR using other CTA techniques, including applied cognitive task analysis (ACTA) (Klein & Associates Inc., 1997) and concept mapping (Crandall, Klein, & Hoffman, 2006). Owing to its collaborative nature, F-CBR offers a time-effective mechanism for gathering consensus expertise from a representative sample of SMEs. This expertise is vetted by competent performers and novices who ensure that what the SMEs provide is understandable at all levels of expertise. F-CBR also offers an efficient mechanism to collect workbasket samples and stories that later find their way into information systems and training. These CTA participants tend to become project ambassadors for the program, thereby creating a demand for the changes to the job that performance solutions based upon the mental model will bring.

References

Berry, D., & Dienes, Z. (1993). Towards a working characterization of implicit learning. In D. C. Berry & Z. Dienes (Eds.), *Implicit learning: Theoretical and empirical issues* (pp. 1–18). Hove, UK: Lawrence Erlbaum Associates.

Broadbent, D. E. (1977). Levels, hierarchies, and the locus of control. *The Quarterly Journal of Experimental Psychology*, *29*(2), 181–201.

Clark, R. E., Feldon, D. F., van Merriënböer, J.J.G., Yates, K. A., & Early, S. (2007). Cognitive task analysis. In J. M. Spector, M. D. Merrill, J.J.G. van Merriënböer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (pp. 1801–1856). Mahwah, NJ: Lawrence Erlbaum Associates.

Collins, A., Brown, J. S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator: The Professional Journal of the American Federation of Teachers*, *15*(3), 6–11, 38–46.

Crandall, B., Klein, G., & Hoffman, R. (2006). *Working minds: A practitioner's guide to cognitive task analysis*. Cambridge, MA: MIT Press.

Czech, C. (1999). Becoming expert faster using situated learning. In B. Sugrue & J. Fuller (Eds.), *Performance interventions* (pp. 219–231). Alexandria, VA: American Society for Training & Development.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, *49*(8), 725–747.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Gilbert, T. F. (2007). *Human competence: Engineering worthy performance* (tribute ed.). San Francisco: Pfeiffer.

Gott, S. P., & Lesgold, A. M. (2000). Competence in the workplace: How cognitive performance models and situated instruction can accelerate skill acquisition. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 239–327). Mahwah, NJ: Lawrence Erlbaum Associates.

Hall, E. P., Gott, S. P., & Pokorny, R. A. (1995). *A procedural guide to cognitive task analysis: The PARI methodology*. Brooks AFB, TX: Armstrong Lab. (Report AL/HR-TR-1995-0108).

Hoffman, R. R., Coffey, J. W., Ford, K. M., & Novak, J. D. (2006). A method for eliciting, preserving, and sharing the knowledge of forecasters. *Weather and Forecasting, 21*, 416–428.

Jonassen, D. H., & Hernandez-Serrano, J. (2002). Case-based reasoning and instructional design: Using stories to support problem solving. *Educational Technology Research and Development*, *50*(2), 65–77.

Jonassen, D. H., Tessmer, M., & Hannum, W. (1999). *Task analysis methods for instructional design*. Mahwah, NJ: Lawrence Erlbaum Associates.

Kieras, D. E., & Santoro, T. P. (2004). Computational GOMS modeling of a complex team task: Lessons learned. Paper presented at the Conference on Human Factors in Computing Systems. Vienna, Austria.

Klein & Associates Inc. (1997). *ACTA instructional package* (booklet). North Fairborn, OH: Author.

Klein, G. A. (1989). Recognition-primed decisions. In W. B. Rouse (Ed.), *Advances in man-machine system research* (Vol. 5, pp. 4–92). Greenwich, CT: JAI Press.

Klein, G., & Militello, L. (2005). The knowledge audit as a method for cognitive task analysis. In H. Montgomery, R. Lipshitz, & B. Brehmer (Eds.), *How professionals make decisions* (pp. 335–342). Mahwah, NJ: Lawrence Erlbaum Associates.

Klein, G. A., Orasanu, J., Calderwood, R., & Zsambok, C. E. (Eds.). (1993). *Decision making in action: Models and methods*. Norwood, NJ: Ablex.

Kolodner, J. L. (1992). An introduction to case-based reasoning. *Artificial Intelligence Review, 6*(1), 3–34.

Kolodner, J. L. (1993). Case-based reasoning. San Mateo, CA: Morgan Kaufmann Publishers, Inc.

Lee, B., & Jackson, M. (2007). Using e-Learning to improve your instructors' teaching skills. Paper presented at the ASTD TechKnowledge Conference. Las Vegas, NV.

Lee, R. L. (2004). The impact of cognitive task analysis on performance: A meta-analysis of comparative studies. Rossier School of Education, University of Southern California.

Legislative Analyst's Office. (2004). *Cal facts 2004: California's economy and budget in perspective*. Sacramento, CA. Available from http://www.lao.ca.gov/2004/cal_facts/2004_calfacts_econ.htm.

Martin, J. (1991). *Rapid application development*. New York: Macmillan Publishing Company.

Militello, L. G., & Hutton, R. J. B. (1998). Applied cognitive task analysis (ACTA): A practitioner's toolkit for understanding cognitive task demands. *Ergonomics*, *41*(11), 1618–1641.

Ockerman, J. J., & Mitchell, C. M. (1999). Case-based design browser to support software reuse: Theoretical structure and empirical evaluation. *International Journal of Human-Computer Studies*, *51*(5), 865–893.

Olson, J. R., & Biolsi, K. J. (1991). Techniques for representing expert knowledge. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 240–285). New York: Cambridge University Press.

Roesch, A., Gruber, H., Hawelka, B., Hamm, H., Arnold, N., Popal, H., Segerer, J., Landthale, M., & Stolz, W. (2003). Computer assisted learning in medicine: A long-term evaluation of the Practical Training Programme Dermatology 2000. *Medical Informatics*, *28*(3), 147–159.

Schank, R. C. (1990). Case-based teaching: Four experiences in educational software design. *Interactive Learning Environments*, *1*(4), 231–253.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66.

Schraagen, J. M., Chipman, S. F., & Shalin, V. L. (Eds.). (2000). *Cognitive task analysis*. Mahwah, NJ: Lawrence Erlbaum Associates.

Schvaneveldt, R. W., & Durso, F. T. (1981). Generalized semantic networks. Paper presented at the Annual meeting of the Psychonomic Society. Philadelphia, PA.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*(2), 127–190.

Stone, D. L., Clark, R. C., Foshay, W. R., & Villachica, S. W. (1999). Cognitive task analysis for HPTers: A panel discussion. Paper presented at the 37th Annual Conference of the International Society for Performance Improvement. Long Beach, CA.

Stone, D. L., & Villachica, S. W. (2004). Lessons learned from Victoria's Secret: Ensuring the ''pull'' in performance improvement. *Performance Improvement*, *43*(8), 5–7.

Stone, D., & Villachica, S. (2005). Want reliable metrics for your organization? Use prototyping! Paper presented at the ISPI Instructional Systems Conference. Las Vegas, NV.

Velmahos, G. C., Toutouzas, K. G., Sillin, L. F., Chan, L., Clark, R. E., Theodorou, D., & Maupin, F. (2004). Cognitive task analysis for teaching technical skills in an inanimate surgical skills laboratory. *The American Journal of Surgery, 187*, 114–119.

Villachica, S. W., & Stone, D. L. (1998a). CORNERSTONE: A case study of a large-scale performance support system. In P. J. Dean & D. E. Ripley (Eds.), *Performance improvement interventions: Performance technologies in the workplace* (pp. 437–460). Washington, DC: International Society for Performance Improvement.

Villachica, S. W., & Stone, D. L. (1998b). Rapid application development for performance technology: Five strategies to deliver better interventions in less time. In P. J. Dean & D. E. Ripley (Eds.), *Performance improvement interventions: Performance technologies in* *the workplace* (pp. 343–399). Washington, DC: International Society for Performance Improvement.

Villachica, S. W., Stone, D. L., & Endicott, J. E. (2006). Performance support systems. In J. Pershing (Ed.), *Handbook of human performance technology: Improving individual and organizational performance worldwide* (pp. 53–566). San Francisco: Pfeiffer.

Williams, S. M. (1992). Putting case-based instruction into context: Examples from legal and medical education. *The Journal of the Learning Sciences*, *2*(4), 367–427.

Zimmerman, L. A., Burns, C. A., & Sestokas, J. M. (2007). High-fidelity simulation/mentoring system for general officers in high-stakes operational environments. Paper presented at the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), Orlando, FL.

PART THREE

INSTRUCTIONAL STRATEGIES

This is the part of the book novice and moderately experienced IDs will probably turn to first. If you are an experienced instructional designer, you will certainly find that some of the proposed strategies in this part will give you new ways to address important training needs. You may also find that some of what you think you know about how to construct instruction—regardless of medium or type—is now greatly elaborated, or it may be outdated. For example:

- The debate over direct instruction versus constructivist strategies has often ignored a great deal of the basic research on instruction. You are likely to conclude that it's a false dichotomy, largely irrelevant to what really matters.
- The so-called ''principles of adult learning'' aren't found here, because they aren't. They ignore what the research tells us really makes a difference about effective instruction of any type for any learner.
- Cognitive load is probably among the most important, and least widely understood, considerations in design of instructional strategies.
- Knowledge structures really matter, and your instruction must help the learner build the right ones. You can actually do damage if you don't get this right when you design your instruction.
- There are emerging new strategies for teaching the highest-value problem-solving skills, which your clients (and many trainers) often regard as unteachable.
- Instructional simulation and serious games are important vehicles for many of these high-value skills, but it is easy to build the wrong one.

- Advances in cognitive learning are leading to exciting new strategies for more efficient, effective, and generalizable psychomotor learning.

This part of the book is about the underlying structure of the training you create. The principles described are general and apply to all types of learners and all types of learning environments in all media. Studying this section will help you think structurally about that portion of the process of instructional design that probably occupies most of your time. Most of what is presented has a strong basis in research, and should be a part of all instructional design practice. However, as we caution below, some of the design practices must be treated as emerging or promising practices.

Chapter Eight: An Analysis of the Failure of Electronic Media and Discovery-Based Learning: Evidence for the Performance Benefits of Guided Training Methods. This first chapter in Part Three presents the basic research on instruction and directly addresses many of the misconceptions that recur in the literature, particularly surrounding electronic learning and constructivism. The chapter interprets a generation of research on the principles of instructional strategy that is foundational to the field of instructional design. Understanding these principles will help you concentrate on what really matters about effective instruction and will help to keep you from the superficial trendiness that plagues much of the training and development field.

Chapter Nine: Instructional Strategies for Receptive Learning Environments. The next chapter summarizes research-based principles for design of any instructional communication—or, for that matter, any technical communication. These principles go well beyond what your first writing teacher taught you and start with the premise that all learning is active learning, in which the essential work is done by the learner, not by the author, not by the instructor. Your job is to facilitate this essential work. Following this principle systematically will lead you to a better understanding of how to structure your instructional communications and will help you avoid many poor (but popular) practices.

Chapter Ten: Instructional Strategies for Directive Learning Environments. As in the preceding chapter, the authors first summarize important lessons from research on learning theory. Then they systematically build from this theory a set of principles for strategies of direct instruction, including practice and feedback. Much of what you find here will be new to you, regardless of your level of experience in developing instruction, and regardless of the medium in which you work.

Chapter Eleven: Assembling and Analyzing the Building Blocks of Problem-Based Learning Environments. The author points out that current cognitive theory has led to some important advances in our understanding of how experts solve problems. We have a much deeper understanding of what experts know and how they know it, especially for problems that are unique and non-recurring. The current challenge in the state of the art is to define instructional strategies to help novices become experts. This chapter presents a ground-breaking framework of problem types and teaching/learning strategies. This is emerging design theory, which will require years of design experiments to fully develop and validate. It is presented here because it represents an important new tool set for high-value capabilities in any organization.

Chapter Twelve: High Engagement Strategies in Simulation and Gaming. This chapter brings out that simulation and gaming are among the most important types of training for teaching cognitive strategies and problem solving. Unless you work in a military environment, there is a good chance that you are unaware of the largest, oldest, and most sophisticated body of research on simulation. This chapter summarizes important design principles for simulation from this work based on current cognitive theory. It also makes the point that these principles apply equally to ''serious'' games.

Chapter Thirteen: Video Game–Based Learning: An Emerging Paradigm for Instruction. The author discusses that, while ''serious'' games are generating a great deal of interest currently, research on effective principles of design, other than the general principles presented in previous chapters, is still in its infancy. This chapter presents an emerging theory of design for online games. We caution that the principles presented here will require years of design experiments to refine and validate. However, we hope this chapter will help you extend the basic principles presented in earlier chapters to this important new genre of training.

Chapter Fourteen: Training Complex Psychomotor Performance Skills: A Part-Task Approach. The final chapter in this part addresses another important area of emerging design theory, psychomotor learning. The chapter shows how current cognitive learning theory can be interpreted to derive new instructional strategies that minimize the need for often-costly whole-task simulation, while increasing both training efficiency and generalizability. Again our caution: this is state-of-the-art work. While initial research on this design strategy is promising, the principles described here will require years of design experiments to refine and validate them further.

CHAPTER EIGHT

An Analysis of the Failure of Electronic Media and Discovery-Based Learning

*Evidence for the Performance Benefits of Guided Training Methods**

Richard E. Clark
Kenneth Yates
Sean Early
Kathrine Moulton

INTRODUCTION

This chapter will present a direct, evidence-based argument that, while media provide economic benefits for training organizations, they have not and will not influence learning, motivation, or work performance. We begin with a discussion of popular instructional design models based on discovery and problem-based learning and argue that a half-century of research has indicated that they

* The project or effort described here has been partially sponsored by the U.S. Army Research, Development, and Engineering Command (RDECOM). Statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred. The authors wish to acknowledge that some sections of this chapter have been drawn from previously published manuscripts or technical reports and were used by permission, and we wish to acknowledge our debt to Dr. David Feldon for some of the material in the discussion of adaptable learning.

are also ineffective for all but a small minority of learners. We will briefly describe the half-century of research that supports our conclusions and describe the consequences for business and education. Contrary to popular belief regarding the importance of media in training, we will suggest that a handful of specific training methods are the only environmental factors that have been found to have a major influence on learning and performance. We will argue that the methods we describe are successful in many different delivery media because they support the mental process by which people learn complex knowledge. We will then describe an example of the current training models that promote guided learning. The chapter will conclude with a description of a powerful tool for selecting the most cost-beneficial media to deliver guided learning methods for nearly any training or performance goal.

Education and training organizations are always alert to new developments that have the potential to increase the effectiveness and cost-benefit of instruction. This chapter will examine the research and best practice evidence for current and future instructional innovations that help and hinder instructional support for learning and performance. The discussion begins with a description of popular instructional approaches that have been found ineffective and/or to cause problems. We then go on to describe recent developments that appear to add significant value to learning and performance support systems.

Some of the innovations that have been a constant focus of attention in training are the exciting developments in new technologies such as computers, multimedia, and virtual reality. The most common assumption is that new media are more motivating than older media and that increased motivation will lead to significantly more learning and performance. The discussion turns next to the evidence for these assumptions.

THE IMPACT OF MEDIA ON LEARNING

The past half-century of research, evaluation, and best practice evidence about learning from instruction has established that the choice of media does not influence learning or motivation. While people have argued about this conclusion (see, for example, a review of the arguments by Clark, 1991, 2001), the current view that is most widely accepted in research and evaluation is that media only deliver instruction but do not influence learning. The reason for this conclusion is that whenever we find a learning or performance benefit from instruction presented in a new medium or mix of media, we also find an equivalent increase in learning from a different medium or mix (Clark, 2001). If, for example, learning requires instructional strategies such as a demonstration of how to solve a class of problems and practice on an example of the problem while providing corrective feedback, the media used for the demonstration, practice, and feedback will not influence learning. In the past decade or so, many schools, business, government, and military training organizations have made extensive use of computer and web or Internet-based ''distance learning.'' Many of those organizations have transferred instruction currently being offered in the classroom to the computer, Internet, or multimedia and have evaluated both offerings before making a long-term commitment to distance education. These are interesting natural experiments because the skills and knowledge being taught in two versions of a course are similar, but the medium being used is different, and sometimes instructional methods such as demonstration and practice are formatted differently in different media. Yet Bernard, Abrami, Lou, and Borokhovski (2004) surveyed all of the 688 comparisons they located of classroom and distance learning offerings of the same course content conducted prior to 2004 and found no differences in either learning or motivation. More recently, Sitzmann, Kraiger, Stewart, and Wisher (2006) reviewed ninety-six studies focused primarily on adults in business and college settings and found the same result. Web-based instruction and instructor-delivered, classroom-based instruction produced the same amount of learning, while classroom and web instruction were equally motivating. While some studies showed definite benefits from distance learning technology over the classroom, others showed better performance with classroom instruction. When this happens in research, it is most often the case that the instructional design is making the difference, not the medium used to deliver instruction. Even more of these studies show ''no significant differences'' between the classroom and the computer. Bernard, Abrami, Lou, and Borokhovski (2004) referenced five previous large-scale reviews of media comparison studies that have reached exactly the same conclusion. It is important to mention that both the Bernard and Sitzmann studies were comprehensive and inclusive. They included only studies in which different media were used to deliver the same course content, even when the instructional methods used had been reformatted to accommodate a newer media.

Arguments About the Impact of Media on Learning

Clark (2001) has reviewed all published studies and reviews of studies where the effects of different media were compared and reaches a conclusion identical to the Bernard and Sitzmann teams. He suggests that the reason some studies show benefits for certain media is because the researchers mistakenly inserted different information content or instructional methods in one of the media but not in the comparison media. In these poorly designed studies, differences in learning were due to providing more or different information and/or learning support to a group using one medium that was necessary to succeed at a test, while the same information or support was unintentionally denied to another group who received a different medium.

Robert Kozma (1994) has argued with the conclusion that media do not cause learning. He emphasizes the potential of different media to tailor instructional support to the unique learning needs of individuals and groups. He also argues that different media offer different kinds of instructional support, and so it is impossible to separate media from their unique learning support capabilities. The debates about this issue extended over a number of years but was finally resolved a few years ago when Kozma acknowledged (see Clark, 2001) that there was no evidence that media caused learning or that any one medium offered a unique learning support. Kozma remains optimistic that in the future we will learn more about the unique learning support capabilities of different media. However, at the present time it is simply not possible to identify any type of instructional support that is a function of any one medium. If more than one medium offers the same learning supports, then the choice between them is based on cost and availability—not on learning benefits.

Why Don't Media Make a Difference in Learning?

The reason media do not make a difference in learning is captured in the analogy that media are ''mere vehicles that deliver instruction but do not influence student achievement, any more than the truck that delivers our groceries causes changes in our nutrition'' (Clark, 1983, p. 445). Nutrition is a result of the way that food is grown, prepared, and consumed. Any food can be delivered by a variety of transportation vehicles, including trucks. Of course, different kinds of food require different vehicle design. Frozen food requires insulation and refrigeration. Yet many different kinds of insulated vehicles can carry frozen food. The point is that, if there is more than one medium that can carry instruction with the same learning outcome, and if this is the only conclusion possible from all the evidence at hand, then the only difference between those different media is their capability to influence the cost of instruction. Thus, it is important to focus our study and selection of media not on learning and performance gains, but on possible economic benefits.

Another analogy that illustrates why media do not cause learning can be found in the way that medication compounds are prepared for delivery. For example, aspirin is a special compound called acetylsalicylic acid and is incorporated into a number of inert ''carrier'' ingredients and delivered to the consumer in a variety of media such as tablets, liquid suspensions, candy, or gum. All of these different media serve to deliver the same ''active'' aspirin ingredient with different levels of efficiency, but with equal effects on our physical symptoms. Media are not the active ingredients in instruction that cause learning, but simply the vehicles by which it is delivered.

If Not Media, What Does Cause Learning?

When the Bernard team (2004) and Sitzmann team (2006) looked more closely at studies in which either the classroom or the distance learning version of a course was more effective, they discovered the factor that appears to cause most of the learning benefit. They called that factor ''instructional methods.'' Instructional methods or strategies are the ''active ingredients'' in instruction in the same way that special compounds are the active ingredients in medications. The problem is that we have been tempted to assume that media are an active ingredient in learning and motivation.

Clark (1991) defines instructional methods as ''any way to shape information that (supports) . . . the cognitive processes necessary for achievement . . . If students cannot (or will not) give themselves an adequate example, an instructional presentation must provide it for them'' (p. 34). An important cognitive process that is essential for learning is to connect new information to similar prior knowledge. This way, students can draw on what they know in order to elaborate and understand something new. For example, after children learn to add and subtract whole numbers, they are asked to learn fractions. The challenge of imagining ''less than one'' is daunting until instruction provides them with the ''slice of pie'' analogy. When reminded that one pie can be shared equally among many by slicing it and that each slice is a fraction of the whole pie, most students connect what they know about sharing pies and cakes and are well on their way to learning to add and subtract fractions. When the Bernard and Sitzmann research teams looked at instructional programs that produced more learning, they found similar instructional methods. Later in the chapter we will describe what appear to be the most powerful instructional methods. At this point, the discussion turns next to evidence about the motivational qualities of media.

THE IMPACT OF MEDIA ON MOTIVATION

Nearly all educators have hoped that newer media will make instruction more engaging and interesting for learners of all ages. Most of us can recall painful memories of feeling trapped in a classroom and subjected to poor instruction. These memories help fuel interest in using technology to foster student excitement about learning. What follows is a description of motivation and how it influences learning, as well as a review of the evidence about the motivational qualities of newer media.

What Is Motivation?

Clark (2003) described motivation as:

> ''The process that initiates and maintains goal-directed performance. It energizes our thinking, fuels our enthusiasm and colors our positive and negative emotional

> reactions to work and life. Motivation generates the mental effort that drives us to apply our knowledge and skills. Without motivation, even the most capable person will refuse to work hard. Motivation . . . nudges us to convert intention into action and start doing something new or to restart something we've done before. It also controls our decisions to persist at a specific work goal in the face of distractions and the press of other priorities. Finally, motivation leads us to invest more or less cognitive effort to enhance both the quality and quantity of our work performance. (p. 21)

The best evidence supports the notion that motivation is the result of three things: (1) our values (we are more likely to choose to start and persist at goals we value); (2) our confidence that we can succeed at specific tasks (we invest more mental effort when tasks are perceived as challenging but possible to achieve); and (3) our mood or emotional state (positive mood states increase the likelihood that we'll start and persist at tasks) (Clark, 1999a). It is important to realize that motivation does not directly influence learning because it energizes the use of effective learning plans and strategies. It energizes us to start, persist, and use adequate mental effort to apply learning methods. Successful learning always requires motivated effort and adequate prior knowledge. Without adequate prior knowledge, exceptionally high motivation will not produce learning, and vice versa.

Are Newer Media More Motivating?

While most people can remember a situation in which they were motivated by outstanding multimedia instruction, most can also remember a teacher or a book that was motivating. The large-scale reviews of media evaluation and research studies by Bernard and Sitzmann's teams both reached the conclusion that people are no more motivated when learning from new technology than they are by classroom instruction. This kind of finding suggests to many people that we should assign students to the setting they prefer if choices are available—or that we should develop and deliver instruction in the medium or mix of media that students prefer. It seems reasonable that if there are any motivational benefits to be gained from media, even for a minority of students, we should take advantage of their preferences. While this is a reasonable assumption, it turns out to be exactly the opposite of the evidence (Clark, 1983).

A landmark study by Salomon (1984) established that most people prefer media that they believe will make learning easier. Salomon asked people from Israel and the United States whether they would rather learn from books or television and then assigned them to either a printed or televised lesson on the same topic. Israelis preferred books, apparently because many of them had experienced an overwhelming amount of televised instruction and thought it was a ''difficult medium.'' Americans had experienced television as an entertainment medium and thought it would make learning easier than books. What happened is that Israelis learned more from the televised lesson than from print and Americans had the opposite result—they learned more from the print than from the televised lesson. Salomon measured the amount of mental effort learners in all conditions invested in learning and found that, in general, we work harder under the conditions that we feel are more challenging. Israelis invested more effort in television because they believed it was difficult, and Americans worked harder when given a printed lesson for the same reason. This study has been repeated many times with many different ages, nationalities, learning tasks, and media and the same result occurs (see, for example, a discussion by Schunk and Pajares, 2004).

It appears that the choice of media does not help either student learning or motivation to learn. In fact, it is likely that allowing students to select media they feel will make learning easier may actually cause many of them to loaf and learn less than if they feel challenged. Recent reviews of instructional research have indicated that learning is supported primarily by a limited set of powerful instructional methods.

What About Computer-Based Educational Games?

A number of studies and reviews of studies that examined the benefits of games have been conducted (for example, Chen & O'Neil, 2005; Gredler, 1996; Mayer, Mautone, & Prothero, 2002; Moreno & Mayer, 2005; O'Neil, Wainess, & Baker, 2005). All of the studies that have been published in reputable journals have reached a negative conclusion about learning from games. Apparently, people who play serious games often learn how to play the game and perhaps gain some factual knowledge related to the game—but there is no evidence in the existing studies that games teach anyone anything that could not be learned through some other, less expensive, and more effective instructional methods. Even more surprising is that there is no compelling evidence that games lead to greater motivation to learn than other instructional programs.

Chen and O'Neil (2005) and O'Neil, Wainess, and Baker (2005) located over four thousand articles published in peer-reviewed journals and found only nineteen studies in which either qualitative and/or quantitative data about learning or motivation from games had been assessed. Their analysis of the learning and transfer measures used in all nineteen studies concluded, ''Positive findings regarding the educational benefits of games . . . can be attributed to instructional design and not to games per se. Also . . . many studies claiming positive outcomes appear to be making unsupported claims for the media'' (O'Neil, Wainess, & Baker, pp. 461–462). Their use of the term ''instructional design'' was intended to highlight the occasional use of instructional methods such as providing examples, classification practice, and problem-solving routines. They conclude that all of the methods used in games have been used effectively in non-game instructional programs and are not unique to games.

We might expect a less conservative and more optimistic view from industry, government, or military sponsored surveys of gaming research because of the high level of investment in those sectors, especially the military. Military trainers in many countries have invested in serious games for training. Yet an excellent technical report by Hayes (2005) for the Air Force training command provides a particularly thorough review of the past forty years of research and reviews of research on instructional games and ''simulation games.'' He concludes, ''The research shows no instructional advantages of games over other instructional approaches . . . . The research does not allow us to conclude that games are more effective than other well designed instructional activities'' (p. 43). He makes the point that only poorly designed studies find learning benefits from games. In most cases, poor design implies that the learning benefit of a game is compared to not receiving any game instruction or engaging in a non-educational exercise. What, he asks, can you conclude about the ''relative'' benefit of games when you do not compare them with any other way to teach or learn?

Chen and O'Neil (2005), O'Neil, Wainess, and Baker (2005), and Hayes (2005) all suggest that most studies that report motivation benefits from games only ask students whether they were motivated—they do not provide any direct measures of motivation (such as increased persistence or mental effort). Student opinions about motivation have been found to be highly unreliable and often in conflict with performance-based measures when both are gathered (see, for example, a recent comprehensive review of reaction measures by Sitzmann, Brown, Casper, Ely, & Zimmerman, 2008). Chen and O'Neil (2005) also note that many games appear to employ unguided, discovery, constructivist, or problem-based learning pedagogy. Since this approach to instruction is also included in many of the applications of newer media, next is a review of discovery approaches to instruction.

THE FAILURE OF DISCOVERY-BASED INSTRUCTION

One of the most popular approaches to instruction found in both schools and industry is based on the assumption that students will learn best if they are given a problem to solve or a task to perform and asked to work alone or to work collaboratively with a team to discover a solution. Those who use this discovery approach assume that the best learning occurs when people discover their own solutions to a problem or task. Discovery learning can be provided in almost any instructional medium. It is often a key element in a computer-based course and is the essential pedagogical element in nearly all simulations and instructional games (Clark, 2007). Teachers in classrooms and trainers in work settings often use discovery to help students learn. Discovery is a defining element of many different approaches to instructional design including constructivism (Duffy & Jonassen, 1992); communities of practice (Brown & Duguid, 1991); problem-based learning (Schwartz & Bransford, 1998); inquiry learning (Kuhn, Black, Keselman, & Kaplan, 2000); collaborative learning (van der Linden, Erkens, Schmidt, & Renshaw, 2005); scaffolding (Pea, 2004), and discovery (Shulman & Keisler, 1966). Its origin probably extends back to work by Jerome Bruner (1961), who used early 1900s Piagetian child development theory (see Piaget, 1928) to support discovery learning.

Evidence About Discovery Learning

Recent reviews of the research and evaluation evidence for discovery approaches by Mayer (2004) and Kirschner, Sweller, and Clark (2006) have provided compelling evidence that discovery is almost always less effective than giving students a guided solution in a demonstration based on task analysis and accompanied by practice and feedback. Mayer (2004) reviewed the past fifty years of research and found consistent evidence against discovery and in favor of guided instruction for all ages, all tasks, and all contexts. Kirschner, Sweller, and Clark (2006) reached the same conclusion as Mayer and focused their examples on the teaching of mathematics in schools and the education of physicians. The difficulty with asking people who are learning complex knowledge to discover all or part of the knowledge they are learning is that the discovery process requires a huge amount of unproductive mental effort. Even if a minority of learners succeed and discover what they need to know, the discovery process does not teach them how to discover, and the effort required could be invested in more efficient learning from demonstrations and practice exercises. Problem solving during learning is desirable, but discovering how to solve a problem in order to learn to solve a problem is not helpful or desirable. If all serious reviews of the evidence about discovery learning have been negative for the past fifty years (see, for example, a very early and very negative research-based critique by Shulman & Keisler, 1966), how is it possible that discovery is our most popular instructional method in schools and industry?

Why Is Discovery Learning So Popular?

Most of us remember vividly those times when we have experienced an important insight after investing effort to solve a problem. In fact, nearly all novel insights throughout the history of human beings have occurred as a result of discoveries. Most psychologists agree that learning is a process whereby people construct new knowledge by adding to what they already know about a topic. Because we are all unique, the result of the knowledge construction process is somewhat different for every individual. Discovering and constructing knowledge are common experiences for all of us. Thus, it seems intuitively correct to assume that, in order to learn, we have to allow people to construct their own knowledge. Yet because learning requires construction, it does not follow that construction or discovery is the most effective or efficient way to instruct or to help people learn. In fact, the evidence best supports the claim that we are born with a mental architecture that makes learning by discovering or constructing knowledge almost impossible for complex tasks. This applies to all but a small minority of the most able and knowledgeable learners (Kirschner, Sweller, & Clark, 2006).

Our Mental Architecture Resists Discovery Learning

In many ways it is not surprising that our most common assumptions and practice in instruction are at odds with the evidence about what works. The approaches we are using now were developed in the past century when our understanding of the architecture of the mind was based on a ''black box'' metaphor. Since we could not directly measure mental reactions to different instructional techniques, we had to make assumptions about what would work from our own experience and the way that learners reacted to different methods of teaching. With the development of direct measures of mental processes using the technology of neuroscience and more sophisticated measures of learning and performance, we are in a transition period to a new instructional psychology. We have achieved a number of new insights about the structure of our minds that will eventually change our approach to instruction. One compelling aspect of these new insights that is relevant to multimedia instruction is called ''Cognitive Load Theory.''

COGNITIVE LOAD THEORY AND MULTIMEDIA DESIGN

Sweller (2007) provides a compelling case for the fact that our minds have evolved to make new learning difficult in order to protect us from quick, extensive, and radical changes in thinking and behavior that might threaten our lives. Because all learning is novel, it is potentially as harmful as it is beneficial. The information-processing system that protects us from learning rapidly has many features. On the one hand, our minds simply will not permit us to think about more than about three to four new things at once. This information-processing limit is a major speed bump since all learning is novel and subject to this limitation. This new estimate of mental processing limit by Cowan (2001) reduces by more than 50 percent our former estimate (Miller, 1956) of a seven- to nine-chunk thinking capacity. This much lower capacity estimate is further reduced by learner anxiety. Not only do these limitations on thinking slow down our learning, but if we try to exceed this three- to four-chunk information processing limit, a processing routine shuts down our conscious minds in the same way a fuse disconnects an overloaded electrical circuit. When we are disconnected, our focus tends to switch to daydreaming (see, for example, a review of the research and a more complete description of this process by Clark, 1999b). Anyone who has attempted to study a difficult chapter in a textbook after reading for ten to twenty minutes or more realizes that he or she cannot remember much of anything that's been read. This describes the phenomenon of cognitive overload.

Another recent insight from extensive research on cognitive load theory (Mayer, 2004; Sweller, 2007) about the destructive power of common features of multimedia instruction raises an even larger cause for concern. Mayer (2001; 2005) has identified and studied the most common multimedia screen and instructional design strategies that overload learners mentally and cause learning problems. In nearly all cases, overload is caused by providing students with information in any form that distracts them from processing the essential conceptual or procedural knowledge required to perform the task they are learning. Since we all have a limited capacity to think when learning, we must use our thinking capacity to process relevant information. When instruction provides distractions such as music, animated agents who give us advice, tabs that allow us to find additional information, pages of text to read on the screen, and key information embedded in irrelevant contextual information, we must spend effort ignoring the irrelevant to select and learn the relevant information (Clark & Choi, 2007). Mayer (2001) identifies a number of multimedia design principles that, if implemented, tend to help us avoid cognitive overload and help learning (see Table 8.1).

Table 8.1 Mayer's (2001) Multimedia Design Principles

Principle	*Guideline*
Multimedia	Students learn better from words and pictures than from words alone.
Spatial Contiguity	Students learn better when corresponding words and pictures are presented near rather than far from each other on the page or screen.
Temporal Contiguity	Students learn better when corresponding words and pictures are presented simultaneously, rather than successively.
Coherence	Students learn better when extraneous words, pictures, and sounds are excluded rather than included.
Modality	Students learn better from animation and narration than from animation and on-screen text.
Redundancy	Students learn better from animation and narration than from animation, narration, and on-screen text.

(*Continued*)

Table 8.1 (*Continued*)

Principle	*Guideline*
Individual Differences	Design effects are stronger for low-knowledge learners than for high-knowledge learners and for high-spatial learners rather than for low-spatial learners.
Signaling	Students learn better when cues (underlining, arrows) are added that highlight the main ideas and organization of the words.
Pacing	Students learn better when they control pacing of segmented narrated animations rather than continuous pace.
Concepts First	Students learn better when new terms are learned before introducing complex processes, principles, or procedures.
Personalization	Students learn better when narration is conversational and uses personal pronouns such as ''you'' and ''yours.''
Human Voice	Students learn better when a human voice is used for narration, rather than a machine voice or foreign accented voice.

Mayer's principles apply to what most people call screen design (for computer or multimedia-based training), or graphic design for the printed page. They are all intended to focus people's attention on only relevant portions of instruction and not to distract them with irrelevant and dysfunctional depictions of information. Yet much more is necessary to support learning than simply helping learners avoid distracting elements in instruction. The additional learning support elements are most often called ''instructional principles'' and related instructional methods or strategies. What follows is an examination of the most powerful methods available that can be included in nearly any instructional design or delivery system.

POWERFUL INSTRUCTIONAL METHODS

Instruction has at least two major components—the first is to provide information that we want people to learn (instructional content) and the second is to structure that information in order to help students learn it (instructional methods) without being distracted by instructional displays (screen or graphic design). Simply giving people information is not instruction—or as Stolovitch and Keeps make clear, *Telling Ain't Training* (2002). Instruction can fail because the information we give students is incomplete or inaccurate and/or because we do not provide adequate learning support. Complicating the search for the best instructional methods is the fact that there are hundreds of different descriptions of instructional methods that have been recommended in the past century (see, for example, discussions by Cronbach & Snow, 1977; Kirschner, Sweller, & Clark, 2006; and Pea, 2004).

A Fragmented Instructional Design Community

What makes the search for the most important instructional design principles and related instructional methods challenging are both the variety one finds in use and the intense advocacy of people who market their own design systems. Advocates of different instructional theories and models tend to define and operationalize instructional support in very different ways. These different theories often spring from different models of learning and sometimes different belief systems, inquiry methods, and philosophies (Cronbach & Snow, 1977; Jonassen, 1991; Merrill, 2002; Romiszowski, 2006). To some extent, these differences reflect the increased specialization and fragmentation in educational research and theory over the past half-century (Ravitch & Viteritti, 2001; Winthrop, 1963) and a growing fragmentation among various sub-specializations in educational research. One result of this phenomenon is that researchers who favor a specific theory or point of view tend to isolate themselves and limit their research, while reading and collaborating with only the journals and professional associations or divisions of associations that emphasize their perspectives. For example, it is our experience that the groups who educate classroom teachers for K–12 education and the large number of faculty who teach in our universities tend to ignore the instructional theory and instructional design models generated by people who develop instructional programs and media for K–12, university, and industry education. Business and military trainers are more open to instructional design, but tend to confuse media with instructional method. Even various instructional design models tend to attract advocates who do not examine alternatives or look carefully at evidence that does not support the model they advocate. Encouraging dialogues between these diverse groups and individuals who are concerned with instruction and learning will help bridge this gap in the future.

David Merrill's Five-Star System

David Merrill (2002; 2006) has bucked the fragmentation trend by selecting the most powerful instructional principles from among the many instructional design systems that are available. In a study supported by the American Society for Training and Development, he identified a number of current systems that could present strong research and evaluation evidence to support their effectiveness. He was curious about the instructional methods that reflect each of the principles that are the common foundation of the design systems he located. After an extensive and systematic analysis, he identified five instructional principles and the related instructional methods that all of these powerful systems use. He then looked broadly at the research on instructional methods

(Merrill, 2006) and confirmed that each of these five received strong validation in both laboratory and field-based experiments.

Five Instructional Design Principles. The five principles Merrill (2002) found in the most powerful design systems suggest that ''learning is promoted when the learner: (1) observes a demonstration—demonstration principle; (2) applies the new knowledge—application principle; (3) undertakes real-world tasks—task-centered principle; (4) activates existing knowledge—activation principle; and (5) integrates the new knowledge into their world—integration principle'' (Merrill, 2006, p. 262). Merrill describes a situation in which the corporate training group NETg revised one of its own Excel spreadsheet application courses to reflect these five principles (Thompson Inc., 2002). When they compared the new and old versions of the application, they found huge differences in performance favoring the five principles. The students in the five principles course completed complex spreadsheet tasks more quickly (twenty-nine versus forty-nine minutes) and effectively (89 percent versus 68 percent performance improvement) than students in their standard course. ''Effectiveness was measured by the learners' ability to complete three complex tasks that required them to develop three different spreadsheets given a set of data and analysis requirements'' (Merrill, 2006, p. 264). He also describes a study reported in a conference paper by Barclay, Gur, and Wu (2004), who analyzed the extent to which these principles were included in the hundreds of training courses presented on over fourteen hundred websites in five different nations (Australia, China, France, Turkey, and the United States). The best training courses in each of these countries implemented only half of the principles, and the national average for all courses in each nation was about one principle per course.

From Instructional Principles to Instructional Methods

Principles predict what will happen if a type of instruction is offered to students. Instructional methods are strategies for implementing the principles. Nearly all principles and methods can be implemented and delivered in nearly all media. One way to translate Merrill's (2002) five principles into instructional methods is as follows:

1. Provide realistic field-based problems for students to solve;
2. Give students analogies and examples that relate their relevant prior knowledge to new learning;
3. Offer clear and complete demonstrations of how to perform key tasks and solve authentic problems;
4. Insist on frequent practice opportunities during training to apply what is being learned (by performing tasks and solving problems) while receiving corrective feedback; and

5. Require application practice that includes ''part task'' (practicing small chunks of larger tasks) and also ''whole tasks'' (applying as much of what is learned as possible to solve the complex problems that represent challenges encountered in operational environments) both during and after instruction.

Merrill's instructional methods and the principles upon which they are based can be integrated into nearly any instructional design system or used to construct a new design system tailored for the needs of specific organizations and/or groups of students. (See Clark and Estes, 2008, for an approach to developing tailored applications of Merrill's principles to different cultural contexts.)

Although comprehensive, Merrill's (2002) principles do not provide a complete model for designing at the lesson level. Moreover, instructional methods alone do not give guidance about what knowledge and skills have to be taught to achieve performance goals. To determine instructional content requires the use of task analysis methods to identify the conceptual and procedural knowledge necessary to perform a task (Jonassen, Tessmer, & Hannum, 1999). As tasks become more complex, they require the use of both controlled (conscious, conceptual) and automated (unconscious, procedural, or strategic) knowledge over an extended period of time (Clark & Elen, 2006; van Merriënböer, Clark, & de Croock, 2002). Thus, a valuable approach to task analysis is to capture both the observable actions and the underlying ''cognitive'' knowledge experts use to successfully and consistently perform a complex task (Clark & Estes, 1996). Optimized instructional design methods, then, should integrate both cognitive task analysis (CTA) and Merrill's five principles. Currently, at least five major instructional design systems take this approach: (1) The Integrated Task Analysis Model (ITAM; Redding, 1995; Ryder & Redding, 1993); (2) the Ten Steps to Complex Learning systematic approach to four-component instructional design (van Merriënböer, Clark, & de Croock, 2002; van Merriënböer & Kirschner, 2007); 3) Task-Centered Instructional Strategy (Merrill, 2007); (4) e-Learning and the Science of Instruction model (Clark & Mayer, 2007); and (5) Guided Experiential Learning (GEL; Clark, 2004, 2006). To more fully illustrate how these integrated instructional design models work, we describe GEL in the next section.

Guided Experiential Learning

Clark (2004, 2006) describes one possible list of the instructional methods that most evidence-based instructional design systems use at the lesson level in a design system called ''guided experiential learning'' (GEL). He specifies how Merrill's five principles could be combined with currently used training methods by requiring that all lessons include the following elements in the following sequence:

1. Objectives (specify actions, conditions, and standards that must be achieved in a lesson);

2. Reasons for learning (advantages of learning and risks of failure to learn and transfer);
3. Overview (knowledge models and content outline);
4. Conceptual knowledge (concepts, processes, and principles necessary to learn to perform a task or solve a problem, with examples and analogies that support learning);
5. Demonstration of the procedure (a clear ''how to'' description for all elements of a task or solution);
6. Part and whole-task practice of procedures with corrective feedback; and
7. Challenging, competency-based tests that include reactions (trainee confidence and value for the learning) and learning performance (memory for conceptual knowledge and application skill for all procedures).

Cognitive and Behavioral Task Analysis. When subject-matter experts are available, implementing Merrill's (2002) five principles and the GEL design at the lesson level also requires the use of cognitive task analysis (CTA) to determine training information content (Clark, 2004; 2006). CTA methods capture accurate descriptions of the performance objectives, equipment, conceptual and procedural knowledge, and performance standards that experts use to perform complex tasks (Clark, Feldon, van Merriënböer, Yates, & Early, 2007). CTA can be viewed as extending, not replacing, behavioral task analysis. CTA not only records observable activities, but it also seeks to capture the unobservable cognitive processes that underlie expert task performance. The results of CTA provide the instructional content that ''populates'' each instructional element in the GEL system.

Why CTA Is Important. Experts are often called upon to provide their knowledge and skills for training development and delivery. Behavioral task analysis has historically served as the primary approach to capturing experts' observable actions for these purposes. However, replicating expert performance that originates from behavioral task analysis alone is problematic, especially for complex tasks involving unobservable cognitive activities, such as analysis, judgments, and decisions (Yates, 2007). Experts achieve high performance in a domain as a result of continuous and deliberate practice in solving problems over a long period of time (Ericsson, Krampe, & Tesch-Römer, 1993). Through practice, experts' new knowledge and skills become automated and unconscious (Anderson & Lebiere, 1998) to an extent that perhaps up to 90 percent of all knowledge is unconscious (Wegner, 2002). As a result, when called upon to describe how they achieve their high performance levels, experts are often unable to completely and accurately recall the knowledge and skills they use, often resulting in significant omissions that can negatively impact the effectiveness of instruction and lead to subsequent difficulties for learners (Chao & Salvendy, 1994; Feldon, 2007; Hinds, 1999). Studies have shown that training based only on expert self-report information lacks approximately 70 percent of the information necessary for training and performance (Feldon & Clark, 2006).

Training based on the results of cognitive task analysis methods, on the other hand, has been shown to be substantially more effective than training developed through other means. In Merrill's (2002) study comparing discovery learning with direct instruction in the use of spreadsheets, the direct instruction training was based on strategies elicited from an expert spreadsheet user. In a problem-solving task, learners in the CTA-based instruction group scored 89 percent versus 34 percent for the discovery group. The CTA-based group also required less time to complete the task with an average of twenty-nine minutes versus an average of more than sixty minutes for the discovery group.

CTA-based instruction has also been shown to be more effective for training in troubleshooting. Schaafstal, Schraagen, and van Berlo (2000) compared a preexisting training course in radar system troubleshooting with a newly designed course based on content elicited using CTA methods. Although participants had equal scores on pre-tests of basic knowledge, participants in the CTA-based training were able to identify more than twice the number of malfunctions, and in less time, than the participants in the traditionally trained group. In addition, these results were replicated in all subsequent implementations of the CTA-based training designs.

Evidence from the use of CTA in various areas of medicine indicates important implications for medical training as well as the treatment of patients. For example, in a CTA study of medical school surgical instruction (Velmahos, Toutouzas, Sillin, Chan, Clark, Theodorou, & Maupin, 2004), researchers found that, when medical professors taught medical students to perform surgery, the professors tended to accurately describe their own visible actions but consistently omitted most of the key decisions they made when describing their approach to a surgery. In this study, information captured from CTA interviews with expert surgeons was used to train half of the annual surgical residents in a large urban teaching hospital, and the other half of the surgical residents experienced a traditional ''see one, do one, teach one'' pedagogy (Halsted, 1904). The experts who taught the traditional group were the same experts interviewed for the CTA. In the year following the training, senior surgeons observed the surgical residents whenever they performed the task without knowing how they'd been trained. Results indicated that the residents who received the CTA-based description of the surgical procedure made about 60 to 70 percent better decisions with patients than those who only observed the procedure and heard expert surgeon explanations. CTA trained students were more accurate about where to perform the procedure, what instruments to choose when patients were seriously injured, and what to do when a step did not have an intended outcome. As a result, the surgeons who experienced the CTA-based training made no serious errors when using the procedure with patients, whereas the experimental group made a number of damaging decision errors (although not more errors than had been typical in the past for this procedure). Similar results in studies of the diagnostic expertise of top neonatal nurses have been reported by Crandall and Gretchell-Leiter (1993), who described a similar study wherein CTA of expert neonatal nurses exposed a strategy for diagnosing life-threatening infections in premature infants that was significantly more effective than the textbook method taught in universities.

To determine the generalizability of the effectiveness of CTA-based instruction, Lee (2004) conducted a meta-analytic review of the training literature. Meta-analysis is a statistical method of aggregating and comparing the findings of different research studies within a common topic (Lipsey & Wilson, 2001). Her search for training studies based on CTA methods resulted in thirty-nine comparisons of the average effect size differences between pre- and post-test measures of training performance. Lee reported an overall average post-training performance gain of about 53 percent ($d = 1.72$) for CTA training when compared to more traditional training design using expert-based task analysis.

How Is CTA Conducted? CTA refers to a variety of interview and observation techniques used to elicit and represent the knowledge, goals, strategies, and decisions that underlie observable task performance. Although there are many types of CTA methods (for a review, see Clark, Feldon, van Merriënböer, Yates, & Early, 2007), CTA methods share a common goal of capturing the knowledge of subject-matter experts (SMEs) who have demonstrated consistent proficiency in performing a task over a long period of time. CTA is most commonly performed in five stages:

1. Identify the tasks to be analyzed and acquire general knowledge of the domain in which the tasks are performed.
2. Identify the types of knowledge required to perform the tasks and subtasks.
3. Elicit the knowledge required to perform the tasks, using multiple SMEs.
4. Analyze and format the elicited knowledge and verify for accuracy and completeness by reviewing transcripts and cross-checking with multiple SMEs.
5. Format the knowledge for its intended application (for example, procedures that include action and decision steps, general strategies or rules of thumb, and job aids).

The GEL design system incorporates a CTA method that includes these five stages. After identifying the required knowledge types and becoming familiar with the area of interest, three multiple subject-matter experts are interviewed followed by cycles of expert self- and peer-review to capture the automated and unconscious knowledge acquired through experience and practice. The initial, semi-structured interview begins with a description of the CTA process by the analyst. The SME is then asked to list or outline the performance sequence of all key subtasks necessary to perform the larger task being examined. SMEs are also asked to describe (or help the interviewer locate) at least five authentic problems that an expert should be able to solve if he or she has mastered the task, the benefits of solving the problem, and the risks of not being able to solve the problem. Problems should range from routine to highly complex whenever possible. The resulting sequence of tasks becomes the outline for the training to be designed or the job description produced after the CTA is completed. Starting with the first subtask in the sequence, the analyst asks a series of questions to collect:

1. The sequence of actions (or steps) necessary to complete the task and all subtasks;
2. The decisions that have to be made to complete the subtask, when each must be made, the alternatives to consider, and the criteria to decide between the alternatives;
3. All concepts, processes, and principles that are the conceptual basis for the experts' approach to the subtask;
4. The conditions or initiating events that must occur to start the correct procedure;
5. The equipment and materials required;
6. The sensory experiences required (for example, the analyst asks whether the expert must smell, taste, or touch something in addition to seeing or hearing cues in order to perform each subtask), and
7. The performance standards required, such as speed, accuracy, or quality indicators.

The interview is repeated for each SME, with each recorded and transcribed verbatim. The transcripts are then analyzed to generate consistently formatted protocols containing the results of each interview. After each SME has corrected his or her own protocol, they are then exchanged with the other SMEs for verification and correction. An aggregated ''gold standard'' protocol is produced and submitted to each SME for final approval.

How Are the Results of CTA Used in the GEL System? The GEL design technology combines the five training principles that Merrill (2002) identified as the active ingredients of the most effective, evidence-based pedagogical systems currently in use with CTA methods that effectively capture the knowledge and skills that underlie expert task performance. Table 8.2 shows how the results of CTA provide the content for a GEL designed course and each of the seven elements within a GEL lesson.

Table 8.2 Incorporating CTA Results for GEL Lesson Design

CTA Result	*GEL Lesson Element*
Course Design	
Five large authentic field problems, ranging from easy to complex, that illustrate the performance of the task	Sequence groups of problems into lessons: first performed in the field are first taught; if no fixed sequence, then easy to difficult.
Conceptual knowledge about field problems: new concepts (definitions and examples); processes (how it works—big picture); principles (what causes something to happen); procedures (how to do it, conditions and consequences)	Identify prior knowledge and pre-requisite knowledge that must be taught first.
Authentic field problems; action and decision steps; conceptual knowledge; standards	Challenging, competency-based assessment—provide whole-task tests that include reactions (trainee confidence and value for the learning) and learning (memory for conceptual knowledge and application skill for all procedures).
Lesson Design	
Conditions or initiating events; performance standards; equipment and materials; sensory experiences	1. Objectives—specify the actions, conditions, and standards that must be achieved.
Authentic field problems; benefits of solving; risks by not solving	2. Reasons for Learning—state the benefits of learning and the risks of failure to learn and transfer.
Authentic field problems; conceptual knowledge	3. Overview—relate the knowledge to be learned to learners' prior knowledge; provide the position of the lesson in the overall course.
New conceptual knowledge for each lesson: concepts, processes, principles	4. Conceptual Knowledge—provide knowledge necessary to perform the task or solve a problem with examples and analogies that support learning.

Authentic field problem; action steps and decision steps with alternatives and criteria for deciding	5. Demonstration—provide a ''how to'' demonstration of the procedure for all elements of a task or solution using an authentic field problem.
Authentic field problems; action and decision steps	6. Part- and Whole-Task Practice—provide opportunity to practice procedures using authentic field problems and provide corrective feedback.
Authentic field problems; action and decision steps; conceptual knowledge; standards	7. Challenging, Competency-Based Assessment—provide part-task tests that include reactions (trainee confidence and value for the learning) and learning (memory for conceptual knowledge and application skill for all procedures).

How Does GEL Develop Flexible Expertise and Learning Transfer? The purpose of learning is to develop flexible skill expertise and to transfer that expertise to solving novel problems. After an extensive review of the transfer literature, Perkins and Grozner (1997) and Clark and Blake (1997) argue that flexibility can be taught in a way that facilitates the solution of novel and challenging problems. They describe strategies that have been used in successful programs. De Corte (2003) and Masui and De Corte (1999) draw on these reviews and others to provide a description of aspects of learning environments that facilitate the development of the necessary characteristics for successful transfer of existing skills to novel problems in which orienting (problem framing) and self-judging were taught according to the following guidelines:

- *Environment*: Skills and knowledge instruction must be taught in environments that reflect the application environment as much as possible to highlight the importance of relevant cues.
- *Motivation*: Task motivation must be linked to tangible and personally relevant outcomes.
- *Increasing novelty:* Training must be sequenced to allow for gradually increasing levels of novelty and challenge (see also extensive research on the design of instruction using worked examples: Atkinson, Derry, Renkl, & Wortham, 2000; Paas & van Merriënböer, 1993; Sweller, 1999).
- *Variable practice*: The characteristics of learning and performance tasks must be variable over the course of instruction to maximize opportunities to develop flexibility.

- *Targeted feedback*: Students must be provided with opportunities to receive targeted feedback and consider alternatives to more effective approaches.

These guidelines reflect a similar list suggested by Merrill (2006), who analyzed the key features of new training design systems that appeared to be successful at developing flexible expertise and recommended similar design features.

Based on the De Corte (2003) and Merrill (2002) design criteria and the studies cited above, the GEL design system attempts to promote the development of flexible expertise through applying all of the empirically identified training methods that promote flexibility:

- *Environment*: Where possible, GEL lessons are situated in the environment in which skills and knowledge will be applied. Environment is reflected in a series of application scenarios (similar to case studies) and demonstration videos. GEL also attempts to prevent cognitive overload by focusing novice trainees on only the key elements of an application environment.
- *Motivation*: GEL requires motivating statements of tangible and personally relevant ''benefits and risks'' associated with each task to be learned.
- *Increasing novelty*: GEL requires the collection of five increasingly novel and challenging scenarios (similar to case studies or authentic problems) for use in practice exercises, checks on learning, and testing. The variation in novelty for a GEL course is greater than any other design system.
- *Variable practice*: GEL requires both part-task practice (during lessons) and whole-task practice wherein trainees are required to apply what they have learned as they attempt to solve the problems and scenarios described in the point above.
- *Targeted feedback*: GEL requires targeted feedback on trainees' attempts to apply what they have learned from demonstrations and attempts to practice when given scenarios and problems. GEL feedback strategies draw on the most current research on feedback and performance to support flexible expertise.

Additional features of GEL designed to promote flexible expertise:

- *Analogical connections to prior knowledge*: GEL requires the presentation of analogies and varied examples in each lesson in order to help trainees connect to prior knowledge and to promote flexible application of skills and knowledge. The strategy reflects research by, for example, Gentner, Lowenstein, and Thompson (2003).

- *Open questions during feedback*: When application practice feedback is given, trainees are asked for their reasoning about their problem-solving strategies and are given the opportunity to examine alternatives rather than being ''told the correct path.''

In summary, we have described a powerful instructional design system that translates Merrill's (2002) five principles to instructional methods and specifies a sequence of cognitive events for the course and each lesson. We have also described a method to capture the knowledge and skills that subject-matter experts use to solve complex problems in the training domain. With the instructional design and content in place, we now turn the discussion to describing a powerful approach for selecting the optimal delivery media, based on supporting the cognitive processes necessary for meaningful learning within the GEL system.

A TWO-STAGE COGNITIVE APPROACH TO MEDIA SELECTION

Sugrue and Clark (2000) provide an in-depth analysis of the use of media, media attributes, and instructional methods as part of their comprehensive approach to media selection for training. They suggest that the difficulty in choosing among media options, either prior to or after instructional design, stems from the confusion between media and methods and between media and media attributes. Each has a role to play in media selection; however, only instructional methods are directly related to the cognitive processes involved in learning, whereas the choice of media has a direct link to cost, access, or time to learn.

To begin the discussion, we return to the definition of an instructional method as an external (environmental) activity that supports internal cognitive process necessary for meaningful learning (Clark, 1983, 1994). The degree to which a method provides a level of support varies according to the amount of intrinsic cognitive load on working memory imposed by the instructional content. A method is further defined according to the cognitive process it supports. For example, presenting a learner with reasons for the training, the benefits of learning a task, and the risks of poor performance support the cognitive process of goal elaboration. Similarly, demonstrating how to perform the steps in a task supports the cognitive process of compiling procedures by presenting examples (Anderson, 1993; Anderson, Bothell, Byrne, Douglass, Lebiere, & Qin, 2004; Anderson & Fincham, 1994).

Sugrue and Clark (2000) refer to Levie's (1989) definition of a media attribute as the specific feature of a medium that provides the functionality of transmitting information to trainees or cognitive processing responses from trainees. Examples include functions that transmit audio and video, display text, provide searchable access to information, or give feedback during practice. A more

Table 8.3 Outcomes of Instruction Influenced by Media and Instructional Methods

	Access	*Cost (Development and Delivery)*	*Efficiency (Time to Learn)*	*Learning and Motivation*
Media	X	X	X	
Methods			X	X

From Clark, Bewley, and O'Neil, 2006. Reprinted with permission.

specific example would be the ''zooming'' attribute of real-time video camera lenses. This attribute allows a designer to ''zoom in or out'' and so to visually select a small aspect of a visual or pull back to a wide shot of a complex visual. A medium, then, is defined as an external resource that contains media attributes or capabilities (Kozma, 1991). In short, a medium's attributes enable the delivery of methods that have cognitive consequences and, therefore, the best approach to selecting media is based on its ability to perform instructional functions relative to other media. Sugrue and Clark's (2000) analysis of the instructional influence of media and methods is summarized in Table 8.3.

Sugrue and Clark (2000) propose that media selection begins first with the selection of instructional methods that support the cognitive processes necessary to perform the task to be trained and then continues with an analysis of media based on their ability to provide the type of method, amount (hint or provide), timing (now or later) and control (learner or media) of methods. Final media selection is based on the most economical, assessible, and cost-efficient media that incorporate the required attributes.

In their discussion of media selection issues, Clark, Bewley, and O'Neil (2006) recommend following Sugrue and Clark's (2000) cognitive process for determining instructional methods. However, in their examination of media attributes, they found that ''three of the most common instructional methods can be only presented via a limited number of media'' (p. 136) based on the methods requirements for (1) sensory information; (2) conditional knowledge; and (3) synchronous feedback.

As a result, we propose a two-stage process for media selection that incorporates both approaches.

Selecting Instructional Methods. Sugrue and Clark (2000) propose a model of instructional methods that includes six categories. The model is based on Glaser's (1992) model of the cognitive components of expertise, Anderson's (1993) theory of learning, and theories of the components of self-regulated learning (Corno & Mandinach, 1983; Flavell, 1979; McCombs, 1988; Pintrich & DeGroot, 1990; Salomon, 1984) and is consistent with Merrill's (2002) five principles and the GEL (Clark, 2004, 2006) system. Each of the categories, the

Table 8.4 Instructional Methods and GEL Components

Instructional Method	*Cognitive Process*	*GEL Component*
Goal Elaboration	Explains the goal and its demands	Objectives, Reasons, Overview
Information	Provides task-related information	Conceptual knowledge, Demonstrations
Practice	Provides opportunities in varied contexts	Practice
Monitoring	Observes performance	Feedback
Diagnosis	Identifies causes of error	Feedback
Adaptation	Modifies goal, information, and practice	Assessments, Demonstration

Based on material in Sugrue and Clark (2000) and Clark (2004, 2006).

cognitive process they support, and a mapping with the instructional components of the GEL system are listed in Table 8.4.

Thus, within the GEL design system, the question to be answered during the first stage of media selection is: What type, amount, timing, and control of objectives, reasons, overview, demonstration, practice, feedback, and assessment methods must be provided? Based on the responses to this question, Table 8.5 provides a procedure for selecting instructional methods.

Table 8.5 Procedure for Instructional Method Selection

Instructional Method Selection Procedure
Step 1: Select type of goal elaboration (description and/or demonstration)
Step 2: Select type of information (description and/or demonstration)
Step 3: Select type of practice (high- and/or low-contextual authenticity)
Step 4: Select type of support for monitoring (data collection or guidance)
Step 5: Select type of support for diagnosis (analysis or guidance)
Step 6: Select type of adaptation (goal/information/practice; or guidance)
Step 7: Select amount of each method (low or high; fixed or variable)
Step 8: Select timing of each method (fixed or variable; immediate or delayed)
Step 9: Select locus of control for each method (system, trainee, or shared)

Based on Sugrue and Clark, 2000

Selecting Media. The second stage of the media selection process is choosing the media that best provide the selected type, amount, timing, and control of methods that support the cognitive processing necessary for learning.

Clark, Bewley, and O'Neil (2006) state that the three instructional methods that often limit instructional media selection are (1) ''the sensory modes required for learning concepts, processes, and procedures; (2) conditional knowledge requirements for the use of learned information; and (3) the need for synchronous feedback when complex knowledge is being learned'' (p. 136).

Some training may require sensory information beyond the visual and aural senses. Firefighting, for example, relies heavily on smell and tactile modes. Currently, electronic media can only provide visual and aural information; therefore, any part of the training that requires smelling, tasting, or touching something must be conducted ''in person.'' For the media selection process, then, the guideline is that if any sensory-based information is absolutely necessary to learn concepts, processes, and procedures, then it must be presented during that particular part of the training (Clark, Bewley, & O'Neil, 2006).

Conditional knowledge about when and where to perform a task must be depicted during training. One way of thinking about conditional knowledge is that it represents the first part of an ''if–then'' statement. A manufacturing example might be: If an order is received, then follow the procedure for processing payments and shipping the product. Some conditions are more complex and require greater authenticity, such as a fire, an urban setting, or a confrontation with people. As a guideline, the media selected must adequately depict the conditions required for learners to apply the new training (Clark, Bewley, & O'Neil, 2006).

Synchronous feedback refers to observation and corrective feedback provided by a live ''expert coach'' when trainees engage in complex practice exercises (Clark, Bewley, & O'Neil, 2006). Complex knowledge is defined as ''requiring the integration and coordinated performance of task-specific constituent skills rather than merely recalling definitions and other conceptual knowledge about concepts, processes, and principles'' (p. 137). In other words, complex knowledge is more than the sum of its parts, so trainees cannot practice each part and then be expected to perform the whole task successfully. Whole task practice is required to integrate and coordinate all parts of a task (see van Merriënböer, 1997, and van Merriënböer, Clark, & de Crook, 2002, for a complete discussion). As a guideline, if complex knowledge is the focus of the training, the media selected for complex practice exercises must support synchronous feedback for trainees through real-time observation and both verbal and visual feedback by a coach.

The objective of the media selection process is to determine the most cost-beneficial media delivery platform for effective training and education. For the purposes of media selection, Clark and his associates (2006) classify media platforms as either classroom or distance, which includes multimedia transmitted over the Internet and/or recorded on CD-ROM or DVD. To apply the procedure for selecting media, they recommend first determining the training requirements with respect to sensory modes, conditional knowledge, and practice and feedback. It is also necessary to collect information about the learning objectives, the location and number of learners, and the cost of delivering the training on all possible platforms. With this information in hand, the procedural steps in Table 8.6 are followed to select the optimal media.

Table 8.6 Training Delivery Platform Selection Procedure

Steps	*Decisions and Actions*
1	Can both a distance and a classroom platform simulate all of the necessary conditions in the job setting where the learners will apply their skills and knowledge? If yes, go to Step 2. If the answer is no for any platform, select the platform that will provide the necessary conditions.
2	Can both platforms provide the required immediate (synchronous) and delayed (asynchronous) information and corrective feedback needed to achieve learning objectives? If yes, go to Step 3. If the answer is no for any platform, select the platform that will provide the necessary feedback.
3	Can both platforms provide the necessary sensory mode information (visual, aural, kinesthetic, olfactory, tactile) required to achieve all learning objectives? If the answer is no for any platform, select the platform that will provide the necessary sensory mode information.
4	If both distance and classroom platforms have survived as viable options, subject both to cost-per-student (Steps 4A and 4B) and (if desired) value-enhanced cost (Step 4C) analysis.
4A	Derive the cost of each platform by listing and summing the costs associated with a specific course. Derive two sums, one for distance delivery and one for classroom.
4B	Divide the projected cost of each platform by the number of learners to be trained to determine the cost-per-student of each platform. Either select the platform with the lowest cost per student or go on to Step 4C.
4C	To determine the value-enhanced cost for classroom or distance platforms, survey key stakeholders to determine their preference or value assigned for each platform. Subtract the percent of average value assigned to the preferred platform by the stakeholders from the cost-per-student of that platform to derive a value-enhanced cost.
5	Select the delivery platform option that survived Steps 1 through 3 and that has the lowest cost-per-student and/or lowest value-enhanced cost from Step 4.

Based on Clark, Bewley, and O'Neil, 2006

To arrive at the cost-per-student value, all direct and indirect costs of resources should be calculated for each platform version under consideration, including design, development, transmission, travel, and cost of materials. The total cost for each platform is then divided by the number of trainees scheduled to complete the course. For example, if a distance-delivered training is projected to cost \$450,000 and will be delivered to 6,500 trainees, then the cost-per-student is \$69.23.

In some instances, key stakeholders may place a value on particular media to deliver training based, for example, on the public relations value of using the latest technology or the ''face time'' with employees that live training provides. Clark and his associates (2006) define value-enhanced cost as ''the percent of value (relative strength of the preferences) stakeholders place on their preferred delivery platform above the value they place on their less preferred option multiplied by the cost-per-student'' (p. 140). For example, assume that the stakeholders prefer distance training to live training by 27 percent and the cost of the distance option is \$69.23 per student. The value-enhanced cost would be \$50.54 calculated as $\$69.23 \times .27 = \$18.69$ and $\$69.23 - 18.69 = \$50.54$. Thus, the value-enhanced cost provides a lower-cost advantage when comparing platforms. In short, the final media selection decision is based on a combination of the instructional methods, cost-benefit ratios, and stakeholder values.

CONCLUSION

In seemingly parallel paths, dramatic changes are occurring in the organizational climate for training, research in cognitive psychology, and the capabilities provided by multimedia technology. However, high expectations that the convergence of these advances would benefit business and education have not been realized. In training environments, research has shown that, for all but a few learners, popular discovery-based instructional methods, including those being used in simulations and ''serious games,'' are largely ineffective. Instructional content is largely drawn from incomplete and inaccurate subject-matter expert descriptions of learning tasks, rather than from capturing the unobservable decisions, judgments, and analysis they use to solve complex problems. And regardless of the evidence that media do not influence learning, most training delivery decisions are based primarily on media preferences, rather than supporting the cognitive processes involved in meaningful learning.

In this chapter, we have illustrated a complete system that integrates three components of evidence-based practice that have been demonstrated to result in successful training and education: (1) a cognitive task analysis process that captures nearly all of the knowledge and skills experts use to solve complex problems in a domain in a way that can be used by novices; (2) a guided experiential design process that integrates expert knowledge and skills with instructional methods that support learners' cognitive processing during learning and transfer of what is learned; and (3) a media selection process to achieve the most effective and efficient delivery of these instructional methods. As education and training stakeholders who hold diverse positions in all three areas engage in dialogs and implement integrated instructional systems, we will benefit not only from additional research data supporting their effectiveness, but also from the impact of highly educated and trained workforce ready to compete in the global marketplace.

References

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–1060.

Anderson, J. R., & Fincham, J. M. (1994). Acquisition of procedural skills from examples. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(6), 1322–1340.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum Associates.

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research*, *70*(2), 181–214.

Barclay, M. W., Gur, B., & Wu, C. (2004). The impact of media on the family: Assessing the availability and quality of instruction on the World Wide Web for enhancing marriage relationships. Paper presented at the World Congress of the Family: Asia Pacific Dialogue, Kuala, Malaysia.

Bernard, R., Abrami, P. L., Lou, Y., & Borokhovski, E. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research*, *74*, 379–439.

Brown, J. S., & Duguid, P. (1991). Organizational learning and communities of practice: Toward a unifying view of working, learning, and innovation. In M. D. Cohen & L. S. Sproull (Eds.), *Organizational learning* (pp. 59–82). London: Sage.

Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*, *31*(1), 21–32.

Chao, C.-J., & Salvendy, G. (1994). Percentage of procedural knowledge acquired as a function of the number of experts from whom knowledge is acquired for diagnosis, debugging and interpretation tasks. *International Journal of Human-Computer Interaction*, *6*, 221–233.

Chen, H.-H., & O'Neil, H. F. (April, 2005). Training effectiveness of a computer game. Paper presented in a symposium titled Research Issues in Learning Environments at the annual meeting of the American Educational Research Association, Montreal, Canada.

Clark, R. C., & Mayer, R. E. (2007). *e-Learning and the science of instruction*. San Francisco: Pfeiffer.

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research*, *53*(4), 445–459.

Clark, R. E. (1991). When researchers swim upstream: Reflections on an unpopular argument about learning from media. *Educational Technology*, *31*(2), 34–40.

Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development*, *42*(2), 21–29.

Clark, R. E. (1999a). The CANE model of motivation to learn and to work: A two-stage process of goal commitment and effort. In J. Lowyck (Ed.), *Trends in corporate training*. Leuven, Belgium: University of Leuven Press.

Clark, R. E. (1999b). Yin and yang cognitive motivational processes operating in multimedia learning environments. In J.J.G.van Merriënböer (Ed.), *Cognition and multimedia design* (pp. 73–107). Herleen, Netherlands: Open University Press.

Clark, R. E. (Ed.). (2001). *Learning from media: Arguments, analysis, and evidence*. Greenwich, CT: Information Age Publishing.

Clark, R. E. (2003). Fostering the work motivation of individuals and teams. *Performance Improvement*, *42*(3), 21–29.

Clark, R. E. (2004). *Design document for a guided experiential learning course*. Final report on contract DAAD 19-99-D-0046-0004 from TRADOC to the Institute for Creative Technologies and the Rossier School of Education.

Clark, R. E. (2006). *Training aid for cognitive task analysis*. Technical report produced under contract ICT 53-0821-0137—W911NF-04-D-0005 from the Institute for Creative Technologies to the Center for Cognitive Technology, University of Southern California.

Clark, R. E. (2007, May/June). Learning from serious games? Arguments, evidence, and research suggestions. *Educational Technology*, pp. 56–59.

Clark, R. E., Bewley, W. L., & O'Neil, H. (2006). Heuristics for selecting distance or classroom settings for courses. In H. O'Neil & R. Perez (Eds.), *Web-based learning: Theory, research and practice* (pp. 133–142). Mahwah, NJ: Lawrence Erlbaum Associates.

Clark R. E., & Blake, S. (1997). Analyzing cognitive structures and processes to derive instructional methods for the transfer of problem solving expertise. In S. Dijkstra & N. M. Seel (Eds.), *Instructional design perspectives. Volume II, Solving instructional design problems* (pp. 183–214). Oxford: Pergamon.

Clark, R. E., & Choi, S. (2007). The questionable benefits of pedagogical agents. *Journal of Educational Computing Research*, *36*(4), 379–381.

Clark, R. E., & Elen, J. (2006). When less is more: Research and theory insights about instruction for complex learning. In J. Elen & R. E. Clark (Eds.), *Handling* *complexity in learning environments: Research and theory* (pp. 283–297). Oxford: Elsevier Science Limited.

Clark, R. E., & Estes, F. (1996). Cognitive task analysis. *International Journal of Educational Research*, *25*, 403–417.

Clark, R. E., & Estes, F. (2008). *Turning research into results: A guide to selecting the right performance solutions*. Charlotte, NC: Information Age Publishers.

Clark, R. E., Feldon, D., van Merriënböer, J.J.G., Yates, K., & Early, S. (2007). Cognitive task analysis. In J. M. Spector, M. D. Merrill, J.J.G. van Merriënböer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd ed.) (pp. 577–593). Mahwah, NJ: Lawrence Erlbaum Associates.

Corno, L., & Mandinach, E. B. (1983). The role of cognitive engagement in classroom learning and motivation. *Educational Psychologist*, *18*(2), 88–108.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–114.

Crandall, B., & Gretchell-Leiter, K. (1993). Critical decision method: A technique for eliciting concrete assessment indicators from the ''intuition'' of NICU nurses. *Advances in Nursing Science*, *16*(1), 42–51.

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods*. New York: Irvington.

De Corte, E. (2003). Transfer as the productive use of acquired knowledge, skills, and motivations. *Current Directions in Psychological Science*, *12*(4), 143–146.

Duffy, T. M., & Jonassen, D. H. (Eds.). (1992). *Constructivism and the technology of instruction: A conversation*. Mahwah, NJ: Lawrence Erlbaum Associates.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363–406.

Feldon, D. F. (2007). Implications of research on expertise for curriculum and pedagogy. *Educational Psychology Review*, *19*(2), 91–110.

Feldon, D. F., & Clark, R. E. (2006). Instructional implications of cognitive task analysis as a method for improving the accuracy of experts' self-report. In G. Clarebout & J. Elen (Eds.), *Avoiding simplicity, confronting complexity: Advances in studying and designing (computer-based) powerful learning environments* (pp. 109–116). Rotterdam, The Netherlands: Sense Publishers.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist*, *34*, 906–911.

Gentner, D., Lowenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, *95*(2), 395–408.

Glaser, R. (1992). Expert knowledge and processes of thinking. In D. F. Halpern (Ed.), *Enhancing thinking skills in the sciences and mathematics* (pp. 63–75). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gredler, M. E. (1996). Educational games and simulations: A technology in search of a research paradigm. In D. H. Jonassen (Ed.), *Handbook of research for educational* *communications and technology* (pp. 521–540). New York: Simon & Schuster Macmillan.

Halsted, W. S. (1904). The training of the surgeon. *Bulletin of the John Hopkins Hospital*, *15*, 267–276.

Hayes, R. T. (2005, November). *Effectiveness of instructional games: A literature review and discussion*. Technical Report 2005-004, Naval Air Warfare Center Training Systems Division, 12350 Research Parkway, Orlando, Florida 32826-3275.

Hinds, P. J. (1999). The curse of expertise: The effects of expertise and debiasing methods on predictions of novice performance. *Journal of Experimental Psychology: Applied*, *5*(2), 205–221.

Jonassen, D. H. (1991). Evaluating constructivist learning. *Educational Technology*, *31*(9), 28–33.

Jonassen, D. H., Tessmer, M., & Hannum, W. H. (1999). *Task analysis methods for instructional design*. Mahwah, NJ: Lawrence Erlbaum Associates.

Kirschner, P., Sweller, J., & Clark, R. E. (2006). Why minimally guided learning does not work: An analysis of the failure of discovery learning, problem-based learning, experiential learning and inquiry-based learning. *Educational Psychologist*, *41*(2), 75–86.

Kozma, R. (1991). Learning with media. *Review of Educational Research*, *61*(2), 179–211.

Kozma, R. B. (1994). Will media influence learning? Reframing the debate. *Educational Technology, Research, and Development*, *42*(2), 7–19.

Kuhn, D., Black, J., Keselman, A., & Kaplan, D. (2000). The development of cognitive skills to support inquiry learning. *Cognition and Instruction*, *18*, 495–523.

Lee, R. L. (2004). The impact of cognitive task analysis on performance: A meta analysis of comparative studies. Unpublished doctoral dissertation, Rossier School of Education, University of Southern California, USA.

Levie, W. H. (1989). Media attributes. In M. Eraut (Ed.), *The international encyclopedia of educational technology* (pp. 398–401). New York: Pergamon.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Masui, C., & De Corte, E. (1999). Enhancing learning and problem solving skills: Orienting and self-judging, two powerful and trainable learning tools. *Learning and Instruction*, *9*, 517–542.

Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist*, *59*(1), 14–19.

Mayer, R. E. (Ed.). (2005). *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.

Mayer, R. E., Mautone, P., & Prothero, W. (2002). Pictorial aids for learning by doing in a multimedia geology simulation game. *Journal of Educational Psychology*, *94*(1), 171–185.

McCombs, B. L. (1988). Motivational skills training: Combining metacognitive, cognitive, and affective learning strategies. In C. E. Weinstein, E. T. Goetz, & P. A. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 141–169). San Diego, CA: Academic Press.

Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research and Development*, *50*(3), 42–59.

Merrill, M. D. (2006). Hypothesized performance on complex tasks as a function of scaled instructional strategies. In J. Elen & R. E. Clark (Eds.), *Handling complexity in learning environments: Research and theory* (pp. 265–282). Oxford: Elsevier Science Limited.

Merrill, M. D. (2007). A task-centered instructional strategy. *Journal of Research on Technology in Education*, *40*(1), 33–50.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Moreno, R., & Mayer, R. E. (2005). Role of guidance, reflection and interactivity in an agent-based multimedia game. *Journal of Educational Psychology*, *97*(1), 117–128.

O'Neil, H. F., Wainess, R., & Baker, E. L. (2005). Classification of learning outcomes: Evidence from the games literature. *The Curriculum Journal*, *16*(4), 455–474.

Paas, F., & van Merriënböer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors*, *35*(4), 737–743.

Pea, R. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education and human activity. *Journal of the Learning Sciences*, *13*(3), 423–451.

Perkins, D. N., & Grozner, T. A. (1997). Teaching intelligence. *American Psychologist*, *52*(10), 1125–1133.

Piaget, J. (1928). *Judgment and reasoning in the child*. New York: Harcourt.

Pintrich, P. R., & DeGroot, E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*, 33–40.

Ravitch, D., & Viteritti, J. P. (2001). *Making good citizens: Education and civil society*. New Haven, CT: Yale University Press.

Redding, R. E. (1995). Cognitive task analysis for instructional design: Applications in distance education. *Distance Education*, *16*, 88–106.

Romiszowski, A. J. (2006). Topics for debate. *Educational Technology*, 61–63.

Ryder, J. M., & Redding, R. E. (1993). Integrating cognitive task analysis into instructional systems development. *Educational Technology, Research and Development*, *41*, 75–96.

Salomon, G. (1984). Television is ''easy'' and print is ''tough'': The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology*, *76*, 774–786.

Schaafstal, A., Schraagen, J. M., & van Berlo, M. (2000). Cognitive task analysis and innovation of training: The case of the structured troubleshooting. *Human Factors*, *42*, 75–86.

Schunk, D. H., & Pajares, F. (2004). Self-efficacy in education revisited: Empirical and applied evidence. In D. M. McInerney & S.Van Etten (Eds.), *Big theories revisited* (pp. 115–138). Greenwich, CT: Information Age Publishing.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, *16*, 475–522.

Shulman, L., & Keisler, E. (Eds.). (1966). *Learning by discovery: A critical appraisal*. Chicago, IL: Rand McNally.

Sitzmann, T., Brown, K. G., Casper, W. J., Ely, K., & Zimmerman, R. D. (2008). A review and meta-analysis of the nomological network of trainee reactions. *Journal of Applied Psychology*, *93*(2), 280–295.

Sitzmann, T., Kraiger, K., Stewart, D., & Wisher, R. (2006). The comparative effectiveness of web-based and classroom instruction: A meta-analysis. *Personnel Psychology*, *59*, 623–664.

Stolovitch, H. D., & Keeps, E. (2002). *Telling ain't training*. Alexandria, VA: American Society for Training and Development.

Sugrue, B., & Clark, R. E. (2000). Media selection for training. In S. Tobias & D. Fletcher (Eds.), *Training & retraining: A handbook for business, industry, government and the military*. New York: Macmillan.

Sweller, J. (1999). *Instruction design in technical areas*. Camberwell, Australia: ACER.

Sweller, J. (2007). What cognitive architecture tells us about constructivism. In J. Elen & R. E. Clark (Eds.), *Handling complexity in learning environments: Research and theory* (pp. 13–26). Oxford: Elsevier Science.

Thompson, Inc. (2002). *Thompson Job Impact Study*. Naperville, IL: Thompson NETg.

van der Linden, J. L., Erkens, G., Schmidt, H., & Renshaw, P. (2005). Collaborative learning. In R. J. Simons, J. L. van der Linden, & T. M. Duffy (Eds.), *New learning* (pp. 33–48). Dordrecht: Kluwer Academic Press.

van Merriënböer, J.J.G. (1997). *Training complex cognitive skills: A four-component instructional design model for technical training*. Englewood Cliffs, NJ: Educational Technology Publications.

van Merriënböer, J.J.G., Clark, R. E., & de Croock, M. B. M. (2002). Blueprints for complex learning: The 4C/ID*-model. *Educational Technology Research and Development*, *50*(2), 39–64.

van Merriënböer, J.J.G., & Kirschner, P. A. (2007). *Ten steps to complex learning: A systematic approach to four-component instructional design*. Mahwah, NJ: Lawrence Erlbaum Associates.

Velmahos, G. C., Toutouzas, K. G., Sillin, L. F., Chan, L., Clark, R. E., Theodorou, D., & Maupin, F. (2004). Cognitive task analysis for teaching technical skills in an inanimate surgical skills laboratory. *The American Journal of Surgery*, *18*, 114–119.

Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.

Winthrop, H. (1963). Some considerations concerning the status of phenomenology. *Journal of General Psychology*, *68*, 127–140.

Yates, K. (2007). Towards a taxonomy of task analysis methods: A search for cognition and task analysis interactions. Unpublished doctoral dissertation, Rossier School of Education, University of Southern California.

CHAPTER NINE

Instructional Strategies for Receptive Learning Environments

Richard E. Mayer
Ruth Colvin Clark

Suppose you were asked to update the sales force on the features and benefits of the new product line. Because the sales force is located world-wide, you need a distance-learning delivery strategy allowing everyone to access consistent information quickly. You could develop a pod-cast that used audio to describe the new product features, or you might post a web page that used text to explain online visuals of the product line. Alternatively, you could record an e-learning session through your virtual classroom tool in which the product developers use audio to describe visual illustrations of the new product features and marketing managers summarize the product rollout plans. In all of these alternatives, the learners will listen, read, or view the content but will not have an opportunity to overtly engage with the content. We refer to learning opportunities that do not require visible responses from the learner as receptive learning environments. Can receptive environments be as effective as more high-engagement strategies such as simulations or games? What are the features that make some receptive environments more effective than others? How can we design receptive learning environments that promote deep learning?

In this chapter, we provide an introduction in which we define key terms and issues; outline a theory of learning based on the distinction among extraneous processing, essential processing, and generative processing during learning; and then explore instructional methods for reducing extraneous processing, instructional methods for managing essential processing, and instructional methods for fostering generative processing in receptive learning environments. We begin the first section by defining receptive learning and evidence-based practice, and distinguish between media and methods, and between technology-centered and learner-centered approaches.

WHAT ARE RECEPTIVE INSTRUCTIONAL ENVIRONMENTS?

Receptive instructional environments—also called low engagement learning strategies—use diverse media elements such as text, graphics, and audio to present content and instructional methods such as examples *without opportunities for overt learner responses*. Typical examples include text readings, web pages with text and visuals, lectures presented in a face-to-face or virtual learning environment, video presentations, and low engagement e-learning in the form of briefings or overviews.

The main distinction between receptive and other learning environments reviewed in this book is that receptive environments provide little or no opportunities *for overt responses* by the learner. For example, they contain no practice exercises or other opportunities for visible interactions. Receptive environments have often been cast in a pejorative light as less effective than moderate to high engagement environments such as simulations and games. For example, receptive environments are sometimes referred to as ‘‘passive’’ learning environments or online ‘‘page-turners.’’ A commonly held assumption is that learning outcomes are generally poorer from receptive learning environments than from high engagement lessons. However, we will see that the research evidence tells a quite a different story.

Receptive Versus ‘‘Active’’ Learning Environments

What research evidence do we have for the benefits or drawbacks of receptive learning environments? One recent experiment compared learning from so-called active and passive lessons that focused on the same topic. Haidet, Morgan, O’Malley, Moran, and Richards (2004) compared learning based on immediate and delated tests of medical residents from one version of classroom-based instruction that involved overt activity with a second version of classroom instruction that did not. In the ‘‘active’’ lesson, the residents were assigned to small groups and asked to discuss a case problem assigned by the instructor. After a brief team discussion, the instructor presented content that resolved the case. The class lasted an hour and discussed four problems. A comparison group (the passive version) listened to a one-hour traditional didactic lecture that covered the same technical information as the discussion class did. Learning was measured immediately after the session and one month later. Learner satisfaction was also evaluated.

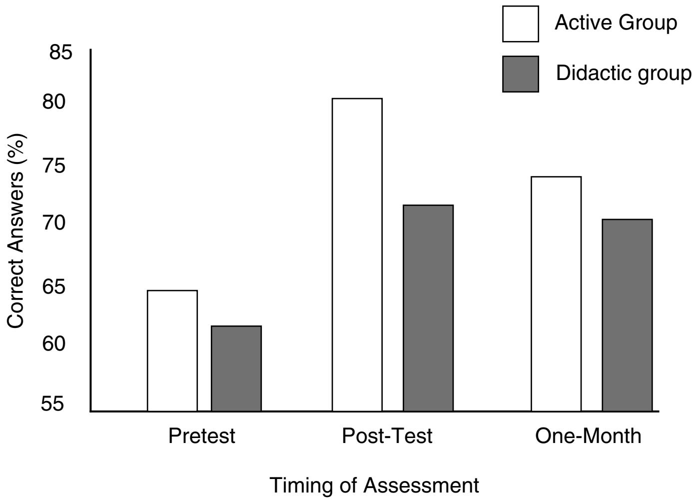


Figure 9.1 There Were No Significant Differences in Learning Between Active and Passive Groups.

Adapted from Haidet, Morgan, O'Malley, Moran, and Richards, 2004

As you can see in Figure 9.1, both groups showed pre- to post-test score gains, and there were no statistically significant differences in learning between the two groups—both on the immediate and on the delayed test. Similarly, Boucheix (2008) has shown that allowing learners to have control over the presentation of animations (such as pausing) does not help learning and can actually harm the learning of students who lack spatial abilities or prior knowledge about the domain of the animation. It's a common assumption that the more ''active'' or ''collaborative'' environment would lead to more learning and greater satisfaction. However, in this research we have evidence that overt activity does not always lead to better learning or greater learner satisfaction!

All Learning Is Active Learning

We know that, for learning to occur, the brain must actively engage with the content. However, as anyone who has learned from reading a chapter such as this one knows, active cognitive engagement does not necessarily require active hands-on activity. The important activity is cognitive activity! In fact, deeper processing may occur in a receptive learning environment than in a highly interactive environment in which activities disrupt the appropriate mental processes. For example, recent research reviews show that discovery methods of instruction—in which learners are free to engage in hands-on learning with little or no guidance—are not effective methods of instruction, whereas providing direct instruction can be highly effective (Kirschner, Sweller, & Clark, 2006; Mayer, 2004).

Clark (2009) and Clark and Mayer (in press) distinguish between implicit and explicit instructional methods to support encoding of new knowledge and skills. Both implicit and explicit methods lead to learning. Implicit instructional methods such as relevant graphics support the organization and integration of new content in the absence of any visible learner activity. The activity is all cognitive. Explicit instructional methods support the organization and integration of new content through well-designed exercises and interactions in which the learner is overtly engaged in a manner that supports the learning objective.

Perhaps the bulk of instructional materials available today reflects a receptive environment. For example, websites identified through Internet research, texts like this one, and audio-visual programs such as videos or lectures all reflect receptive learning environments. Most formal instructional environments that include active engagement opportunities also incorporate receptive elements, including both classroom lecture and distance education. For example, an instructor shows learners how to complete a task, after which they practice the task themselves. The demonstration portion of the class is a receptive event. Because of the large volume of receptive instruction produced annually, it's important that all instructional professionals be aware of the evidence and psychology to guide their construction of receptive environments that lead to active cognitive processing.

WHAT IS EVIDENCE-BASED PRACTICE?

Instructional programs should be guided by scientific evidence rather than opinion, ideology, or common practice. This simple statement reflects the call for *evidence-based practice*, which has been the centerpiece of educational science for a hundred years. As early as 1906, the famous educational psychologist E.L. Thorndike (1906, p. 206) called on ''leaders in education [to] direct their choices of methods by the results of scientific investigation rather than general opinion.'' Similarly, in a recent report commissioned by the National Academy of Sciences, Shavelson and Towne (2002, p. 1) noted the ''rising enthusiasm for evidence-based education policy and practice'' is based on the idea that ''one cannot expect . . . efforts in education to have significant effects without research-based knowledge to guide them.''

Practitioners are also embracing evidence-based practice. Among the top-ten teaching and learning issues facing instructional professionals in 2007, the

EDUCAUSE advisory committee points to (1) establishing and supporting a culture of evidence, (2) demonstrating improvement of learning, and (3) translating learning research into practice as their top three priorities (Campbell & Oblinger, 2007). In an industry that invests over $50 billion annually in training, organizational training professionals cannot afford to ignore evidence of what works as they design, develop, or select workforce learning environments. In this chapter, we take an evidence-based approach by seeking to identify instructional design principles that are grounded in empirical research on the effects of instructional methods (Clark, 2009; Clark & Lyons, 2004; Clark & Mayer, 2008; Clark, Nguyen, & Sweller, 2006; Mayer, 2008).

Suppose you want to determine the effectiveness of an instructional method or feature—such as using conversational style rather than formal style. A useful approach is to compare the test performance of a group that received the regular lesson (control group) with one that received the regular lesson along with the new feature (experimental group). By using an appropriate statistical test, you can determine whether the experimental group scored significantly higher than the control group—where *significantly* means that there is less than a 5 percent chance that you would be wrong in saying that the test score of the experimental group is greater than the test score of the control group.

In addition to statistical tests, it is useful to determine the practical importance of the effect, using a measure called *effect size* (Cohen, 1988). Effect size tells you about the strength of the effect. For example, effect size can be measured by subtracting the mean score of the control group from the mean score of the experimental group, and dividing by the pooled standard deviation (a measure of how spread out the scores are). The resulting measure of effect size (called Cohen's *d*) tells you how many standard deviations better the experimental group did as compared to the control group, with $d = .2$ considered a small effect, $d = .5$ considered a medium effect, and $d = .8$ considered a large effect. A useful feature of effect size is that it allows researchers to combine data across many different experiments that examine the effect of the same method or feature, because they all can be translated into a common metric—effect size. This technique for combining results across different experiments based on average effect size is called *meta-analysis*.

In the evidence we summarize in this chapter, we always compare the problem-solving transfer test performance of a group that receives a lesson with a particular instructional method or feature against a group that receives the same lesson without the instructional method or feature. In these studies, the learners are predominantly adults—usually college students (but occasionally high school students)—and the training materials generally explain how something works.

INSTRUCTIONAL MEDIA VERSUS INSTRUCTIONAL METHODS

In our introductory scenario, we described several alternative delivery media for our product-knowledge lesson. For example, we considered pod-casts that deliver audio, web pages that deliver text and pictures, and virtual classroom technologies that can deliver visuals, audio, and text. You might be tempted to ask: Which of these technologies is most effective for learning?

In designing or selecting multimedia training materials, it is tempting to become fascinated with the capabilities of various instructional media—such as computer graphics or web-based interactivity or virtual classrooms. However, we urge you to consider the idea that instructional methods cause learning, not instructional media. You would think it odd to ask whether books are good or bad delivery devices, and it is just as odd to ask whether computers are good or bad. Books and computers do not cause learning; rather, the instructional method you use with them is what causes learning. Richard Clark (2001, p. ix) summarizes this point: ''The evidence justified the definite conclusion that media do not cause learning and I recommended that people stop asking research and evaluation questions about learning from media.'' In this chapter we focus on various instructional methods and features, many of which have the same effects on learning in a paper-based environment and a computer-based environment. This is why we focus on instructional methods rather than instructional media in designing multimedia training materials. However, we do know that not all delivery media can readily deliver all training methods. For example, a book can deliver text and visuals but not audio, while a pod-cast delivers audio but not text or visuals. Therefore we must consider the trade-offs in any training delivery system when selecting media and designing training for a given medium.

LEARNER-CENTERED VERSUS TECHNOLOGY-CENTERED APPROACHES

In designing instructional training, it is tempting to focus on incorporating the capabilities of cutting-edge technologies into the lesson. Mayer (2001) calls this taking a *technology-centered approach*, because the learner is being asked to adapt to cutting-edge technology. The problem with taking a technology-centered approach is that the instructional program may not be suited to the needs of the learner. For example, advances in graphics and communication technologies make it possible to deliver elaborate animations and stunning video, but in some situations people tend to learn just as well or better from a series of static graphics (Mayer, Hagarty, Mayer, & Campbell, 2005). Cuban

(1986) has shown how educational technologies of the 20th century tended to gain high popularity at first, and then fizzle out as they proved to be ineffective—ranging from motion pictures in the 1920s, educational radio in the 1930s and 1940s, educational television in the 1950s, computer-based programmed instruction in the 1960s, and so on. Although web-based instruction is currently a popular venue for training, this technology will likely also prove fruitless if it is used without regard to how people learn (O'Neil, 2005; O'Neil & Perez, 2006).

In contrast, we advocate taking a learner-centered approach to educational technology, in which you focus on adapting technology to serve as an aid to human learning (Mayer, 2001). Instead of asking, "How can we use cutting-edge technology in multimedia training materials?" we ask, "How can we adapt technology to aid human learning?" In order to take a learner-centered approach, it is essential to understand how people learn. This topic is addressed in the next section.

HOW LEARNING WORKS

What Is Learning and Instruction?

Learning is a change in knowledge due to experience (Mayer, 2008). This definition has three components: (a) learning involves a change in a person, (b) learning involves the person's knowledge and can only be inferred indirectly from the person's behavior, and (c) learning is caused by experience such as participating in an instructional program.

Instruction refers to the instructor's construction of environments that will afford experiences for learners that are intended to lead to learning (Mayer, 2008). This definition has two components: (a) the instructor creates experiences for people, and (b) the goal is to promote learning in people. In organizations, the changes in knowledge are intended to result in behavioral changes that promote operational goals. In short, the goal of most training programs is to foster a change in the learner, particularly a specified change in what the learner knows. These definitions are broad enough to include a variety of kinds of multimedia training materials ranging from paper-based booklets, to PowerPoint presentations, to interactive computer-based programs.

How Does Learning Work?

Figure 9.2 summarizes a cognitive model of how people learn from lessons containing words and pictures. As you can see, we begin on the left side with instructional materials consisting of words (in spoken or printed form) and pictures (in static form as drawings, graphs, maps, or photos, or in dynamic form as animation or video). In the next step to the right, pictures and printed

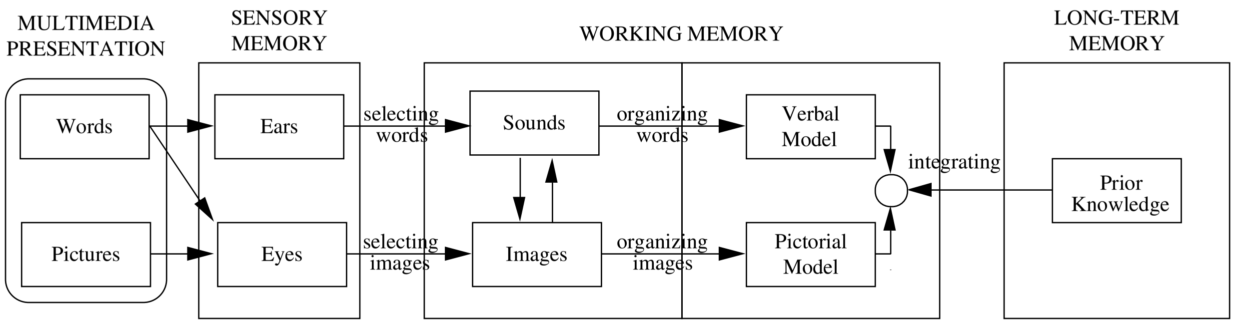


Figure 9.2 A Cognitive Theory of Multimedia Learning.

From Mayer, 2005a

words enter the learner's sensory memory through the eyes and spoken words enter through the ears. If the learner pays attention to this material before it decays, it can be transferred to working memory as indicated by the *selecting words* and *selecting images* arrows. In working memory, the learner can engage in deeper processing of the selected material by mentally organizing the incoming words and images into coherent cognitive structures (as indicated by the *organizing words* and *organizing images* arrows) and by integrating the resulting verbal model and pictorial model with each other and with prior knowledge from long-term memory (as indicated by the *integrating* arrows). The knowledge that is constructed in working memory can then be stored in long-term memory.

How Does Instruction Work?

As you can see in Figure 9.2, the goal of instruction is to guide three appropriate kinds of cognitive processing during learning—selecting relevant information (that is, making sure the learner attends to the essential information), organizing the information into a coherent mental representation (that is, making sure the learner builds verbal and pictorial models), and integrating the new representation with other knowledge (that is, making sure the learner connects verbal and pictorial models with each other and with relevant prior knowledge from long-term memory).

All of this processing must take place within a cognitive system that has separate verbal and visual channels, each of which is limited in capacity. According to recent theories of instruction (Mayer, 2005a), inspired by Sweller's (1999, 2005) cognitive load theory, the central problem for instructional designers is to be sensitive to three kinds of cognitive processing:

Extraneous processing—cognitive processing that does not support the instructional goal, and which is caused by extraneous material in the lesson or poor layout of the lesson. For example, an animated decorative visual can distract the learner from the main instructional message.

Essential processing—cognitive processing of the essential material in the lesson, which is caused by the complexity of the essential material, resulting in selecting the essential material to be represented in working memory for further processing. For example, essential processing is required to understand an explanation of how a new data management system works, but this task can be simplified by explaining how each major part of the system works.

Generative processing—cognitive processing aimed at making sense of the essential material by mentally organizing it and integrating it with other knowledge. For example, generative processing is reflected in thinking about how to apply new content to a job-relevant task.

In order to understand the material in a lesson, a learner needs to engage in essential and generative processing. However, given that learners have only a limited amount of processing capacity, if they use their capacity for extraneous processing, they may not have sufficient capacity left over for essential and generative processing. Therefore, we offer three overarching goals for instructional designers:

Reduce extraneous processing—You should eliminate anything in a lesson that encourages the learner to waste precious processing capacity, so extraneous material should be omitted and confusing layouts should be simplified.

Manage essential processing—When the essential material is complex, you can't eliminate it, but you can present it in ways that help the learner manage it. These techniques include breaking it into smaller segments, providing pretraining on the key concepts, and presenting words in spoken form rather than printed form.

Foster generative processing—When properly implemented, techniques for reducing extraneous processing and techniques for managing essential processing leave the learner with cognitive capacity that can be used for generative processing. Your final goal is to foster generative processing, that is, to encourage the learner to use available processing capacity to make sense of the presented material. Some techniques for fostering generative processing include using appropriate pictures that correspond to the words, and using words in conversational style.

In the remainder of this chapter we explore techniques for accomplishing these three goals within a receptive learning environment.

PRINCIPLES FOR REDUCING EXTRANEOUS PROCESSING

The most common problem with training materials is that they are more complicated than they need to be. Sometimes they contain extraneous material—content that is not directly related to the instructional goal—so processing the extraneous material is a form of extraneous cognitive processing. Sometimes they are laid out in ways that induce extraneous processing, such as when printed words are placed in captions at the bottom of the screen far away from the corresponding part of the graphic. An important goal of instructional design is to reduce the need for the learner to engage in extraneous processing. In this section, we explore four principles for reducing extraneous processing: coherence, signaling, redundancy, and spatial contiguity principles.

APPLYING THE COHERENCE PRINCIPLE

The coherence principle is: People learn better from multimedia training lessons when extraneous material is excluded rather than included. In designing training materials, you might be tempted to spice up the lesson by adding extraneous material in the form of background sounds, interesting but irrelevant stories, attention-grabbing graphics, additional text, and technical details. For example, in a lesson on how lightning works, you might think it would be good to insert short (ten-second) video clips of lightning strikes with thunderclaps or background music. Alternatively, you might decide to insert interesting but irrelevant facts such as, ''When flying through updrafts, an airplane ride can become bumpy. Metal airplanes conduct lightning well, but they sustain little damage because the bolt, meeting no resistance, passes right through.'' Instead, you could add more text or include technical details about the physics formulas that are involved in lightning formation. Our advice is simple: Don't do it. Based on the coherence principle, Mayer and Moreno (2003) recommend that instructional designers engage in *weeding*, that is, eliminating interesting but extraneous material from the lesson in order to reduce extraneous processing by the learner. The coherence principle includes five specific recommendations: omit extraneous sounds, omit irrelevant stories, omit attention-grabbing graphics, omit technical details, and keep the text short.

Rationale for the Coherence Principle

Consider the cognitive consequences of adding extraneous material to a lesson. The learner has limited cognitive capacity, so when the learner uses precious cognitive capacity to process extraneous material, less capacity is available for engaging in essential and generative processing—the processes that lead to meaningful learning.

Evidence for the Coherence Principle

In a review of eleven experiments comparing the test performance of people who learned about lightning or braking systems with a concise lesson or an embellished lesson (with added interesting but extraneous material), people in the concise group performed better on problem-solving transfer tests than did people in the embellished group (Mayer, 2005b). The median effect size favoring the concise group was $d = 1.32$, which is considered a large effect.

Applications of the Coherence Principle

The bottom line for practitioners applying the coherence principle is that *Less Is More!* It's a common misconception that adding extensive explanations will ensure understanding or adding interesting stories to a dry technical topic will

Figure 9.3 A Screen to Add Interest to the Database Lesson.

From Clark and Mayer, 2008

motivate learners and improve the learning experience. In fact, controlled research studies have shown that such well-intended additions more often than not depress rather than improve learning. Here we summarize a few proven guidelines for applying the coherence principle.

1. **Minimize Interesting Stories and Visuals Related to the Topic But Irrelevant to the Learning Goal.** In Figure 9.3 we show a screen capture from an asynchronous lesson on constructing databases. To add interest, the author sprinkled some facts about database abuse such as this one throughout the lesson. In fact, Harp and Mayer (1997, 1998) showed that adding stories in the form of text and/or visuals to a lesson reduced learning compared to learners who studied the basic lesson minus the stories. We are not suggesting that all stories be deleted from all instruction. Rather, we recommend that you ask yourself whether (a) the story is directly related to the instructional objective and (b) whether it might interfere with the selecting, organizing, or integrating processes we summarized previously.
2. **Write Lean Text.** Many subject-matter experts and some instructional designers tend to write text or give lectures that are very lengthy. Subject-matter experts often want to tell everything they know on the topic. Instructional professionals working in asynchronous e-learning may want to add additional details since the training will not have an

instructor to elaborate. However, several research experiments that compared lean and flabby versions of instruction (for example, a one-hundred-word description of the value of creating a primary key in a database compared to forty words that communicated the same idea) consistently found better learning from the lean versions (Mayer, Bove, Bryman, Mars, & Tapangco, 1996; Mayer and Jackson, 2005).

3. **Delete Extraneous Audio.** Many individuals like music playing while they are working or studying. It's also a common practice to add a music track to e-learning to increase the motivational impact of the message and to accommodate the MTV generation! Research that compared learning from lessons with and without background music recommends that you omit music when generative learning is your goal (Moreno & Mayer, 2000). The experiments found that lessons that omitted environmental sounds and music resulted in 105 percent better learning than lessons that included the extra audio. You may want to add a brief musical background to a title screen or to segments where your goal is motivational rather than learning. However, where your goal is deep learning (that is, learning that promotes transfer performance), we recommend eliminating any auditory input not directly related to the instructional goals.

In short, our message is focus the instructional elements—words, graphics, and audio—on just what is needed to concisely and clearly convey your content. Delete or rewrite flabby scripts, keep music only in non-instructional segments, and avoid the temptation to spice up your instruction with stories that do not directly relate to the instructional objective.

APPLYING THE SIGNALING PRINCIPLE

Sometimes it is not possible to weed out all of the extraneous material in a lesson. This situation calls for the signaling principle: People learn better from multimedia training materials when cues are added that signal the organization of the essential material. Signaling cues include inserting an outline, using headings keyed to the outline, and including pointer words such as "first, second, third" or "the result is . . .". Signaling can also involve highlighting key printed words through italics, bolding, underlining, color, font size, or spacing, and highlighting key spoken words through increasing the volume or decreasing the pitch of the voice.

Rationale for the Signaling Principle

Signaling helps to direct the learner's attention toward the essential material, and thereby reduces extraneous processing by the learner. Reducing extraneous processing results in more cognitive capacity for making sense of the essential material—by mentally organizing it and relating it with prior knowledge.

Evidence for the Signaling Principle

In a review of three experiments, Mayer (2005b) reported that people who learned about airplane flight or lightning with a signaled lesson performed better on problem-solving transfer tests than did people who learned without signals. The median effect size favoring the signaled treatment was $d = 0.60$, which is considered a medium-to-large effect.

Applications of the Signaling Principle

When your instructional material is complex, you can guide the learners' eyes and ears to important elements of the content by providing signals in the form of arrows, bolding, vocal emphasis, and phrases that organize content such as ''*First* . . . .'', and ''*The three main items to consider are* . . . .''. As an example, in this chapter our goal is to help you process our content through signals such as the headers and sub-headers, text treatments such as bolding and italicizing, and organizing statements such as: ''*We offer three overarching goals for instructional designers:*'' and ''*In this section, we explore four principles for reducing extraneous processing—coherence, signaling, redundancy, and spatial contiguity principles.*''

In displays that include complex visual information, you will rely on visual cues such as arrows, circles, and highlighting. Signals are most important when the materials are lengthy or complex. For example, animated displays tend to show a great deal of visual information in a transitory manner. Adding cues such as circles or highlights to animations will help direct the learner's eye to the appropriate section of the visual display.

APPLYING THE REDUNDANCY PRINCIPLE

Suppose you have a multimedia training segment consisting of animation with concurrent narration. You might be tempted to add redundant on-screen text, for example, printed sentences at the bottom of the screen that correspond to the sentence being spoken by the narrator. Again, our advice is: Don't do it! This advice is embodied in the redundancy principle: People learn better from graphics and narration than from graphics, narration, and on-screen text.

Rationale for the Redundancy Principle

What happens in the learner's cognitive system when the learner is exposed to identical streams of spoken and printed words? The learner is encouraged to engage in extraneous processing (a) by trying to reconcile the two streams and

(b) by trying to scan between the printed words at the bottom of the screen and the corresponding elements in the graphic. Both activities waste precious cognitive capacity, leaving less capacity for making sense of what the words say.

Evidence for the Redundancy Principle

In a review of ten experiments involving lessons on mathematical problem solving, electrical engineering, lightning, and environmental science, people who received narrated animation performed better on problem-solving transfer tests than did people who received narrated animation with concurrent on-screen text (Mayer, 2005b). The median effect size was $d = .69$, which is in the medium-to-large range. In some cases, however, using redundant on-screen text may be warranted, such as when the narration is difficult, contains jargon terms, or is not in the learner's native language.

Applications of the Redundancy Principle

Many practitioners feel they should provide learners with on-screen text and simultaneous narration of that text to accommodate different learning styles or to meet 508 compliance requirements (governmental regulations concerning access for people with disabilities). For example, they produce screens such as the one illustrated in Figure 9.4 in which a visual is explained by onscreen text as well as narration of that same text.

Based on the redundancy principle, we recommend that you use only one media element to describe visuals. For example, you could use audio narration. Alternatively, you could use text. In general, however, you should avoid using audio and redundant text to describe a visual. As we described previously, learning requires a maximal amount of mental capacity devoted directly to the learning task. When scarce mental resources are used in unproductive ways, learning is depressed. To accommodate Section 508 as well as the modality principle (described below), we recommend that, in e-learning lessons that can technically support audio, you make the audio the default option but allow learners to turn off the audio and view the words in text. Similarly, English language learners (people who are in the process of becoming proficient in English) may benefit from seeing printed text or at least having the option to see printed text.

APPLYING THE SPATIAL CONTIGUITY PRINCIPLE

The contiguity principle is: People learn better when corresponding printed words and pictures are near rather than far from each other on the screen or page. In compliance with this principle, Mayer and Moreno (2003) recommend *aligning*, that is, placing printed words near corresponding parts of the graphic in order to reduce the need for visual scanning.

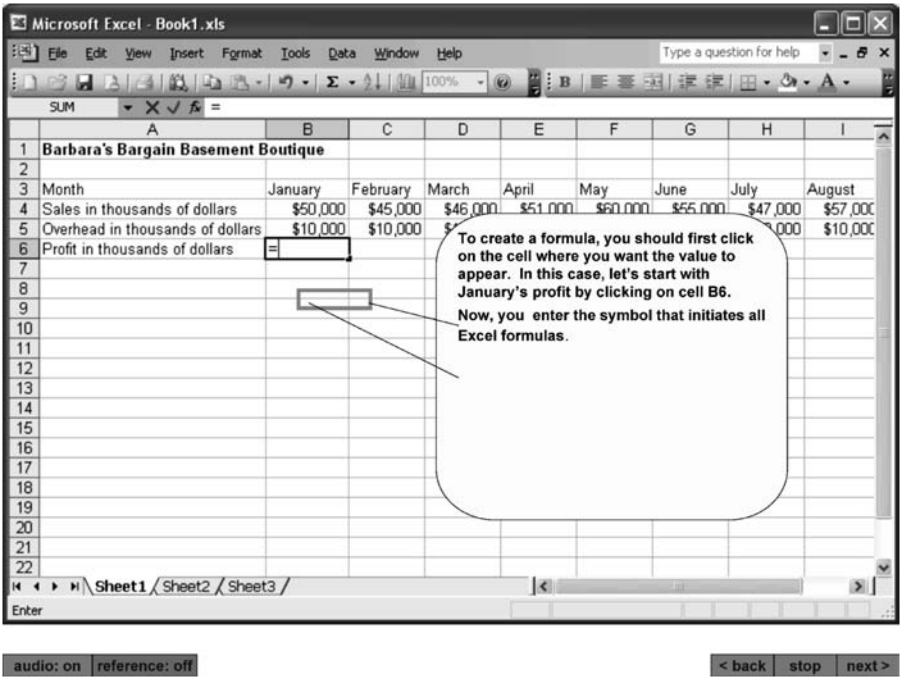


Figure 9.4 Visual Described by Onscreen Text and Narration.
From Clark and Mayer, 2008

Rationale for the Contiguity Principle

Sometimes extraneous processing is caused by poor layout rather than by having extraneous material in the lesson. When printed text captions are presented at the bottom of an animation or a static graphic, the learner has to scan between the words in the caption (for example, ''negative ions fall to the bottom of the cloud'') and the corresponding part of the graphic (for example, small circles with negative signs moving from the top to the bottom of the cloud). Looking back and forth creates *split attention*, in which the learner cannot be viewing the animation if he or she is reading the text and cannot be reading the text if he or she is viewing the animation. Placing the words near the corresponding part of the graphic reduces the need to scan—a form of extraneous processing—and helps guide the learner's attention.

Evidence for the Contiguity Principle

In a review of eight experiments, Mayer (2005b) reported that people who learned about braking systems, lightning, electrical engineering, or solving mathematics problems with an integrated layout (words near corresponding graphics) performed better on problem-solving transfer tests than did people who learned with a separated layout (words far from corresponding graphics). The median effect size favoring the integrated treatment was $d = 1.11$, which is considered a large effect. In another meta-analysis of research involving contiguity, Ginns (2006) also found consistent support for the contiguity principle.

Applications of the Contiguity Principle

We have all seen frequent violations of the contiguity principle both in print and e-learning delivery formats. A common violation in books such as this one consists of explanatory text on one page and an illustrative graphic on the back of the page. Neither the text nor the graphic alone tells the full story. To make sense of the message, both the text and the visual must be processed. We've all experienced the frustration of flipping pages back and forth to make sense of this type of split attention layout. That frustration is your working memory feeling the extra mental load of having to integrate content while trying to understand it. Digital media often violate contiguity in their placement of text and graphics. For example, in a scrolling screen, text appears at the top of the screen and the graphic underneath it. When the viewer scrolls down to the graphic, the text is no longer visible and vice versa. A simple solution to many violations of contiguity is to place text callouts on top of the visual with lines pointing to the relevant section of the instruction, as shown in Figure 9.4. Violations of contiguity in print and online media are common and in most cases can be remedied by some simple realignment of text and visuals. Although some might make the case that principles such as coherence and contiguity do not apply to the younger generation who are fluent working with various forms and formats of digital inputs and devices, keep in mind that the research showing that ''less is more'' is based on learners in the eighteen-to-twenty-two-year age range.

PRINCIPLES FOR MANAGING ESSENTIAL PROCESSING

In the foregoing section, you saw how to create lessons that minimize the need for learners to engage in extraneous cognitive processing—mainly by keeping the lesson as simple as possible. However, it is not possible to simplify the lesson when the essential material is complex. In this case what is needed is techniques for helping the learner manage essential processing, that is, building a mental representation of the core material. In this section, we examine three principles for managing essential processing: segmenting, pretraining, and modality principles.

APPLYING THE SEGMENTING PRINCIPLE

Suppose you have a narrated animation on lightning formation that runs for about two and half minutes. The lesson describes sixteen major steps and is rapidly paced. Some students may not fully comprehend one step in the process before the next one is presented and may not see the causal relation between one step and the next. What can be done to help learners more fully comprehend this complex material? We offer the segmenting principle: People learn better when multimedia training lessons are presented in user-paced segments rather than as a continuous unit. For example, the lightning lesson can be broken into sixteen bite-size segments, each containing one or two sentences and about eight to ten seconds of animation with a ''CONTINUE'' button in the lower-right corner. After each segment, the lesson stops until the learner clicks on the ''CONTINUE'' button for the next segment. In this way, the learner can digest one step before moving on to the next one.

Rationale for the Segmenting Principle

Consider what happens in the learner's working memory when he or she encounters a fast-paced lesson that contains complex material. The learner attends to some of the words and pictures (as indicated by the *select words* and *select images* arrows in Figure 9.2), but before he or she can process them more deeply (by mentally organizing them and integrating them with prior knowledge), the next set of new words and images appears. When a lot of complex material is presented at a fast pace, the learner is less likely to be able to engage in the cognitive processes of organizing and integrating, which are required for deep understanding.

Evidence for the Segmenting Principle

In a review of three experimental tests involving multimedia lessons on lighting and on how electric motors work, Mayer (2005c) reported that people who received narrated animations that were segmented performed better on problem-solving transfer tests than did people who received the lesson as a continuous unit. The median effect size was $d = .98$, which is considered a large effect.

Applications of the Segmenting Principle

In print media such as books like the one you are reading, you can move at your own pace and easily go back to review sections. We call these types of environments highly learner controlled. In asynchronous multimedia, however, the designer may set the lesson to play continuously in lengthy segments or may play short segments, stop the play, and ask learners to move ahead when they are ready by pressing a navigational button. Our recommendation here is pretty straightforward: Always allow the learners to control the rate at which they progress through online content, such as by pressing the continue button.

Some common media practices violate the segmenting principle. For example, in recordings of online classes, video examples, and narrated animations, the material plays continuously. Even though you provide pause and play buttons, recent research showed that in a narrated animation sequence, most learners did NOT use the pause function. Instead they replayed the entire animated lesson several times (Schar & Zimmermann, 2007). We recommend that you present short sequences of content and pause the program, allowing the learners to proceed when they are ready.

APPLYING THE PRETRAINING PRINCIPLE

Another evidence-based technique for managing essential processing is reflected in the pretraining principle: People learn better from multimedia training material when they know the names and characteristics of the main concepts. For example, consider a narrated animation on how a car's braking system works, which consists of six major steps and lasts about forty-five seconds. The lesson concerns several components that may not be familiar to the learner such as ''piston in master cylinder,'' ''fluid in brake tube,'' ''piston in wheel cylinder,'' ''brake shoe,'' and ''brake drum.'' Pretraining involves helping students learn the name and location of each part, as well as its characteristics, for example, a piston can be forward or back in a cylinder.

Rationale for the Pretraining Principle

Learning from a fast-paced narrated animation can overload working memory because the learner has two tasks: (a) building a causal model in which a change in one part causes a change in the next part, and so on, and (b) building component models concerning the name and characteristics of each part. Consider what happens when the learner encounters a sentence such as, ''When the driver steps on a car's brake pedal, a piston moves forward in the master cylinder.'' In building a causal model, the learner should comprehend how a change in one part can cause a change in another part (the pedal going down causes the piston to move forward). However, if the learner does not know what a piston is, the learner may have to grapple with trying to find out how it works and where it is, which will leave less cognitive capacity for building the causal model. Through learning the names and characteristics of key parts in advance, some essential processing is off-loaded to the pretraining session.

Evidence for the Pretraining Principle

In a review of seven experiments involving braking systems, tire pumps, geology, and electrical engineering, people performed better on a problem-solving transfer test when they studied a technical lesson *after* they had received pretraining in the names and characteristics of the key concepts or parts. The pretraining did not provide any additional information, but gave people the chance to master part of the essential material before moving on to the task of building a casual model. The median effect size was d = .92, which is considered a large effect.

Applications of the Pretraining Principle

The pre-training principle can be applied by adding pre-work to a class or by re-sequencing your content within a lesson. For example, if you plan to teach a lesson on how to enter a valid formula into an Excel spreadsheet, the learner must know what a formula is, the correct formula format, as well as the steps to enter and execute the formula in an Excel spreadsheet. To apply the pretraining principle, you could sequence some of the core concepts, such as what is a formula and formula formatting conventions, as pre-work for the class or as the initial concepts within the body of the lesson. Clark (2008b) refers to core lesson concepts as ''supporting knowledge'' and recommends that you teach all major supporting knowledge topics before you teach the task in which the learner will apply that knowledge. During the design phase of your training, outline your content sequence by placing important concepts first, followed by the key steps or guidelines of the lesson procedure.

APPLYING THE MODALITY PRINCIPLE

Suppose you have a fast-paced animation (or a series of graphic slides) that depicts the steps in lightning formation, and that you wish to accompany it with a concurrent verbal script. Does it matter whether you present the words in printed form (for example, as a caption below the animation or below each slide) or in spoken form (for example, as narration)? According to the modality principle, the answer is yes. The modality principle is: People learn better from multimedia training material when words are presented in spoken form rather than printed form.

Rationale for the Modality Principle

Consider what happens in people's visual channels when they have to process pictures and printed words. As you can see from the arrows on the left side of Figure 9.2, pictures enter the visual channel through the eyes and printed words enter the visual channel through the eyes. The problem is that all this visual processing can overload the visual channel because people cannot look at printed words and pictures at the same time. Although both activities represent essential processing, neither one gets the full attention it deserves. The solution to this problem is to present the words in spoken form, so the eyes can be used solely for processing pictures and the ears for processing words. Mayer and Moreno (2003) refer to this technique as *off-loading*, because the task of initially processing words is shifted from the eyes to the ears.

Evidence for the Modality Principle

There is overwhelming evidence for the modality principle. In a review of twenty-one experiments involving lessons on lightning, braking systems, electric motors, math problems, electrical engineering, and botany, people who received graphics with concurrent spoken commentary performed better on problem-solving transfer tests than did people who received graphics with concurrent printed text. The median effect size favoring the spoken group was $d = .97$, which is considered a large effect. In another review involving forty-three experimental tests, Ginns (2005) also found consistent support for the modality principle. There are, of course, limitations on the applicability of each principle, including this one. In particular, the modality principle is not likely to apply in situations in which the text is long and complex, has technical terms or symbols, and is not in the learner's native language, nor when the material is presented at a slow rate, is paced by the learner, or is already familiar to the learner. In addition, learning will suffer if the narration and animation are out of sync.

Applications of the Modality Principle

Many practitioners believe that in multimedia instruction it's best to explain online visuals with words in text and audio to accommodate different learning styles or to give learners a double opportunity to process the information—one in text and a second via audio. Previously, we showed evidence that using two redundant expressions of the same content—one in audio and one in text—overloads memory. We also recommended that you make audio the default option but allow learners to turn off the audio and select text. For example, in Figure 9.5 we show a screen from an animated multimedia lesson demonstration of how to use Adobe Acrobat software to review and comment on documents. The instructor explains the demonstration with audio narration, not with text. However, should the learners need a textual display, they can turn off the audio and see text displayed in captions placed close to the application area to apply the contiguity principle. Authoring software such as Adobe Captivate that automatically captures your on-screen movements and your narration can easily accommodate the modality principle when you deselect the "automatic" captions option.

When applying the modality principle with animations, keep your scripts succinct in alignment with the coherence principle. Many learners grow impatient listening to lengthy narrations on each screen. When narrations are lengthy, learners may opt for text so they can quickly scan the content and thus lose the psychological benefits of audio. Also, when the interface is complex, apply the signaling principle by using circles, arrows, or highlights to draw attention to the portion of the interface being described. Finally, break demonstrations into fairly short chunks to apply the segmenting principle. Let

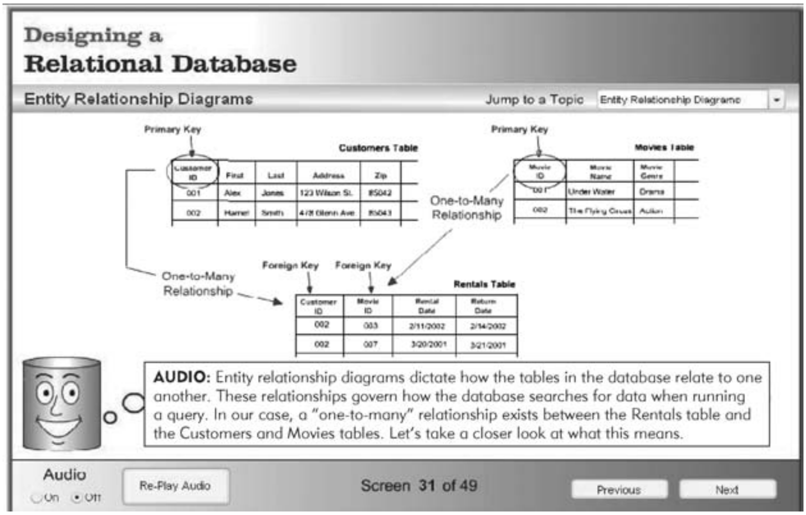


Figure 9.5 Audio Is Used to Describe Steps in a Software Demonstration.

the learner view and hear a brief segment and then press a continue button to move to the next segment.

PRINCIPLES FOR FOSTERING GENERATIVE PROCESSING

The foregoing two sections focused on techniques for avoiding cognitive overload by reducing extraneous processing and managing essential processing. In this section, we explore a different problem: People have cognitive capacity available but do not use it to make sense of the presented material, that is, they do not engage sufficiently in generative processing. Two principles for fostering generative processing are the multimedia and personalization principles.

APPLYING THE MULTIMEDIA PRINCIPLE

The conventional way to provide instruction is to use words—either in printed form (such as in a book or an online entry) or in spoken form (such as a lecture). Is there anything to be gained by supplementing words with corresponding pictures (in the form of illustrations, graphs, video, animation, etc.)? According to the multimedia principle, the overwhelming answer is yes. The multimedia principle is: People learn better from words and pictures than from words alone. The multimedia principle is a primary principle of instructional design because most of the other principles take it as their starting point and go on to deal with how best to create multimedia training materials (Fletcher & Tobias, 2005).

Rationale for the Multimedia Principle

The multimedia principle has its roots in dual coding theory (Paivio, 1986, 2007), which holds that people have separate information processing systems for words and pictures. According to the cognitive theory of multimedia learning summarized in Figure 9.1, people engage in generative processing when they make connections between corresponding words and pictures. For example, a learner may hear the narrator say, ''negative ions fall to the bottom of the cloud,'' and look at a portion of the animation depicting little circles with negative signs in them moving from the top to the bottom of the cloud. By mentally integrating the verbal and visual representations, the learner develops deeper understanding of what they both mean. The main idea behind the multimedia principle is that the mental act of integrating corresponding words and pictures is a generative activity—one that produces deeper understanding. A contrasting view—the information equivalence view—is that when pictures provide the same information as words, the pictures become redundant, that is, the pictures add nothing to the words.

Evidence for the Multimedia Principle

In a review, Mayer (2001) identified nine experiments comparing the problem-solving transfer test performance of people who learned from words and people who learned from words and pictures. The training materials covered pumps, brakes, lightning, or electrical generators and were presented either on paper or on a computer screen. Overall, people who received words and pictures performed better than did people who received words alone, with a median effect size of $d = 1.50$, which is considered large. More recently, Fletcher and Tobias (2005) conducted a review and also concluded that the research evidence supports the multimedia principle. However, not all graphics are equally effective so you need to be concerned with what makes an effective graphic (Clark & Lyons, 2004). For example, contrary to popular belief, a series of static diagrams can be as effective or even more effective than an equivalent animation (Betrancourt, 2005; Mayer, Hagarty, Mayer, & Campbell, 2005).

Applications of the Multimedia Principle

Text is fast and easy to produce. Also, in online environments, text uses less bandwidth than visuals, especially complex visuals such as animations. Many practitioners do not have access to graphic support beyond what they can find in clip art. Therefore, many instructional environments—both print and online—fail to leverage the potential of visuals. Either they rely heavily on text or they use decorative types of visuals that enliven the slide or page but do not

Table 9.1 Types of Graphics for Learning

Graphic Type	*Description*	*Examples*
Decorative	Visuals added for aesthetic appeal or for humor	1. A person riding a bicycle in a lesson on how a bicycle pump works; 2. Baseball-related icons as a game theme in a lesson on product knowledge
Representational	Visuals that illustrate the appearance of an object	1. A photograph of equipment in a maintenance lesson; 2. A screen capture
Organizational	Visuals that show qualitative relationships among content	1. A matrix such as this table; 2. A concept map; 3. A tree diagram
Relational	Visuals that summarize quantitative relationships	1. A bar graph or pie chart; 2. A map with circles of different sizes representing location and strength of earthquakes
Transformational	Visuals that illustrate changes in time or over space	1. An animated demonstration of a computer procedure; 2. A video of how volcanoes erupt; 3. A time-lapse animation of seed germination
Interpretive	Visuals that make intangible phenomena visible and concrete	1. Drawings of molecular structures; 2. A series of diagrams with arrows that illustrate the flow of blood through the heart

Adapted from Clark and Lyons, 2004

contribute to learning. Here we offer a couple of proven guidelines about visuals:

1. **Use Explanatory Visuals.** Table 9.1 summarizes six common types of visuals found in instructional materials. Of the six types, we recommend that you make heavy use of the last four, which we consider explanatory. Explanatory visuals illustrate various types of relationships among ideas in your content. For example, the visual from a pharmaceutical website shown in Figure 9.6 shows an invisible process—the attack of a cell by the AIDS virus. Visuals that illustrate invisible or abstract content, what we call interpretive visuals, can dramatically enhance the understanding of a technical process such as this one.

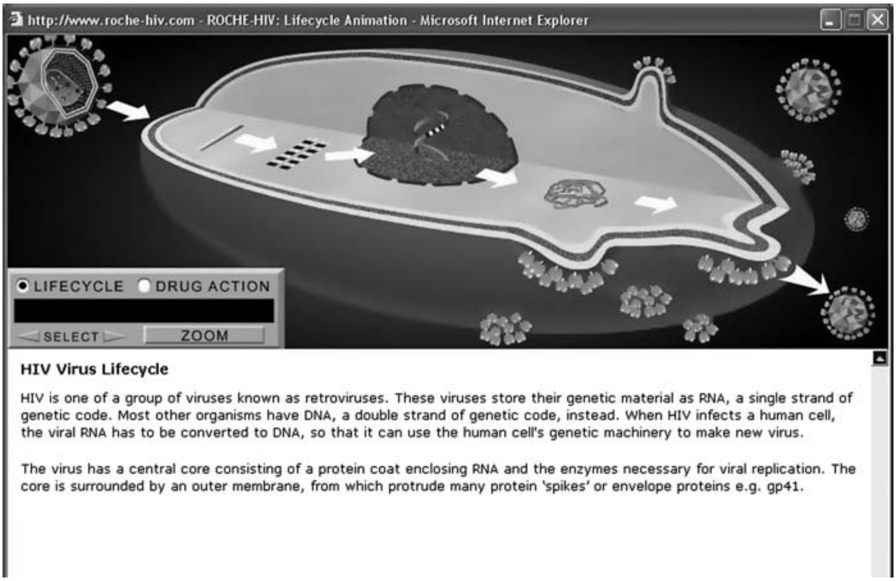


Figure 9.6 A Graphic Illustrating the Process of AIDS Infection.
With permission of Roche, Basel, Switzerland. http://www.roche-hiv.com/front.cfm

2. **Keep visuals simple.** You usually do not need high-end, three-dimensional realistic visuals to successfully illustrate your content. As we mentioned above, a series of static visuals to illustrate processes such as how a toilet tank flushes proved more effective than an animated version. Along similar lines, Butcher (2006) found better learning of blood circulation from simpler line drawings than from realistic images. In many cases, higher degrees of visual fidelity add extraneous mental load and at the same time do not illustrate the relationships any better than a simpler rendition.

APPLYING THE PERSONALIZATION PRINCIPLE

Social cues can also play an important role in encouraging people to work harder to make sense of the material in a training lesson, that is, in fostering generative processing. The personalization principle is that: People learn better when the words in multimedia training materials are presented in conversational style rather than formal style. You can create conversational style by using first- and second-person constructions (using ''I'' and ''you'') rather than third-person constructions, and by directly addressing the learner (saying, ''Let me ask you a question''). For example, in a narrated animation on how the human respiratory system works, formal style involves using third-person constructions such as, ''During inhaling, the diaphragm moves down, creating more space for the lungs, air enters through the nose or mouth, and moves down through the throat and bronchial tubes to tiny air sacs in the lungs.'' To personalize this speech, we can convert ''the'' to ''your'' throughout the lesson, such as, ''During inhaling, your diaphragm moves down, creating more space for your lungs, air enters through your nose or mouth, and moves down through your throat and bronchial tubes to tiny air sacs in your lungs.'' As another example, a portion of an explanation of plant growth in formal style is, ''In very rainy environments, plant leaves have to be flexible so they are not damaged by rainfall,'' whereas a conversational version is, ''This is a very rainy environment and leaves of your plant have to be flexible so they're not damaged by rainfall.''

Rationale for the Personalization Principle

How does personalization work? According to Mayer (2005d), social cues such as personalization prime a social response in the learner so that he or she accepts the instructor as a conversational partner who is trying to explain something. When you feel that you are in a conversation, you try harder to make sense out of the material (Nass & Brave, 2005; Reeves & Nass, 1996). When you try harder to understand the material, you engage in generative processing.

Evidence for the Personalization Principle

There is evidence that personalization can improve learning from instructional materials. In a review of ten experiments involving both spoken and printed text in multimedia lessons on lighting, botany, and the human respiratory system, Mayer (2005d) found that people who learned with conversational text performed better on problem-solving transfer tests than did people who learned with formal text. The median effect size was $d = 1.30$, which is a large effect. There is also preliminary evidence that people learn better when online tutors speak in a human voice rather than in a machine-synthesized voice (see the voice principle; Mayer, 2005d) and when online tutors use polite wording for suggestions and feedback rather than direct wording (see the politeness principle; Wang, Johnson, Mayer, Rizzo, Shaw, & Collins, 2005).

Applications of the Personalization Principle

As we summarized in the previous section, applying the personalization principle can greatly enhance learning and, in some cases, personalizing your materials takes only a few minor adjustments. In this section we offer two tips for personalizing your training.

1. **Use Conversational Language.** When designing training materials such as printed handouts for instructor-led classes or online screens for

virtual classes, engage your learners by using first- and second-person language. For example, Clark (2008) recommends that you use an informal format to present learning objectives to the students. For example, consider a formally stated objective, such as ''*Given a calculation goal, supporting data, and a spreadsheet, the learner will enter the correct formula to accurately achieve the goal with no errors.*'' When you are ready to write the learner materials, convert this type of objective to something more reader-friendly, such as ''*You will learn how to construct and enter a formula into a spreadsheet to achieve a calculation goal.*'' In this revision we changed ''the learner'' to ''you,'' thereby speaking directly to the participant. In addition, we shortened the objective to make it more readable. For another example, Moreno and Mayer introduce an online game with the following formal language: ''*This program is about what type of plants survive on different planets.*'' A more conversational version stated: ''*You are about to start a journey where you will be visiting different planets*.''

To apply this principle, you only need to add a few first- and second-person constructions to make your scripts conversational. We have had clients report that their corporate communications departments opposed use of informal language since it did not meet official standards. We recommend that you review this evidence with the stakeholders and collaborate to adjust standards in ways that accommodate human learning processes.

2. **Add Online Agents.** In multimedia courses, you can add screen-characters such as the little ''database'' character shown in the lower-left corner in Figure 9.5. Be sure that your agent serves some useful instructional purpose such as explaining an example or giving feedback and is not just a decorative graphic. To apply the modality principle, present the agent's comments in audio narration using a native language voice. To apply the redundancy principle, avoid using audio and text that repeats the audio at the same time. We are just beginning to learn more about what agent features do and do not affect learning. Evidence to date suggests that the image is not that important. Instead, you should invest effort in what the agent says to ensure it serves a useful instructional role.

SUMMARY

As you can see, a seemingly passive instructional medium—receptive learning environments—can lead to active cognitive learning. What is needed to promote active learning is an instructional method that guides appropriate cognitive processing during learning, namely, selecting relevant information, organizing it into a coherent cognitive structure, and integrating it with other knowledge. To accomplish this goal, instructional designers should be sensitive to the cognitive load required by the instructional materials. In particular, you should seek training materials that minimize extraneous processing (in line with the coherence, signaling, redundancy, and contiguity principles), manage essential processing (in line with the segmenting, pretraining, and modality principles), and promote generative processing (in line with the multimedia and personalization principles).

We started our discussion with several options regarding how to best disseminate important information about the features and benefits of a new product release to a global sales force. We mentioned several alternatives, including a pod-cast (audio only), a web page with text and visuals, and a recording of a virtual presentation that uses visuals and audio. All of these options represent receptive learning environments because they do not provide for overt learner interaction. However, all of them have the potential for learning by promoting *mental activity during learning*.

Because pod-casts use audio only, they lose the benefits of the multimedia principle—the power of visuals. If the product features and benefits involve content that is visual in nature such as software interfaces, equipment components, or other concrete features, the pod-cast may not be as useful as an alternative environment that can display visuals. The website page applies the multimedia principle by adding visuals. If it applies the contiguity and coherence principles in construction and display of the text in conjunction with relevant visuals, you may get a good learning outcome. The virtual class recording allows you to take advantage of the multimedia and modality principles if effective visuals are explained by concise and relevant audio narration. Because most recorded lessons play a continuous stream of content, you may want to create a series of quite short modules, each of which focuses on a specific feature or benefit.

Our point is that all delivery media have tradeoffs. Your challenge is to consider the constraints of your technical environment and learners and provide the optimal mix of media and instructional methods to make new content easily accessible and to promote mental processing of the content. Whatever delivery medium you use, you should strive to apply the proven principles we have summarized in this chapter.

References

Betrancourt, M. (2005). The animation and interactivity principles in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 287–296). New York: Cambridge University Press.

Boucheix, J. (2008). Young learner's control of technical animations. In R. Lowe & W. Schnotz (Eds.), *Learning with animation* (pp. 208–234). New York: Cambridge University Press.

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology*, *98*(1), 182–197.

Campbell, J. P., & Oblinger, D. G. (2007). Top-ten teaching and learning issues, 2007. *Educause Quarterly*, *30*(3), 15–22.

Clark, R. C. (2009). *Building expertise* (3rd ed.). San Francisco: Pfeiffer.

Clark, R. C. (2008). *Developing technical training* (3rd ed.). San Francisco: Pfeiffer.

Clark, R. C., & Kwinn, A. (2007). *The new virtual classroom*. San Francisco: Pfeiffer.

Clark, R. C., & Lyons, C. (2004). *Graphics for learning: Proven guidelines for planning, designing, and evaluating visuals in training materials*. San Francisco: Pfeiffer.

Clark, R. C., & Mayer, R. E. (2008). *e-Learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning* (2nd ed). San Francisco: Pfeiffer.

Clark, R. C., & Mayer, R. E. (in press). Learning by viewing versus learning by doing. *Performance Improvement*.

Clark, R. C., Nguyen, F., & Sweller, J. (2006). *Efficiency in learning: Evidence-based guidelines to manage cognitive load*. San Francisco: Pfeiffer.

Clark, R. E. (2001). *Learning from media*. Greenwich, CT: Information Age Publishing.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920*. New York: Teachers College Press.

Fletcher, J. D., & Tobias, S. (2005). The multimedia principle. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 117–134). New York: Cambridge University Press.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, *15*, 313–331.

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and the temporal contiguity effects. *Learning and Instruction*, *16*, 511–525.

Haidet, P., Morgan, R. O., O'Malley, K., Moran, B. J., & Richards, B. F. (2004). A controlled trial of active versus passive learning strategies in a large group setting. *Advances in Health Sciences Education*, *9*, 15–27.

Harp, S. F., & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology*, *89*, 92–102.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, *90*, 414–434.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*, 75–86.

Mayer, R. E. (2001). Multimedia learning. New York: Cambridge University Press.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist*, *59*, 14–19.

Mayer, R. E. (2005a). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). New York: Cambridge University Press.

Mayer, R. E. (2005b). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 183–200). New York: Cambridge University Press.

Mayer, R. E. (2005c). Principles for managing essential processing in multimedia learning: Segmenting, pretraining, and modality principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 169–182). New York: Cambridge University Press.

Mayer, R. E. (2005d). Principles for multimedia learning based on social cues: Personalization, voice, and image principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 201–212). New York: Cambridge University Press.

Mayer, R. E. (2008). *Learning and instruction* (2nd ed.). Upper Saddle River, NJ: Pearson Merrill Prentice Hall.

Mayer, R. E., Bove, W., Bryman, A., Mars, R., & Tapangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of Educational Psychology*, *88*(1), 64–73.

Mayer, R. E., Hagarty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia learning. *Journal of Experimental Psychology: Applied*, *11*, 256–265.

Mayer, R. E., & Jackson, J. (2005). The case for coherence in scientific explanations: Quantitative details can hurt qualitative understanding. *Journal of Experimental Psychology: Applied*, *11*, 13–18.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, *38*, 43–52.

Moreno, R., & Mayer, R. E. (2000). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology*, *92*(1), 117–125.

Nass, C., & Brave, S. (2005). *Wired for speech: How voice activates and advances the human-computer relationship*. Cambridge, MA: MIT Press.

O'Neil, H. F. (Ed.). (2005). *What works in distance learning*. Greenwich, CT: Information Age Publishing.

O'Neil, H. F., & Perez, R. S. (Eds.). (2006). *Web-based learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.

Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach*. Mahwah, NJ: Lawrence Erlbaum Associates.

Reeves, B., & Nass, C. (1986). *The media equation: How people treat computers, television, and new media like real people and places*. New York: Cambridge University Press.

Schar, S. G., & Zimmermann, P. G. (2007). Investigating means to reduce cognitive load from animations: Applying differentiated measures of knowledge representation. *Journal of Research on Technology in Education*, *40*, 64–78.

Shavelson, R. J., & Towne, L. (Eds.). (2002). *Scientific research in education*. Washington, DC: National Academy Press.

Sweller, J. (1999). *Instruction in technical areas*. Camberwell, Australia: ACER Press.

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). New York: Cambridge University Press.

Thorndike, E. L. (1906). *Principles of teaching based on psychology*. New York: Seiler.

Wang, N., Johnson, L., Mayer, R. E., Rizzo, P., Shaw, E., & Collins, H. (2005). The politeness effect: Pedagogical agents and learning gains. In C. Looi, G. McCalla, B. Bredeweg, & J. Breuker (Eds.), *Artificial intelligence in education: Supporting learning through intelligent and socially informed technology* (pp. 668–693). Amsterdam: IOS Press.

CHAPTER TEN

Instructional Strategies for Directive Learning Environments

Ruth Colvin Clark
Richard E. Mayer

Suppose you were asked to develop a course for underwriters on how to qualify loan applicants. Because underwriters are hired at various times in small numbers in regional offices, you recommend an asynchronous e-learning course. You consider several ways to design the course. You could prepare a presentation to be recorded and delivered through your virtual classroom technology. The presentation would reflect a *receptive learning environment* as discussed in Chapter Nine. In receptive learning environments such as a recorded presentation, content is presented with few opportunities for learners to respond overtly.

Alternatively, you could identify the main tasks involved in researching and qualifying loan applicants and organize them into brief lessons sequenced from easier to more difficult. Each lesson starts with key concepts and then teaches the steps. For example, in a lesson on evaluating credit reports, the initial topics explain major credit concepts and assign practice on those concepts. Then the lesson presents guidelines for interpreting credit reports with practice on evaluating sample credit reports. We refer to this type of design as a *directive learning environment*.

In a third approach, you could start the training with a case study of a loan applicant and present a simulated environment allowing access to varied information about the applicant including credit reports, financial statements, and references, to name a few. In this approach, the learner is in charge of deciding what to research and the best research sequence. The learner selects a loan funding option from a menu and receives feedback from the loan committee that approves or rejects the recommendations. The learner can work through the loan application simulation multiple times, seeing the effects of different actions on the accuracy of the final decision. We refer to learning environments that allow learners a great deal of freedom to try different actions and learn from experience as *guided discovery environments*. In Chapter Nine we focused on receptive learning environments. In this chapter we examine directive approaches and contrast them to receptive learning environments summarized in Chapter Nine, and to guided discovery approaches discussed in Chapter Eleven.

We will begin with an introduction in which we define key terms and issues and review a theory of learning presented in Chapter Nine based on the distinction among extraneous processing, essential processing, and generative processing during learning. Then we review guidelines and evidence for instructional methods in directive learning environments that reduce extraneous processing, manage essential processing, and foster generative processing.

WHAT ARE DIRECTIVE LEARNING ENVIRONMENTS?

Directive learning environments are characterized by (1) a linear sequence of lessons that teach prerequisite knowledge and skills first, (2) short lessons with brief content presentations in the form of explanations or demonstrations, (3) frequent learner response opportunities to practice exercises, and (4) immediate feedback to learner responses. Directive course designs are also called *rule-example-practice* or *tell-show-do* learning environments. Directive designs differ from receptive approaches in the inclusion of frequent practice exercises accompanied by immediate feedback. Receptive approaches, in contrast, present information but do not incorporate opportunities for overt learner responses.

For example, Figure 10.1 shows the menu from a directive course that teaches how to use an electronic spreadsheet. The lessons start with basic content, including naming cells, entering data, using formulas, and navigating in the spreadsheet; and then move to more complex skills involving absolute references, pivot tables, and report formatting. In each lesson, the learner views a brief narrated explanation followed by practice exercises. If the learner gives an incorrect answer, she receives explanatory feedback and is asked to try again. With a second incorrect response, the program provides the correct answer.

Three Views of Learning

Over the past one hundred years, learning psychologists have evolved three different views of learning and instruction, as summarized in Figure 10.2. One

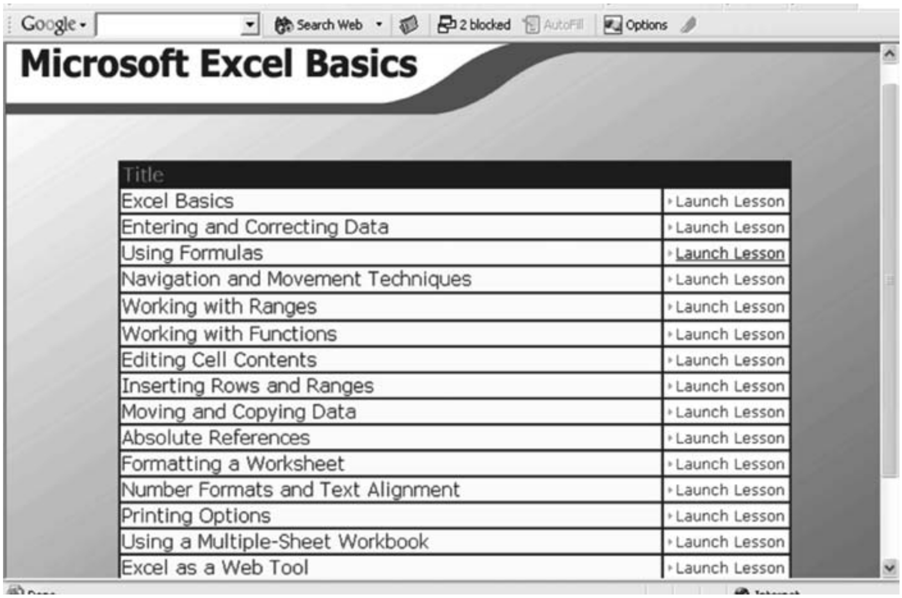


Figure 10.1 A Directive Course Menu.

With permission from Clark, Nguyen, and Sweller, 2006

early view that is prevalent among many practitioners today is *information acquisition*. In this view the role of the instructor is primarily to provide information, and the role of the learner is to absorb it. Teaching according to the information acquisition view involves transmission of content. Learning is believed to occur by adding new content to memory, a bit like a sponge absorbs water. For example, some lessons that rely heavily on PowerPoint slide presentations reflect an information acquisition view of learning.

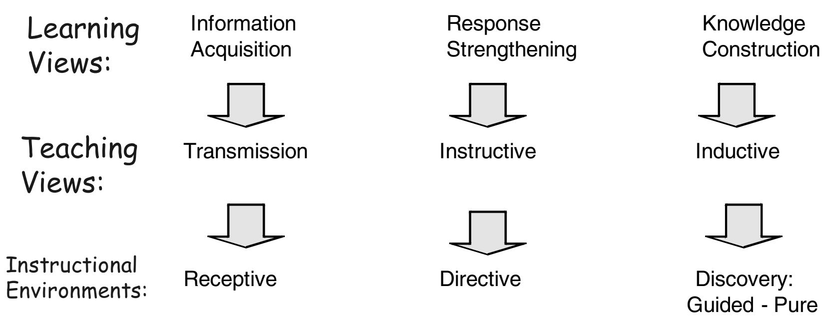


Figure 10.2 Three Views of Learning and Instruction.

Reflecting the influence of behaviorism, a *response strengthening* view predominated training designs through much of the first half of the 20th century. In the response strengthening view, the role of the instructor is to provide the learner with brief content explanations, questions, and feedback, and the role of the learner is to strengthen mental associations through responses to questions. Teaching based on a response strengthening view of learning uses an instructive approach in which the training is highly guided and focuses on helping learners respond correctly to practice questions. Much procedural training today, such as online software training developed with tools like Adobe Captivate, reflect a response strengthening view of learning.

Most recently, learning psychologists have adopted a *knowledge construction* view of learning. In this view, the role of the instructor is to provide an environment that helps learners build new mental models in memory, and the role of the learner is to actively engage with the learning environment to build new mental models. Some implementations of a knowledge construction perspective see learning as an inductive event during which learners engage in various experiences such as solving problems or viewing multiple examples from which they derive new knowledge and skills. Inductive instructional environments can provide high or low levels of structure and guidance (frequently referred to as scaffolding) to form either guided discovery or pure discovery learning environments.

On the surface, it appears that the information acquisition view can be associated with receptive learning environments, and the response-strengthening view was the historical source of directive learning environments. However, on closer inspection we can conclude that all learning requires an active knowledge construction process on the part of the learner. Clark (2008b) and Clark and Mayer (2009) distinguish between *implicit* and *explicit* instructional methods to help learners actively build new mental models. Table 10.1 summarizes these methods. Implicit methods promote mental models *in the absence of external activity*. For example, graphics help learners build mental models by forming

Table 10.1 Implicit vs. Explicit Instructional Methods to Build Mental Models

Method Type	*Description*	*Examples*
Implicit	Training techniques that promote building of mental models in the *absence of* visible physical activity	Graphics, Examples, Analogies
Explicit	Training techniques that promote building of mental models through physical activity that leads to productive mental activity	Practice exercises, Discussions, Case studies

two mental codes—a visual one and a verbal one. In Chapter Nine we described how receptive learning environments can use various implicit instructional methods to promote the mental activities that lead to learning. In contrast, explicit instructional methods help learners build mental models through external activities, such as responding to questions or participating in discussions. In this chapter we will look at ways that directive learning environments can use both implicit and explicit methods to promote the mental activities that lead to learning. Because we discussed implicit methods in Chapter Nine, we will emphasize explicit methods in this chapter.

HOW LEARNING HAPPENS

In Chapter Nine we reviewed the main processes involved in learning, including (1) selecting words and images, (2) organizing words and images into coherent representations, and (3) integrating verbal and pictorial representations with each other and with existing knowledge (Mayer, 2005). As a result of these processes, learners construct knowledge that enables them to transfer knowledge and skills acquired in the learning setting to situations and environments outside of the instructional environment. In workforce learning, transfer of learning means that learners build new mental models that they can apply later on the job.

The selecting, organizing, and integrating processes all occur in our two memory systems responsible for learning and thinking: working memory and long-term memory. While the actions of selecting, organizing, and integrating occur in working memory, the goal of these activities is to build mental models in long-term memory that serve as a permanent repository of your knowledge and skills. Working memory is the center of conscious processing, including the activities that lead to learning. However, its capacity for storage is very limited. Therefore, you must manage cognitive load during learning. In Chapter Nine we defined three forms of cognitive load: extraneous, essential, and generative. To briefly recap:

1. *Extraneous cognitive load* comes from mental work that is irrelevant to the learning goal and drains working memory resources in unproductive ways. Much extraneous cognitive load originates from suboptimal use of instructional modes such as displaying text in a distant location from a visual being described.
2. *Essential cognitive load* comes from the complexity of the content and the learning goal. For example, a goal to identify a correctly formatted Excel formula is less complex than a goal to construct and input an Excel formula into a spreadsheet to solve a problem. Of course, neither of these goals is challenging to someone experienced with Excel. Thus

essential cognitive processing will vary according to the learning task and to the background knowledge of the learner.

3. *Generative cognitive load* originates from mental work that helps learners achieve the learning goal. Generative cognitive processing promotes the selecting, organizing, and integrating processes essential to learning and transfer. For example, when reviewing a demonstration of entering a formula into a spreadsheet, a learner must select the relevant part of the spreadsheet, organize the graphics and words to form a coherent mental representation, and integrate this representation with prior knowledge from long-term memory to build an expanded mental model.

Your job as an instructional professional is to minimize extraneous cognitive load to free up working memory resources for essential and generative processing. For example, as you prepare a demonstration to illustrate how to construct and input a formula into Excel, you minimize extraneous load by placing explanatory text into the spreadsheet close to where the formula appears. In this chapter we will focus on how to manage essential, extraneous, and generative cognitive load in order to promote learning in directive learning environments.

Cognitive Load and Directive Learning Environments

As a result of their behaviorist legacy, directive learning environments incorporate many features that minimize extraneous cognitive processing, manage essential cognitive processing, and optimize generative processing. First, they manage essential load by breaking content into bite-size pieces and organizing them into a series of short lessons that teach skills in a prerequisite order. In multimedia lessons, learners control their access through the content displays using navigational devices such as continue and back buttons. Second, directive learning environments encourage generative cognitive processing through frequent practice opportunities that follow content explanations. The practice exercises are highly guided with immediate corrective feedback following learner responses. Third, extraneous cognitive load is minimized by applying many of the techniques we summarized in Chapter Nine. For example, explanations and visuals are limited to just those related to the learning objectives (Coherence Principle), and explanatory text is placed close to the relevant sections of the visuals (Contiguity Principle).

EVIDENCE-BASED PRACTICE

One way to ensure a return on the more than $100 billion invested annually in workforce learning is to adopt instructional strategies that are rooted in evidence. Establishing a culture of evidence and translating learning research into practice are the number one and three priorities of the 2007 Educause advisory committee (Campbell & Oblinger, 2007). We recommend that you consider research on instructional methods proven to support learning as one factor in your course design, development, and selection decisions. We prefer experimental research evidence as the most valid indicators of instructional methods that work. In experimental research, learners are randomly assigned to two or more versions of a lesson that are the same except for the one variable under study. For example, one lesson uses text only to present content, while a comparison lesson uses text and graphics. After studying their assigned lesson, all learners are tested with the same test, and the average scores are submitted to statistical analysis to determine whether any outcome differences are real and practically significant. In this chapter, we review experimental evidence regarding the main features that characterize directive instructional environments, including:

- Segmenting and sequencing of content,
- Five laws of practice,
- Leveraging examples, and
- Instructional methods to support learning transfer.

As we review the evidence on these features, we will suggest application principles previewed in Table 10.2.

Table 10.2 Proven Principles for Design of Directive Learning Environments

Principle	*Description*
Segmenting and Sequencing	Break content into short segments; allow learners to progress at their own pace; sequence supporting concepts prior to major lesson content.
Deep Processing	Design practice activities that promote elaborative mental rehearsal that leads to deep understanding; avoid regurgitation exercises.
Practice Amount	Determine the number of practice exercises based on degree of initial proficiency required by the job.
Spaced Practice	Distribute practice exercises throughout a lesson and among lessons rather than together in time or place.
Explanatory Feedback	Include tailored feedback for all correct and incorrect response options that explain why an answer is right or wrong.

(*Continued*)

Table 10.2 (*Continued*)

Principle	*Description*
Learning-Goal Feedback	Design feedback that shows progress over time to achieving the learning objective rather than a comparison of one's progress with others.
Task-Focused Feedback	Provide feedback that focuses learners' attention to the task and away from themselves.
Practice Challenge	Tailor practice exercises to offer appropriate challenge that will extend performer expertise.
Worked Examples	Replace some practice with faded worked examples; include self-explanation questions adjacent to worked out steps.
Identical Elements	For near transfer learning, construct examples and practice that mirror the application (work) environment.
Varied Context Examples	For far transfer learning, construct a series of examples that vary surface features but hold the underlying principles constant.

SEGMENTING AND SEQUENCING IN DIRECTIVE LEARNING ENVIRONMENTS

Directive environments break content into small chunks and sequence those chunks in a prerequisite order so the learner progressively builds a knowledge base by first understanding basic concepts and skills and then moving to more advanced. For example, in Figure 10.1 we showed a typical directive lesson sequence in which Excel basics are taught in the initial lessons followed by more advanced tasks. Within directive lessons, content is broken into brief topics and basic concepts sequenced prior to the lesson task. For example, in an Excel lesson on formulas, following an introduction, the topics of cell references and formula formats are presented prior to the procedures of inputting a formula into a spreadsheet cell.

While the behaviorist goal for segmenting and sequencing was to help learners grow stimulus-response chains of knowledge, a more contemporary view suggests that these design features are effective in managing essential cognitive load. Recall that essential cognitive load refers to the inherent complexity of a learning objective and its associated content. When your content and objectives are complex with many tasks involving subtasks, processes, and concepts, one way to reduce the load is to break content into small pieces and sequence those pieces from less to more complex. This hierarchical approach to segmenting and sequencing content is widespread in practice in academic and workforce learning. But do we have evidence that it works? In the next section, we review evidence for segmenting content as well as for sequencing prerequisite topics first.

Evidence for Segmenting

Mayer and Chandler (2001) compared learning from a continuous 2.5-minute lesson on how lightning forms with a segmented lesson in which the same content was presented in approximately ten-second clips. After viewing each brief clip, learners in the segmented versions advanced the lesson by clicking on the continue button. In three different experiments using different content, segmented versions resulted in better transfer learning with a median effect size of 1, which is high (Clark & Mayer, 2008). Essentially, the segmented versions presented short chunks of information and allowed the learner to determine the pace at which he would access each chunk. Hasler, Kersten, and Sweller (2007) found similar benefits for segmented (learner controlled pacing) over continuous (instructional controlled pacing) lesson animations.

Evidence for Prerequisite Sequencing

By prerequisite sequencing, we mean teaching related concepts in a lesson prior to the main task. Mayer, Mathias, and Wetzell (2002) compared learning how a hydraulic braking system works among learners who did and did not view training on the names and actions of each part *prior to viewing the complete explanation*. The pretraining was presented either in print, in which a diagram showed labeled parts, or via multimedia, in which a learner could click on a part and see its name while viewing its motion. Learners who received pretraining and thus were familiar with the individual part names and actions learned more from the full lesson than those who did not have pretraining, with an effect size of .9, which is large.

Kester, Kirscher, and van Merriënböer (2006) also found a learning advantage to sequencing concepts separately from procedures in lessons teaching troubleshooting of malfunctioning electrical circuits. Their lessons included a concept portion with topics such as ''What is voltage?'' and ''What is amperage?,'' a procedural portion with topics such as ''how voltmeters are connected in circuits'' and ''how power flows from poles in the power source,'' along with ten practice tasks. The sequence of the concept and procedure topics was varied so that some lessons taught the concept segments followed by the procedural segments, some taught procedural segments followed by concept segments, and some presented both segments together. They found that segmented versions were more effective, concluding that: ''It is clearly better to teach declarative information and procedural information piece by piece instead of simultaneously'' (p. 181).

In summary, we offer a Segmenting and Sequencing Principle for directive learning environments: *Break content into short segments allowing learners to progress from one to the next at their own pace; sequence basic concepts prior to major processes or tasks*.

Linear Versus Non-Linear Sequencing

Because directive learning environments foster a gradual building of new knowledge and skills from more basic to more complex, a linear sequence of lessons is implied. In a typical directive course, each lesson is intended to be taken one after the other. In multimedia learning, it is common practice to give learners navigational control over the sequence of lessons or topics. However, the default layout in directive designs is linear. In contrast, guided discovery environments often initiate a lesson with a design task or a realistic work problem and offer learners many routes to access relevant data, learn related content, and complete the assignment. For example, in Figure 10.3 you can see the interface of a guided discovery course designed to teach loan analysis. To make a decision about a loan applicant, the learner can click on various objects in the virtual office, including books, telephone, fax machine, and computer in any sequence. What evidence do we have for the benefits of a linear course design?

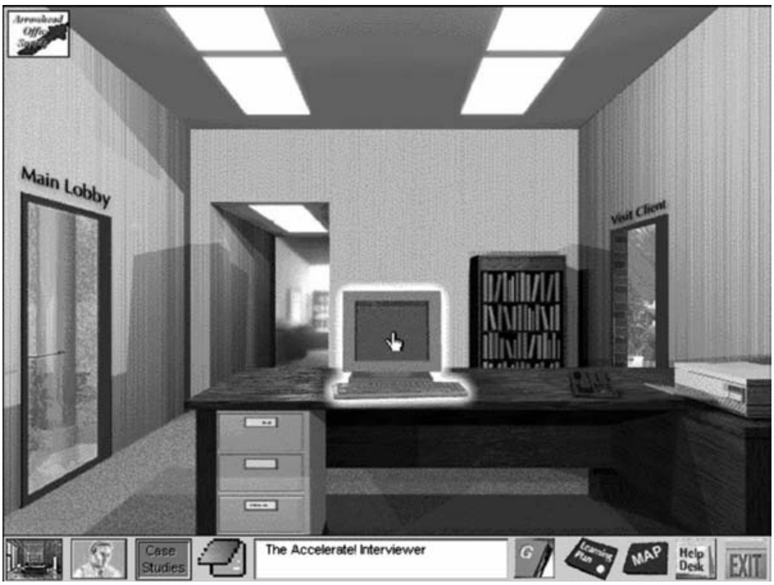


Figure 10.3 A Guided Discovery Approach to Loan Analysis Training.
With permission from Moody's Investment Service

Evidence for Instructional Control

Perhaps the most relevant evidence comes from research on learner control. In controlled experiments, one group of learners is required to follow a prescribed sequence of topics and lessons in a more or less linear sequence. We call this scheme *high instructional control*. In contrast, a second group of learners can select or skip the lessons and topics at their own discretion. We call this scheme *high learner control*. As mentioned above, although directive lessons delivered in multimedia typically allow learners to select various lessons, the default plan illustrated in Figure 10.1 uses a prerequisite linear sequence. Research comparing learning from instructional control with learning from high learner control has shown that instructional control leads to better learning by novices. High instructional control is more effective because (1) having to make content decisions adds extraneous cognitive load, which absorbs limited mental resources, and (2) most novice learners lack the background knowledge to make good decisions regarding the content sequences. Glenberg and his colleagues (1992) describe the poor self-assessments of novice learners as ''illusions of knowing.'' In contrast, learners with relevant prior knowledge can usually make good instructional decisions and can effectively exploit learning environments, such as the one shown in Figure 10.3, that allow learner control. Additionally, experienced learners may have unique learning goals that they can productively pursue in a highly learner-controlled environment.

We need more research and better quality research to guide our decisions regarding the circumstances under which high learner control is beneficial (Scheiter & Gerjets, 2007). However, based on evidence to date, Clark and Mayer (2008) recommend that your lessons present basic content before more advanced content as the default navigational scheme for novice learners.

FIVE EVIDENCE-BASED LAWS OF PRACTICE

One of the hallmarks of a directive learning environment is the regular placement of structured interactions, also called practice exercises, throughout the training. Moreno and Mayer (2007) define interactivity as a two-way action between learner and instruction in which the goal of the action is to help learners build new knowledge that supports the learning objective. Thus, actions such as clicking on the continue button or engaging in a game-show type activity would not qualify as interactions because these activities are unlikely to promote mental activity consistent with the outcome objective.

In this section we review five research-based guidelines for the design of practice; exercises that promote selecting, organizing, and integrating; as well as transfer of new knowledge and skills. As we discussed in Chapter Nine, learning requires active processing on the part of the learner. Receptive learning environments promote active processing through implicit methods that do not require visible activity on the part of the learner. In contrast, directive learning environments incorporate frequent exercises that require overt responses. The goal of the exercises is to support the mental activities of selecting, organizing, integrating new knowledge and skills to result in a transferable mental model.

1. Design Practice That Promotes Deep Understanding

Practice exercises can lead to rehearsal of new lesson content in working memory, resulting in integration of that content with existing knowledge from long-term memory. Exercises can prompt two types of rehearsal: maintenance or elaborative. Maintenance rehearsal involves the repetition of information. Repeating a telephone number until you dial it is an example of maintenance rehearsal. An exercise that asks the learner to state the meaning of the four formula operators in Excel is another example of a maintenance exercise. We refer to maintenance rehearsals as ''regurgitation'' practice. While maintenance rehearsal keeps content active in working memory, unless it is very extensive, it does not support integrating new content with existing content to form mental models. In other words, mere repetition of presented content leads to shallow learning.

In a classic experiment, Craig and Watkins (1973) manipulated the number of times participants repeated individual words in a long list of words. They then asked participants to recall the word list and compared recall frequency with the number of times the word had been repeated. There was little relationship between repetition frequency and later recall. In this experiment, the number of repetitions was relatively few: around twelve. As we will see below, when repetitions are extensive, learning can occur through automaticity.

Rather than maintenance rehearsal, we recommend that you design practice opportunities that support elaborative rehearsal that leads to understanding. Elaborative rehearsal requires generative cognitive processing. That is, elaborative rehearsal helps learners integrate new knowledge with existing knowledge to build new mental models. In workforce learning, elaborative rehearsal should require learners to respond in ways that they will respond on the job. For example, rather than asking learners to state the meaning of four operators in an Excel formula, a better exercise states a calculation goal and asks the learner to identify which of several formulas reflect the correct format to achieve the goal. This exercise requires the learner to process content more deeply in a manner that emulates how they will process it on the job.

To design elaborative practice exercises, be sure to incorporate the context of the job. For example, in airline gate agent training, a maintenance rehearsal question asks learners to restate the five reasons they can prohibit passenger boarding. This repetition of information already provided does not require any transformation of content by the learner and offers little opportunity for the organization or integration activities required for active building of new mental models. In contrast, an elaborative rehearsal presents the learner with five passengers. When the learner clicks on a passenger, she hears a brief dialog, after which she can drag the passenger onto the aircraft or onto the concourse. This interaction requires the learner to apply lesson guidelines to a work-relevant scenario. Although this interaction is a drag-and-drop format, it could as effectively be constructed as a multiple-choice or true-false item. The focus of this guideline *is not on the format of the practice* but on a design that encourages learners to process new information in a way that mirrors how it will be used on the job. Later in the chapter, we provide more detailed guidance with examples of maintenance and elaborative practice exercises that use common response formats such as multiple choice and short answer.

In summary, we recommend you apply the Deep Processing Principle: *Design practice exercises that promote elaborative mental rehearsal that builds deep understanding appropriate to the application environment.*

2. Adjust the Amount of Practice Based on Performance Criticality

Practice does make perfect, although at a diminishing rate of return. You get your biggest learning gains from the first few practice sessions. For example, Keehner and her associates (2006) measured the time novices needed to accurately complete a simulated surgical laparoscopic maneuver over a number of practice exercises. Figure 10.4 shows the average improvement over practice

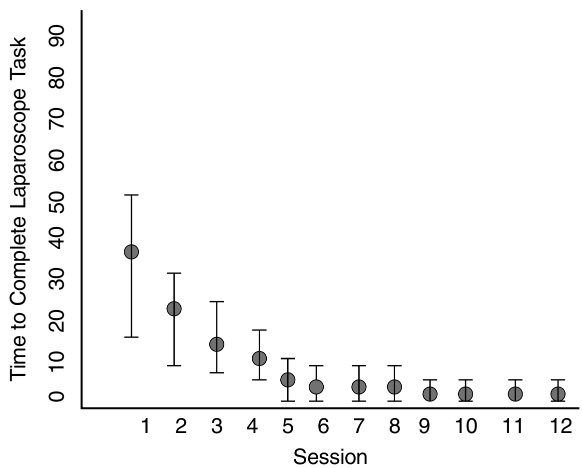


Figure 10.4 Performance Was Faster and More Consistent with More Practice.

Based on data from Keehner, Lippa, Montello, Tendick, and Hagarty, 2006 sessions as well as the average time distribution among subjects with each practice session. As you can see, average time to complete the task diminished rapidly over the first five practice sessions. After the fifth session, improvement was much smaller. These results reflect the well-established Power Law of Practice, stating that performance continues to improve with more practice but at a logarithmic rate. Thus, the greatest improvement occurs with the first few exercises. Notice also that the variations among average time to complete the task (illustrated by the vertical lines in Figure 10.4) decrease with each practice session. These data show that performance consistency improves greatly over the first few practice iterations.

Schnackenberg, Sullivan, Leader, and Jones (1998) compared learning from better and poorer learners (defined by grade point average) from e-learning with many (sixty-six) and fewer (twenty-two) practice exercises. They measured test scores as well as time to complete the training for low- and high-ability-level students randomly assigned to lessons with high and low numbers of practice. Both high- and low-ability learners assigned to lessons with more practice exercises had higher test scores. As expected, lessons with more practice required longer completion times. In accordance with the Power Law of Practice, the amount of improvement was *not directly proportional* to the amount of extra practice time invested. For example, poorer learners invested 75 percent more time for a 14 percent gain in scores.

To apply the power law of practice, consider the criticality of the performance accuracy needed immediately after training. In many cases, learners attain acceptable proficiency after a few practice sessions and gain greater proficiency through on-the-job experience. In contrast, some situations, such as tasks with safety consequences, demand high levels of skill proficiency the first time they are performed in the job setting. For these situations, the greater time invested to improve skill performance is justified. In short, consider the required degree of accuracy of task performance when deciding how much practice to include in a directive learning environment. Based on a legacy of experiments that reflect the Power Law of Practice, we recommend the following *Practice Amount Principle: Determine the number of practice exercises based on the degree of proficiency required by the job*.

When the situation requires high levels of task proficiency, you can use repetitive practice methods called *drill and practice*. After practicing a skill for many repetitions—numbering in the 100s—that skill becomes automatic. When procedural skills become automated, more cognitive capacity is available for the learner to engage in deeper processing of the material (Sweller, 1999). To make drill and practice more engaging, embed it in a computer game format in which rewards are assigned based both on accuracy and on speed of response. Once automated, learner responses become highly accurate, very rapid, and resistant to change.

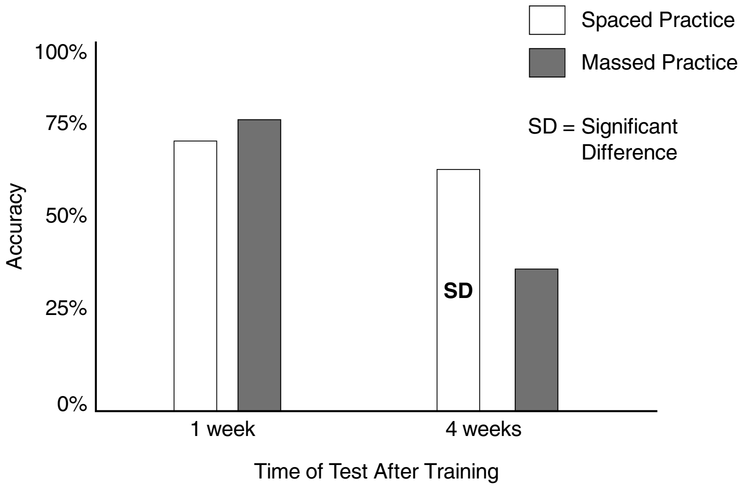


Figure 10.5 Better Long-Term Retention with Spaced-Out Practice.
Adapted from Rohrer and Taylor, 2006

3. Distribute Practice Throughout Learning Events

One way to get more mileage out of the same number of practice exercises is to distribute them throughout your lessons, rather than lumping them in a single lesson or in one place in a lesson. For example, Rohrer and Taylor (2006) compared two groups learning a mathematical procedure. Each group worked ten practice problems. One group completed all of the practice in a single session. The other group practiced five problems in week one and the remaining five problems the following week. As you can see in Figure 10.5, there were no differences in learning on an immediate test, but those in the spaced practice group had much better retention four weeks later. The authors recommend that, rather than placing a large number of practice problems of the same type at the end of a lesson, the practice items should be spread among several lessons. For example, ''each lesson is followed by the usual number of practice problems, but only a few of these problems relate to the immediately preceding lesson. Additional problems of the same type might also appear once or twice in each of the next dozen assignments and once again after every fifth or tenth assignment thereafter''(p. 1218).

Since most training programs do not measure learning after the training event, the value of spaced practice is rarely salient to training practitioners. However, based on the consistent evidence of long-term benefits of distributed practice, we recommend the *Spaced Practice Principle: Distribute practice exercises throughout a lesson and among lessons rather than placing them together in one time or place.*

4. Follow Practice Responses with Effective Feedback

Growing out of the response-strengthening view of learning, directive learning events stress immediate corrective feedback to practice responses. Historically, feedback was thought to serve as a reinforcement for correct responses or as a mechanism to reduce incorrect associations that were viewed as detrimental to learning. In the active constructive view of learning, we emphasize the cognitive value of feedback more than its reinforcement value. To gain the most benefit from feedback, we recommend providing feedback that is (1) explanatory rather than corrective, (2) learning-goal directed rather than normative, and (3) task-focused rather than self-focused.

Corrective Versus Explanatory Feedback. Corrective feedback informs the learner that a response was correct or incorrect. Many authoring programs include simple controls for providing corrective feedback. For example, after an incorrect response to a multiple-choice question, the program responds with a message: ''That is incorrect. Please try again.'' Similarly, in a drag-and-drop exercise, an icon dragged to an incorrect location ''snaps back'' to its original location. In contrast, explanatory feedback not only tells the learner whether he is right or wrong but also provides a rationale. This requires the instructional professional to construct a brief explanation for learner responses. In e-learning, the author must write explanations for correct and all incorrect options.

What evidence do we have for explanatory feedback? Moreno (2004) and Moreno and Mayer (2005) compared learning from two versions of a lesson on botany called Design-A-Plant. In the lesson, participants select roots, leaves, and stems to build a plant best suited to an imaginary environment. The goal is to teach the adaptive benefits of plant features for specific environments such as heavy rainfall, sandy soil, and so forth. In the research study, either corrective or explanatory feedback was offered by a pedagogical agent in response to a plant design. For explanatory feedback, the agent made comments such as: ''Yes, in a low sunlight environment, a large leaf has more room to make food by photosynthesis'' (for a correct answer) or ''Hmmm, your deep roots will not help your plant collect the scarce rain that is on the surface of the soil'' (for an incorrect answer). Corrective answer feedback told learners whether they were correct or incorrect but did not offer any explanation. Better learning resulted from explanatory feedback with a high effect size of 1.16. In addition, students rated the version with explanatory feedback as more helpful than the version with corrective feedback. This research is the basis for our *Explanatory Feedback Principle: Include tailored feedback for all correct and incorrect options that tell the learners they are right or wrong AND also give an explanation why the answer is right or wrong.*

Learning Goal Versus Normative Feedback. Goal orientation is an important motivational factor that predicts learning strategies. Individuals with a learning goal orientation focus on increasing their individual competence over time. In contrast, a performance goal orientation focuses on performance relative to others—in other words, on ''looking good.'' Research shows that individuals with a performance orientation will diminish effort in the face of failure and seek less challenging goals for which success is likely (for reviews, see Clark, 2008a; Shute, 2008). In contrast, those with a learning goal orientation tend to persist in the face of failure and pursue more challenging goals. Feedback that helps learners focus on their own progress over time rather than their own progress compared to that of other learners will promote a more productive motivational orientation. For example, a progress bar that shows skill proficiency of an individual over a series of lessons is a more productive form of feedback than a progress bar that illustrates how an individual's score compares to others. This research is the basis for our *Learning-Goal Feedback Principle: Design feedback that informs the learners of their progress toward attaining a learning goal; avoid feedback that leads to a comparison with other learners*.

Task-Focused Versus Self-Focused Feedback. Receiving feedback is a potentially ego-threatening event. In a large meta-analysis of feedback, Kluger and DeNisi (1996) show that much feedback actually reduces performance! For positive results, they recommend feedback that minimizes attention to the self and instead directs attention to the task. For example, feedback that is normative (as discussed in a previous paragraph), feedback that includes praise or discouraging comments, or feedback from a person (rather than a computer) may direct attention to the self rather than the task. Attention to the self increases the probability of unproductive ego involvement. Instead, provide feedback that shows progress from previous attainments, emphasizes correct solutions, avoids both praise (*Great Job!*) and negative comments (*Sorry. You are wrong*), and comes from a less personal source such as a computer. This research is the basis for our *Task-Focus Feedback Principle: Design feedback that directs attention to the task rather than the learner*.

Feedback is a complex instructional method and could easily be the topic of an entire chapter. For more details, please consult the recent review by Shute (2008).

5. Maximize Performance Potential with Deliberate Practice

From musicians to athletes to chess players to medical practitioners, over twenty years of research on high-level performers demonstrates the positive relationship between practice and expertise. Across many different domains of expertise, a ''Ten-Year Rule'' has emerged. According to the Ten-Year Rule, elite levels of expertise require about ten years of sustained training and effort. However, we all know individuals who pursue an avocation such as golf or music for over ten years but reach only acceptable levels of performance. So there is more to expertise than just years of practice. Ericsson (2006) distinguishes between *routine practice* and *deliberate practice*. He found that all expert violinists spent over fifty hours a week on music activities. But the best violinists spent more time per week on activities that had been specifically tailored to improve their performance. ''The core assumption of deliberate practice is that expert performance is acquired gradually and that effective improvement of performance requires the opportunity to find suitable training tasks that the performer can master sequentially . . . typically monitored by a teacher or coach'' (Ericsson, 2006, p. 692). In other words, deliberate practice requires good performers to concentrate on specific skills that present a challenge that is neither too difficult nor too easy for that performer.

A regimen of deliberate practice will leverage generative cognitive load by helping performers invest effort on specific aspects of critical tasks that need improvement. Schnotz and Kurschner (2007) recommend that practice tasks be tailored to the individual so that they are not so difficult as to impose too much cognitive load, but at the same time offer sufficient challenge to engage learning processes. An optimal level of challenge can be achieved by adjusting the difficulty of the practice assignment or by adjusting the amount of guidance offered in conjunction with the exercise. Practice difficulty can be increased by assigning an exercise that incorporates multiple elements (rather than just a few factors), varies the context of the practice scenarios, or reduces guidance such as cueing or hints. Research on practice leading to high levels of expertise is the basis for our *Practice Challenge Principle: Tailor your practice exercises so that the difficulty level and skill focus offer the appropriate challenge to extend expertise of each learner*.

Because directive learning environments are characterized by relatively simple tasks that escalate in complexity over the sequence of lessons as well as by extensive guidance, we recommend these environments for beginning stages of learning. As learners gain expertise, the level of task difficulty should increase, and the level of guidance should decrease to maintain an optimal level of cognitive load for learning.

Tailored regimens of practice are difficult to administer in large-class instructor-led settings. However, in asynchronous e-learning, frequent diagnostic assessments can monitor individual levels of learning and adjust the lesson difficulty (task assignments or amount of support) based on outcomes. Creating diagnostic assessments and branching logic is a labor-intensive task and makes sense when you can save considerable learning time in an audience with a heterogeneous background or rate of learning. Deliberate practice can also be provided by a coach who observes performance and assigns specific practice exercises. The low ratio of instructor to performer makes coaching a costly intervention and is justified when very high levels of expertise will benefit the organization. Senior management development is one example for which a deliberate practice regimen may be cost-effective.

PRACTICE FORMATS VERSUS PRACTICE FUNCTIONALITY

It is not the format of a practice exercise that ensures its efficacy. Rather, it's the design of the exercise to promote deeper processing. For example, multiple-choice practice items can be constructed to promote maintenance or elaborative processing of content. In an Excel lesson, a maintenance exercise may ask the learner to select the correct meaning of a formula operator from four descriptions. In contrast, an elaborative multiple-choice item may describe a calculation goal and ask the learner to select the formula that will produce the correct result. Likewise, drag-and-drop, matching, and true-false type of formats can promote shallow processing or can promote deeper processing. Too often, practice exercises are written at a shallow (maintenance) level because these types of items are faster to construct and because some practitioners are not aware of the benefits of elaborative exercises. In this section we summarize guidelines from Clark (2008) and from Foshay, Silber, and Stelnicki (2003) for designing practice activities that promote deep processing of concepts and of near and far transfer tasks. Table 10.3 summarizes these guidelines.

Table 10.3 Examples of Practice for Content Types

Content	*Maintenance Practice*	*Elaborative Practice*
Concepts	Describe an adverb; select the features of a valid formula	Select all adverbs in ten sentences; circle the valid formula
Procedural Tasks	List the steps to enter a formula into a spreadsheet; sequence the steps into the correct order	Enter a formula into the spreadsheet; calibrate the micrometer
Strategic Tasks	Describe the four principles of effective report writing; select the important guidelines for negotiating a dispute	Write a report; negotiate a dispute

Design Guidelines for Concept Practice

A concept is a mental representation of a class of objects, events, or symbols that applies to diverse instances. Some common concepts include formula, chair, training, and practice. Concepts are defined by specific features that are common to most members of the group, which in turn may differ on irrelevant features. For example, all chairs have a seat, back, and some support to the floor. However, chairs vary in regard to size, presence of arms, the type of floor support such as wheels or rockers, and so forth. A major goal of nearly all training and educational programs is to build conceptual understanding through concept schemas in long-term memory. A concept schema allows you to identify an instance of a given concept, even though you have never seen it before. For example, baggage screeners must identify luggage with prohibited items. Spreadsheet users must identify correct formats for formulas. Nearly all job tasks include important concepts that must be applied to complete the task. Practice to build concept mental models should ask learners to identify or construct valid instances of the concept class. New spreadsheet users, for example, should be provided with several formulas from which they identify those with valid formats. This type of exercise could be formatted as a multiple-choice, a drag-and-drop, or a true-false item. Following a recognition exercise, learners may be given a calculation goal and be asked to construct the correct formula.

Design Guidelines for Task Practice

For procedural task outcomes, the practice exercise should require the learner to perform the steps to complete the task. In directive learning environments, job procedures are typically broken into bite-size tasks so that the learner practices a limited number of steps at a time. Often a part-task approach is used whereby a complex task is broken into a group of smaller tasks, each of which is practiced in isolation. As learning progresses, smaller tasks are combined until eventually the whole procedure is mastered. Initial procedural practice may be heavily guided. For example, if the learner makes an error, she is immediately corrected and asked to repeat the step. In the same way, a working aid with a list of the steps involved in the procedure is provided for guidance.

The format for procedural task practice is some type of hands-on activity in which the learner engages with the real performance environment or a close simulation of it. For example, in an Excel spreadsheet class, the learner may practice with real spreadsheets. To practice laparoscopic surgical procedures, the learner may work with a simulation. Such operational simulations have proven useful for practice of procedures that have adverse safety or cost consequences of errors in job settings.

In contrast to procedural skills, practice can also be designed to build strategic skills. A strategic skill is a capability to perform tasks for which there are no single invariant approaches or responses. Strategic tasks involve product design or problem solving that require judgment. Unlike procedures, strategic tasks require the learner to adjust her approach each time to accommodate changing circumstances. To build a more robust mental model needed for strategic tasks, the instruction may teach process content as well as include varied context examples and practice. We discuss these techniques in the section on learning transfer to follow. The format of these practice exercises is also hands-on, initiated by a case-study or role-play activity in which the learners are asked to solve a problem or develop a product congruent with the learning objective. Conceptual simulations are often used to build deeper understanding associated with strategic task performance. For example, the simulation shown in Figure 10.3 allows learners to conduct research and make a loan decision in a virtual environment.

Some learning psychologists recommend using a guided discovery learning environment for strategic tasks. Guided discovery environments such as the example shown in Figure 10.3 typically use a whole-task approach in which the learner is asked to perform authentic tasks or resolve realistic problems. An inductive learning environment characteristic of guided discovery encourages learners to attempt tasks and learn from the experience. For example, in the loan course shown in Figure 10.3, learners are assigned a loan to research and approve and have great freedom to access various resources about the loan applicant. In contrast, a directive environment uses a more instructive approach by imposing greater structure in the content sequence and the inclusion of focused practice with immediate explanatory feedback in each lesson.

While guided discovery environments are quite popular among some practitioners, we do not yet have sufficient research to indicate when a strategic task is best trained in a directive versus a guided discovery environment.

MAXIMIZING LEARNING FROM EXAMPLES IN DIRECTIVE LEARNING ENVIRONMENTS

Although frequent practice with feedback is one of the core instructional methods that characterize directive learning environments, research shows that too much practice in early stages of learning imposes extraneous cognitive load (Clark, Nguyen, & Sweller, 2006). We recommend that you substitute some worked examples for practice exercises during initial learning.

What Are Worked Examples?

A worked example is a step-by-step demonstration of how to solve a problem or complete a procedure. For example, in Figure 10.6 you can see part of a worked

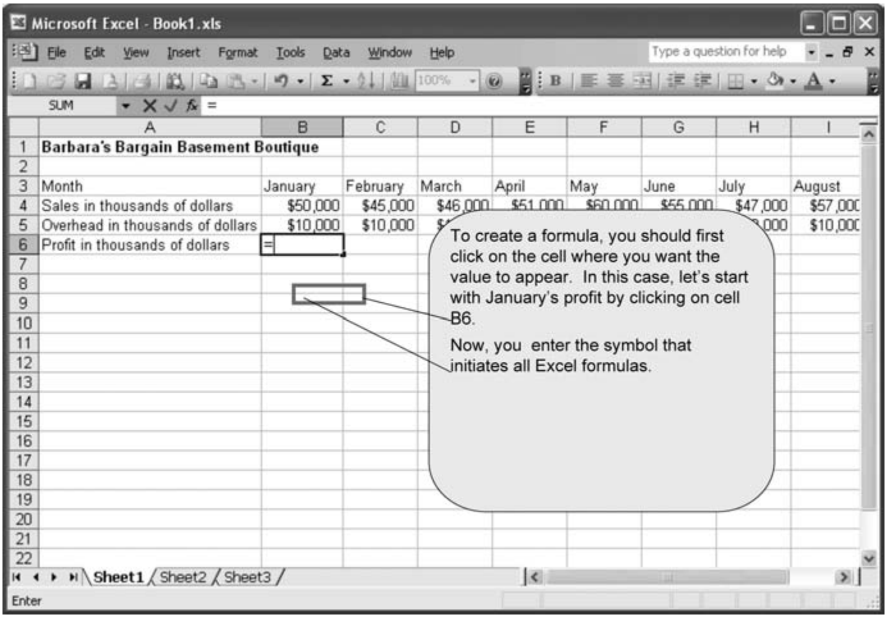


Figure 10.6 A Worked Example from an Excel Course.
With permission from Clark, Nguyen, and Sweller, 2006

example of how to construct a formula in Excel. The early research on worked examples compared learning from mathematics lessons that included twelve practice exercises with lessons that included six worked examples and six practice exercises (Sweller & Cooper, 1985). Therefore, students in both lessons were exposed to the same twelve problems. One group worked all twelve problems as a practice assignment. The comparison group viewed six problems as worked-out examples and solved the other six problems as a follow-up to each example. It's not surprising that the all-practice lesson versions took almost six times longer to complete. After all, it is more time-consuming to work out twelve problems than to study six and work out six. The good news is that the more efficient lessons that included pairs of examples and problems led to fewer errors both during the training and on the test.

Worked examples proved more efficient because they reduced extraneous cognitive load imposed by working practice exercises. When studying a worked-out example, the learner has more mental resources to perform the mental activity, which builds more new mental models than when having to work a practice. Ironically, an instructional method that seems more passive, for example, viewing an example rather than working a practice, actually can produce greater mental learner activity by freeing mental capacity for processing. The power of worked examples has proven very robust and can be applied to a variety of skills, including soft skills such as negotiations, problem-solving, and troubleshooting as well as the more procedural tasks, as illustrated in Figure 10.6.

Who Benefits from Worked Examples?

Worked examples benefit novice learners the most. That is because a worked-out example serves as a knowledge substitute for novices. Once a learner has built her own mental model of the task, she actually profits more from exercising her model in practice activities than from studying worked examples. Kalyuga, Chandler, Tuovinen, and Sweller (2001) compared lessons with worked example-problem pairs with all-problem assignments in lessons on writing programs for relay circuits of various levels of complexity. Learners were randomly assigned to lessons with either all-practice or example-practice pairs. Learning of both groups was measured at several points during the training period. In the beginning, the worked-example-problem pairs led to better learning than the all-problem lessons, as described in the previous paragraph. However, as the learners gained expertise, gradually the lessons with worked examples offered no advantage and eventually became disadvantageous. Instructional methods that work well for novices but have no effect or even a negative effect for experienced workers reflect a phenomenon called the *Expertise Reversal Effect* (Kayluga, 2007). How can you deploy worked examples most effectively in directive learning environments?

Transition Gradually from Worked Examples to Practice Through Fading

When applying a fading technique, you provide worked examples during the beginning stages of learning and transition gradually to practice as the learner gains skills. To create a faded worked example, start with a full worked-out example such as the one illustrated in Figure 10.6. Follow with a second example in which one or more of the steps are worked out and the learner finishes the remaining steps. For example, in a three-step problem, the first two steps are demonstrated and the learner finishes the problem by completing the last step. The next example would illustrate only the first step for the learner, who then finishes the last two steps. Finally, the learner faces a full practice exercise in which she works all three steps.

A number of experiments have demonstrated better learning from faded worked examples compared to learning from worked example/problem pairs. For example, Atkinson, Renkl, and Merrill (2003) designed lessons on probability that included three-step faded worked examples. Half the participants were assigned to lessons with faded worked examples, while the other half were assigned to lessons with worked examples paired with problems. The amount

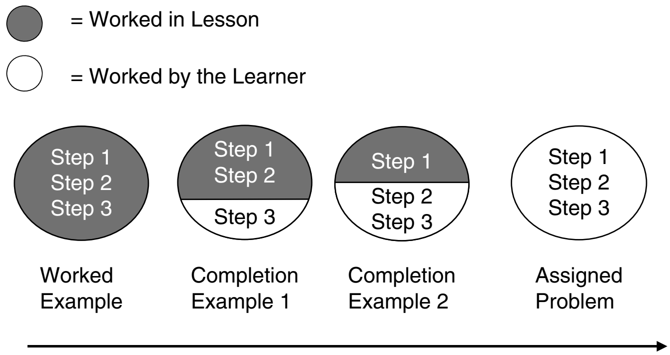


Figure 10.7 Fading from a Full Worked Example to a Practice Problem.

of study time was the same in both groups. Learning was better among those studying the lessons with faded examples than among those with example-problem pairs. The fading procedure offers the highest cognitive load support to beginning learners who need it the most. As mentioned previously, studying worked examples saves time and improves learning of novices. As learning progresses and the learners build their own mental models, this support is gradually removed to allow more experienced learners to engage in practice. As illustrated in Figure 10.7, the fading strategy offers a gradual transition from a complete worked example to a practice exercise as the learner gains expertise.

Encourage Self-Explanations of Examples

One drawback to examples is that many learners don't really make effective use of them. Some may skip them altogether. Others review them but only at a shallow level. Chi and her colleagues (1994) found that students who studied physics examples and generated self-explanations learned more than students who did not self-explain. A self-explanation is student ''self-talk'' during which he or she processes the example deeply. For example, the learner might relate the example to the concepts discussed in the lesson. Alternatively, she might study a few steps and then predict the final steps of the example. The research team found that learners who scored higher on a physics test generated, on average, 15.5 self-explanations per example, compared to 2.75 from poorer learners.

By studying the type of self-explanations associated with better learning, the research team found that effective self-explanations either help learners identify and correct misconceptions or expand on the ideas presented in the example by making inferences. How can we extend the benefits of self-explanations to all learners?

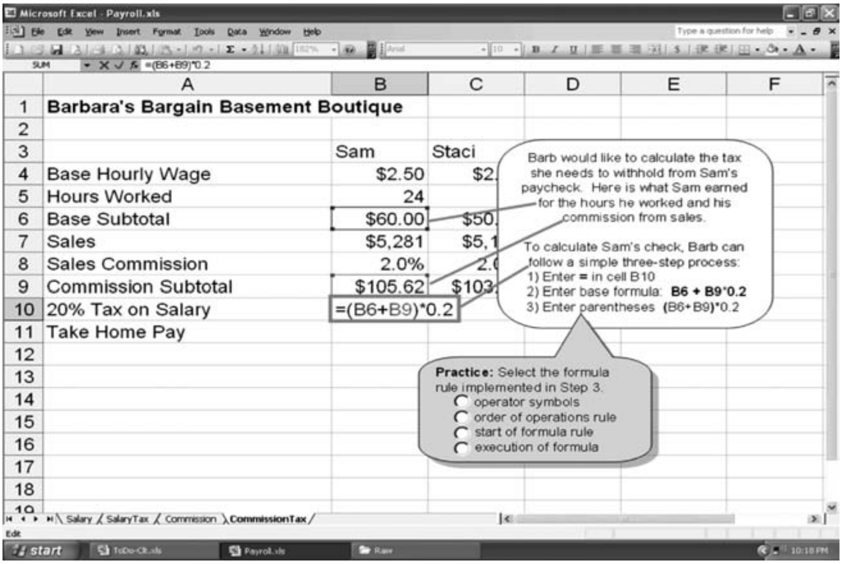


Figure 10.8 A Self-Explanation Question in an Excel Worked Example.

With permission from Clark, Nguyen, and Sweller, 2006

One approach is to embed self-explanation opportunities within the worked-out example. You can do so by adding a question next to a worked-out step that requires the learner to identify important concepts or principles illustrated in that step. For example, in our Excel lesson, Figure 10.8 shows a practice question linked to step three of the demonstration. The question requires the learner to identify the formula rule being implemented. Answering this question ensures that the learner (1) reviews the example carefully and (2) associates the example with important related content. Atkinson and his colleagues (2003) found that lessons that included self-explanation questions associated with probability worked steps led to better learning than worked examples without a self-explanation question. One way to implement self-explanation questions is to think about what type of concepts or principles underlie each of your demonstrated steps. Then construct a multiple-choice question for one of your worked steps that includes the correct concept or explanation along with some distractors. A multiple-choice menu of options constrains the number of possible self-explanations and avoids the potential for incorrect self-explanations that might arise with a more open-ended response format. There may be other productive formats for self-explanation questions, such as *predict what step follows this step* or *predict the outcome of this step* with options presented in a multiple-choice format. However, the research on self-explanation questions is new, and we will need additional guidance on the types of questions that maximize learning from worked examples.

Exploiting Examples: A Summary

Examples and practice are two of the defining elements of directive learning environments nicknamed *rule-example-practice* or ‘‘tell-show-do’’ instructional formats. The research evidence to date recommends that, rather than assigning new learners numerous practice exercises, use a fading technique in which you start with a fully worked-out example and progress through several examples during which the learner gradually assumes greater responsibility for completing the example until reaching a full practice exercise. To help all learners maximize the learning value of examples, we recommend placing self-explanation questions with at least some of the worked-out steps. When responding to these questions, learners will need to carefully process the worked example and in so doing actively engage the selecting, organizing, and integrating processes associated with learning.

Based on research to date, we offer a *Worked Example Principle: Replace some practice with a series of worked-out examples that gradually assign more steps to the learner. To optimize processing of worked examples, include self-explanation questions that require learners to deeply process worked-out steps*.

MAXIMIZING LEARNING TRANSFER IN DIRECTIVE LEARNING ENVIRONMENTS

It is not sufficient to build a mental model in long-term memory if that mental model cannot be applied to performance at a future time. For example, if a learner masters the main tasks associated with Excel spreadsheets in a training class, she will need to perform similar tasks back on the job. Transfer of learning refers to the ability to apply new knowledge and skills acquired in a training setting to the application environment, which is usually the work setting. Unfortunately, a long history of research on transfer has shown that transfer failure is more common than transfer success. In this section we focus on how to incorporate features into directive environments that maximize transfer.

Near Versus Far Transfer Tasks

Near transfer tasks are also called procedures and involve tasks that are similar to those presented during training. For example, logging into a specific e-mail account and completing an online customer order are typical near transfer tasks. In contrast, far transfer tasks, also called strategic tasks or ill-structured problem-solving tasks, will require learners to adapt the knowledge and skills acquired during training to unique situations. Conducting a hiring interview or designing a web page are two examples of far transfer tasks. Far transfer tasks are performed by adapting domain-specific knowledge to the unique context of the application environment. For example, when designing a web page, the goal of the web page, the technological issues, customer branding guidelines, and usability factors must all be integrated into a unique solution.

Design Guidelines for Near Transfer Learning

Instructional psychologists have learned the critical role of the context of learning to promote transfer. A basic law of transfer called the Identical Elements Principle states that the *cues of retrieval* must be embedded in memory *at the time of learning*. In other words, the context of the performance environment must be incorporated into the training environment. For example, when learning how to use a spreadsheet, the examples and practice should use spreadsheets and spreadsheet problems that are similar to those the learner will encounter on the job. Based on the *Identical Elements Principle, we recommend that you present examples and practice during training that mirror the application setting*.

Design Guidelines for Far Transfer Learning

Unfortunately, the Identical Elements Principle will constrain the degree of transfer you can achieve. Since far transfer of strategic tasks will have varied contexts each time they are performed, there will be no one context that will support learning transfer. Instead you will need to use instructional methods that build a more flexible mental model that can be adapted to diverse application situations. In short, you will need to promote deeper levels of learning. To achieve far transfer, we recommend two main strategies: process-based lessons and diverse context examples and practice.

Strategy A: Build Understanding of ''How It Works''

Imagine that your instructional goal is to teach learners to make a specific indicator light up on a control panel that contained a number of dials and switches. You could teach learners a step-by-step procedure that involved six steps to make the indicator light up. Alternatively, you could extend the procedural training by adding an explanation of how the system worked. Specifically, you could explain the meaning of each element on the control panel and its influence on the entire system. Teaching how things work is the focus of process-based lessons. Adding process content would require additional training time. Is that time justified?

Kieras and Bovair (1984) experimented with two different training approaches, summarized in the previous paragraph. They measured how long it took learners to master the procedure and to resolve system problems. The results are summarized in Table 10.4. As you can see, the group that received system training along with procedural training required overall more training time. However, as a result of greater understanding of how the system worked, learners who studied the process lesson were able to make procedures more efficient as well as to adjust procedures to resolve problems. The research

Table 10.4 Learning a System Model Improved Retention and Led to More Shortcuts

Outcome	*Rote Training*	*Model Training*	*Improvement*
Training Time–System Model	NA	1141	NA
Training Time–Procedure	270	194	28 percent
Retention	67 percent	80 percent	19 percent
Shortcuts	8 percent	40 percent	400 percent
Execution Time	20.1	16.8	17 percent

From Kieras and Bovair, 1984, p. 263.

team concluded that performance transfer can be enhanced by helping learners build an understanding of a system. We recommend system training in situations such as troubleshooting for which an understanding of the system supports enhanced judgment of the performer.

Strategy B: Use Varied Context Examples and Practice

We have seen the power of a combination of examples and practice to build new knowledge and skills. Since far transfer tasks will require an adaptation by the learner each time they are performed, you can build a more flexible mental model by varying the surface features of your examples and practice while helping learners to focus on the common principles or guidelines that underpin them. For example, if your goal is to teach how to work with formulas in an Excel spreadsheet, you can promote learning by showing diverse examples and assigning diverse practice exercises that require learners to construct different formulas to achieve varied calculation goals. Diverse examples and practice exercises will take more time to construct and will impose more mental load during learning. Do we have evidence that this diversity supports far transfer?

Evidence for Varied Context Examples

Quilici and Mayer (1996) created examples to illustrate three statistical tests, including t-test, correlation, and chi-square. Each of these tests requires a different solution formula and applies to different types of data. For each test type, the research team created three examples. In one version, all example sets used the same surface story. For example, the three t-test problems used data regarding experience and typing speed; the three correlation examples used data regarding temperature and precipitation; and the three chi-square examples included data related to fatigue and performance. These example sets, which used the same cover story, were called surface emphasizing examples. A second lesson version used examples that varied the cover story. For example, the t-test was illustrated by one example that used experience and typing speed, a second example about temperature and precipitation, and a third example about fatigue and performance. The examples in these lessons were called structure emphasizing examples. After studying either the surface-emphasizing or the structure-emphasizing lesson versions, learners were asked to (1) sort a new set of problems according to the statistical test needed or (2) select which set of calculations they should use to implement the correct statistical tests. The structure-emphasizing examples led to significantly greater discrimination among test types. By building a more robust mental model that can be applied to diverse unpredictable performance situations, training efficiency is maximized.

Consistent with many experiments that have shown a transfer benefit of using a range of examples that vary the surface features, we recommend a *Varied Context Example Principle: Construct a series of examples that use diverse surface features to help learners build more flexible mental models to support far transfer learning*.

WHEN TO USE DIRECTIVE LEARNING ENVIRONMENTS

We began this chapter by describing an instructional goal that required training for loan underwriters using asynchronous e-learning delivery. We considered using a receptive approach in which the key guidelines are explained and illustrated; a directive approach in which underwriting tasks are broken into short segments with practice assigned to each segment; or a guided discovery approach in which learning is mediated by solving of cases studies using an interface, such as that shown in Figure 10.3. Which design approach is best? How can you determine when to use a directive environment?

First, we know that, whatever lesson design is used, the goal is to help learners build mental models, which can be transferred to the job. To decide which design approach is most appropriate, first consider whether the outcome goal is primarily to help learners acquire conceptual knowledge or also to build new skills. For underwriters, the operational objective requires them to analyze data and to make funding decisions that minimize risk to the lending institution. This goal involves both knowledge and skills. Because of the importance of the skill-building outcomes, we recommend against a receptive design in lieu of a design that provides overt opportunities to try out new skills and get feedback.

Therefore we need to adopt either a directive or a guided discovery design. Since the learners will be primarily novice, we recommend starting with a directive course design. A number of brief lessons would describe and illustrate the various tasks performed as part of the loan analysis process, each providing examples, practice, and feedback. Toward the end of the course, you might switch to a guided discovery design using a whole-task approach in which loan application case studies are presented and learners have freedom to access relevant applicant data and make their recommendation. This guided discovery segment could be mediated by online simulations such as the one illustrated in Figure 10.3 or by way of a coach guiding work through structured case scenarios at the local office level.

There is some heated debate on the learning value of receptive, directive, and guided discovery instructional environments (Tobias & Duffy, 2009). Some practitioners eschew receptive environments as being too passive and thus ineffective. Others condemn guided discovery environments because they impose too much cognitive load for productive learning (Kirschner, Sweller, & Clark, 2006). However, evidence points to the value of receptive, directive, and guided discovery environments depending on the instructional goals and the learners' background knowledge and skills. The key to success is to support the critical underlying learning events of selection, organization, and integration of new knowledge in a manner that is appropriate to your outcome goals and your learners' level of prior knowledge. Rather than a ''one or the other'' perspective, we recommend that instructional professionals adapt all of these designs in ways that complement their goals and learners.

In the case of directive lessons, use instructional modes and methods that manage intrinsic cognitive processing, minimize extraneous cognitive processing, and foster generative cognitive processing. For example, break loan analysis tasks into a series of small tasks to manage intrinsic load. Use audio to explain relevant visuals and eliminate any irrelevant words or visuals to maximize limited working memory capacity. Finally, distribute practice exercises with explanatory feedback throughout all lessons to promote generative cognitive load.

Although more traditional learning environments such as receptive or directive course designs are sometimes regarded as outmoded, less effective, or less engaging than the newer guided discovery approaches, research shows that each of the three designs can be appropriate for certain kinds of learners and learning outcomes. Further, whichever design you use for a given situation, its instructional power will benefit from evidence-based instructional methods such as those we have described in Chapters Nine and Ten.

References

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked out steps. *Journal of Educational Psychology*, *95*(4), 774–783.

Campbell, J. P., & Oblinger, D. G. (2007). Top-ten teaching and learning issues, 2007. *Educause Quarterly*, *30*(3), 15–22.

Chi, M.T.H., DeLeeuw, N., Chiu, M., LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*, 439–477.

Clark, R. C. (2008a). *Building expertise: Proven methods for training and performance improvement* (3rd ed.). San Francisco: Pfeiffer.

Clark, R. C. (2008b). *Developing technical training* (3rd ed.). San Francisco: Pfeiffer.

Clark, R. C., & Mayer, R. E. (2008). *e-Learning and the science of instruction* (2nd ed.). San Francisco: Pfeiffer.

Clark, R. C., & Mayer, R. E. (2009). Learning by viewing vs. learning by doing. *Performance Improvement*, *47*(9), 5–13.

Clark, R. C., Nguyen, F., & Sweller, J. (2006). *Efficiency in learning*. San Francisco: Pfeiffer.

Craig, F.I.M, & Watkins, M. J. (1973). The role of rehearsal in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 12*, 599–607.

Ericcson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance*. New York: Cambridge University Press.

Foshay, W. R., Silber, K. H., & Stelnicki, M. B. (2003). *Writing training materials that work*. San Francisco: Pfeiffer.

Glenberg, A. M., Wilkinson, A. C., & Epstein, W. (1992). The illusion of knowing: Failure in the self-assessment of comprehension. In T. O. Nelson (Ed.), *Metacognition: Core readings*. Boston: Allyn & Bacon.

Hasler, B. S., Kersten, B., & Sweller, J. (2007). Learner control, cognitive load and instructional animation. *Applied Cognitive Psychology*, *21*, 713–729.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, *19*, 509–539.

Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of Educational Psychology*, *93*, 579–588.

Keehner, M., Lippa, Y., Montello, D. R., Tendick, F., & Hagarty, M. (2006). Learning a spatial skill for surgery: How the contribution of abilities changes with practice. *Applied Cognitive Psychology*, *20*, 487–503.

Kester, L, Kirschner, P. A., & van Merriënböer, J.J.U.G (2006). Just-in-time information presentation: Improving learning a troubleshooting skill. *Contemporary Educational Psychology*, *31*, 167–185.

Kieras, D. E., & Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science*, *8*, 255–273.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback intervention on performance: A historical review, a meta analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254–284.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology*, *93*, 390–397.

Mayer, R. E., Mathias, A., & Wetzell, K. (2002). Fostering understanding of multimedia messages through pretraining: Evidence for a two-stage theory of mental model construction. *Journal of Experimental Psychology: Applied*, *8*, 147–154.

Moreno, R. (2004). Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Instructional Science*: *32*, 99–113.

Moreno, R., & Mayer, R. E. (2005). Role of guidance, reflection, and interactivity in an agent-based multimedia game. *Journal of Educational Psychology*, *97*, 117–128.

Moreno, R., & Mayer, R. E. (2007). Interactive multimedia learning environments. *Educational Psychology Review*, *19*, 309–326.

Quilici, J. L., & Mayer, R. E. (1996). Role of examples in how students learn to categorize statistics word problems. *Journal of Educational Psychology*, *88*(1), 144–161.

Rohrer, D., & Taylor, K. (2006). The effects of over learning and distributed practice on the retention of mathematics knowledge. *Applied Cognitive Psychology*, *20*, 1209–1224.

Scheiter, K., & Gerjets, P. (2007). Learner control in hypermedia environments. *Educational Psychology Review*, *19*, 285–307.

Schnackenberg, H. L., Sullivan, H. J., Leader, L. R., & Jones, E.E.K. (1998). Learner preferences and achievement under differing amounts of learner practice. *Educational Technology Research and Development*, *46*, 5–15.

Schnotz, W., & Kurschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychology Review*, *19*, 469–508.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153–189.

Sweller, J. (1999). *Instructional design in technical areas*. Camberwell, Australia: ACER Press.

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, *2*, 59–89.

Sweller, J., van Merriënböer, J.J.G., & Paas, F.G.W.C (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251–296.

Tobias, S., & Duffy, T. (Eds.). (2009). *Constructivist theory applied to instruction: Success or failure?* New York: Taylor & Francis.

CHAPTER ELEVEN

Assembling and Analyzing the Building Blocks of Problem-Based Learning Environments

David H. Jonassen

Over the past decade, I have worked to articulate a meta-theory of learning to solve problems (Jonassen, 1997, 2000, 2001, 2002, 2003a, 2003b, 2004, 2006, 2007a, 2007b, 2008; Jonassen & Hung, 2006, 2008). The primary rationale for that effort is overcoming the dearth of instructional design principles and methods for helping learners to develop problem-solving skills, especially given the ubiquity of problem solving in everyday and professional lives. Problems are everywhere. Employees in corporations, the military, and other agencies are hired, retained, and rewarded for solving problems. While most corporate trainers prefer to interpret problems as opportunities, the reality is that our lives are suffused with opportunities for which there are no readily apparent answers, ergo problems. How do I get my DSL to work? Which is the best route to avoid this traffic jam? How do I prevent my boss from criticizing me? How can I get that new contract? What shall we make for supper this evening? How shall we market this new product to maximize cash flow? What do I have to do in order to attract recognition in this agency? We are deluged with problems every day. In his book of essays, Karl Popper (1999) averred, ''All life is problem solving.'' Unfortunately, we have rarely been taught how to solve problems, especially the full range of problems that are encountered in daily and professional lives.

Another rationale for constructing a meta-theory of problem solving is that traditional pedagogies that emphasize direct instruction are inadequate for helping people to learn how to solve the range of problems. Why? Because problems vary, and the most foundational principle of instructional design is that different learning outcomes require different instructional conditions (Gagné, 1965).

PROBLEM-SPECIFIC MODELS OF PROBLEM SOLVING

My theory of problem solving diverges from traditional approaches, known as phase models, to problem solving that articulates single approaches to solving all kinds of problems (e.g., Bransford & Stein, 1994). These models generally suggest that all problems can be solved if we (1) identify the problem, (2) generate alternative solutions, (3) evaluate those solutions, (4) implement the chosen solution, and (5) evaluate the effectiveness of the solution. However, problems and problem solving vary in several ways. Foremost among these differences is the continuum between well-structured and ill-structured problems (Jonassen, 1997; Simon, 1969; Voss & Post, 1988). Most problems encountered in formal education are well-structured problems. Well-structured problems typically present all elements of the problem; engage a limited number of rules and principles that are organized in a predictive and prescriptive arrangement; possess correct, convergent answers; and have preferred, prescribed solution processes. For example, training employees on how to apply rules or procedures in their company represents well-structured problems, and because well-structured problems are easier to teach, they often become the default objectives.

Ill-structured problems, on the other hand, are the kinds of problems that are encountered in everyday practice. Ill-structured problems have many alternative solutions to problems; vaguely defined or unclear goals and constraints; multiple solution paths; and multiple criteria for evaluating solutions; so they are more difficult to solve. Learning to troubleshoot complex systems or learning how to make policy decisions or to adapt accounting techniques are ill-structured problems. There are no accepted solution paths, and optimal solutions are rarely, if ever, known.

Problems also vary in complexity. The complexity of a problem is a function of the breadth of knowledge required to solve the problem, the level of prior knowledge, the intricacy for the problem-solution procedures, and the relational complexity of the problem (number of relations that need to be processed in parallel during a problem-solving process) (Jonassen & Hung, 2008). Ill-structured problems tend to be more complex; however, there are a number of highly complex well-structured problems, such as playing chess or writing computer programs.

Dynamicity, another characteristic of problems, may be thought of as another dimension of problem complexity. In dynamic problems, the relationships among variables or factors change over time. Changes in one factor may cause variable changes in other factors that often substantively change the nature of the problem. The more intricate these interactions, the more difficult it is to ascertain a solution. Ill-structured problems tend to be more dynamic than well-structured problems that tend to be static. That is, the problem elements are stable.

Another dimension of problems and problem solving is domain specificity. In contemporary psychology, there is a common belief that problems within a domain rely on cognitive strategies that are specific to that domain (Smith, 1991; Sternberg & Frensch, 1991). These are often referred to as strong methods, as opposed to domain-general strategies (weak methods). For example, Lehman, Lempert, and Nisbett (1988) concluded that different forms of reasoning are learned in different graduate disciplines. Graduate students in the probabilistic sciences of psychology and medicine perform better on statistical, methodological, and conditional reasoning problems than students in law and chemistry, who do not learn such forms of reasoning. The cognitive operations are learned through the development of pragmatic reasoning schemas rather than exercises in formal logic. Graduates in different domains develop reasoning skills through solving situated, ill-structured problems that require forms of logic that are domain-specific.

How do problems vary within these dimensions? Finally, problems vary by context. That is, the way problems are solved depends on the context in which they are solved. Context may describe the purpose of the organization in which the problem is solved (education, corporations, military, and so forth). Problems also vary between different organizations. Accounting problems, for example, are solved differently in various corporations, just as problems are solved differently within different divisions of the same corporation. All of these differences make problem solving ill-structured.

If problems vary, then there must be different kinds of problems. Jonassen (2000) described a typology of problems that vary primarily along a continuum from well-structured to ill-structured, including puzzles, algorithms, story problems, rule-using problems, decision making, troubleshooting, diagnosis-solution problems, strategic performance, systems analysis, design problems, and dilemmas. This typology assumes that there are similarities in the cognitive processes for solving problems within classes. Within each category of problems that is described, problems may vary with regard to abstractness, complexity, and dynamicity. For example, troubleshooting a lawn mower engine is less complex than troubleshooting a Formula One race engine. A goal of mine is to develop and test an instructional design model for each kind of problem. To date, models exist for story problems (Jonassen, 2003a), troubleshooting problems (Jonassen & Hung, 2006), and policy analysis problems (Jonassen, 2004). Recently, I began the deconstruction of design problems and have generated an initial interpretation of design problem solving (Jonassen, 2008) that may lead to an instructional design model. Also, I am currently working on a model for decision making, which is performed quite differently by experienced decision makers (Klein, 1998) than the ways that we are taught to make decisions. To provide an example of these models, I shall briefly present the design model for troubleshooting problems.

Troubleshooting

Troubleshooting is among the most commonly experienced kinds of problem solving in the professional world. From troubleshooting a faulty modem or multiplexed refrigeration systems in modern supermarkets, troubleshooting attempts to isolate fault states in a dysfunctional system. Troubleshooting is also an essential cognitive component of medical diagnoses and psychotherapy. Having found the fault, the part or subsystem is replaced or repaired. Troubleshooting is most often taught as a checklist of decisions that directs the fault isolation. This approach may work for helping novices solve simple troubleshooting problems; however, it is inadequate for training competent, professional troubleshooters. Effective troubleshooting requires system knowledge (conceptual knowledge of how the system works), procedural knowledge (how to perform problem-solving procedures), and strategic knowledge (strategies such as search-and-replace, serial elimination, and space splitting) (Pokorny, Hall, Gallaway, & Dibble, 1996). These different kinds of knowledge comprise the troubleshooter's mental model of the process. In order to solve troubleshooting problems, problem solvers must identify the fault state(s) and symptoms, describe goal state, identify the subsystem that the fault occurs in, and diagnose the problem. Diagnosis is most commonly facilitated by recalling prior problems and how they were solved. If the troubleshooter is unable to recall a similar case, then he or she needs to begin iterative hypothesis testing based on the troubleshooter's conceptual model of the system until the problem is solved. Those conceptual models include failure modes and probabilities of system components, and the hypothesis generation involves manipulation of the conceptual model to predict the behavior of the system with fault probabilities (Hall, Gott, & Pokorny, 1995). Additionally, expert troubleshooters often consider the cost and information value of test/replacement operations.

In order to learn to troubleshoot, novices must engage in problem-solving behavior that is supported by a rich conceptual model of the problem space and by pseudo-experiences. Figure 11.1 illustrates a design architecture for building troubleshooting learning environments (Jonassen & Hung, 2006). The model assumes that the most effective way to learn to troubleshoot is by

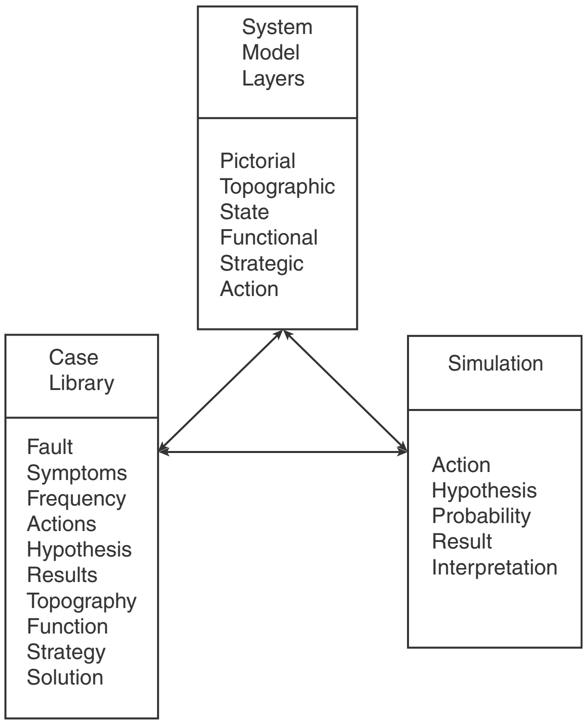


Figure 11.1 Model for Designing a Troubleshooting Learning Environment.

troubleshooting problems. Learning to troubleshoot problems requires presenting learners with the symptoms of novel problems and requiring them to solve them. The major components of the troubleshooting learning environment are a case library of previously solved problems, a diagnostic simulation that enables the learner to practice troubleshooting, and a rich conceptual model of the system being troubleshot. The conceptual model supports the construction of system knowledge; the troubleshooter supports the construction of procedural and strategic knowledge; and the case library supports the construction of the experiential knowledge that integrates all of the other kinds of knowledge.

The conceptual model of the system includes multiple representations of the system being troubleshot, including the:

- *Pictorial layer*—pictures of the device or system as it actually exists;
- *Topographic layer*—components of the system and their interconnections;
- *State layer*—normal states, fault states, and probability or frequency of faults;

- *Functional layer*—information, energy, or product flows through the system;
- *Strategic layer*—alternative decisions regarding the states; and
- *Action layer*—descriptions of procedures for conducting various tests along with their cost and time.

The heart of the learning environment is the diagnostic simulation. After seeing a description of the problem that describes the symptoms just before the car ceased to work, the learner (like an experienced troubleshooter) first selects an action, such as ordering a test, checking a connection, or trying a repair strategy. The student must select a fault hypothesis that he or she is testing. Requiring the learner to justify the action taken is an implicit form of argumentation. The learner then receives the results of action (test results, system status, and so on) and must interpret those results. If the interpretation is inconsistent with the action, hypothesis, or subsystem, then an error message is triggered.

The case library or fault database contains stories of as many troubleshooting experiences as possible. Each case represents a story of a domain-specific troubleshooting instance. Case libraries, based on principles of case-based reasoning, represent a powerful form of instructional support for ill-structured problems such as troubleshooting (Jonassen & Hernandez-Serrano, 2002). The case library indexes each case or story according to its system fault, the system or subsystem in which the fault occurred, the symptoms of the fault, similar to the troubleshooter. The failure mode, hypotheses or strategies that were tested, the results of those tests, and what lessons were learned from the experience are also contained in the case library. The case library represents the experiential knowledge of potentially hundreds of experienced troubleshooters. Why? Because troubleshooters almost invariably store their knowledge of problems and solutions in terms of their experiences. It is important to note that troubleshooting requires different skills than other kinds of problem solving, such as design. Thus, the engineer who designed a system is not likely to be as good at troubleshooting it and so should not be used as the only SME when designing a troubleshooting learning environment.

The most common exemplars of troubleshooting are automobile mechanics (Why won't my car start?) or electronics troubleshooting (Why can't I connect to the server?). However, troubleshooting pervades medicine, psychotherapy, and a host of softer domains. For example, offices of all sizes in all contexts are plagued with social dynamics problems that result in an inability to communicate (Why can't these folks work together?). A troubleshooting environment to support learning how to diagnose communications problems would include a problem scenario that the learner is supposed to diagnose.

In the simulation, the learners are presented with a list of actions they could take (questions they may ask to certain people, documents they want to examine, or histories they want to collect). Before selecting the action, they may want to consult the case library of stories about others communications problems that had been diagnosed to see what others had done, or they may want to study a model of the communication system they are diagnosing, including the organization, politics, social dynamics, psychological make-up, or history of the organization in order to better understand the specifics of the problem. The system model shows not only who the stakeholders are but also how they affect each other, what their normal states are, strategies for interactions, and so on. After selecting an action, the student must explain his or her hypothesis (Why are they doing this?) and predict how probable it is that this is a problem. If the hypothesis agrees with the action, then he or she would receive an answer, which the student must interpret. The students would continue to take actions that are informed by stories from the case library or information in the systems model until they understood the problem well enough to suggest a solution. Yes, the learning curve and the cognitive load in such a learning environment is steep. But mental models are not constructed from reproductive interactions.

Summary: Problem-Specific Models of Problem Solving

My research over the past decade has focused on articulating how problems are different and therefore how problem-based learning environments must also differ. More recently, my research has led me to test an alternate hypothesis: that although problems vary, there are similarities in the cognitive requirements for solving different kinds of problems. But first, I describe the building blocks of all problem-based learning environments—cases. Like problems, the functions that cases play in learning also differ. In the final section, I will describe how those cases can be related to each other.

THE ROLE OF CASES IN LEARNING TO SOLVE PROBLEMS

Problem-based learning has many different meanings that are located along a continuum from the methods suggested by Barrows and Tamblyn (1980), where all learning throughout the curriculum emerges from problems to problem-based modules found in science classrooms. In between, many courses are now being redesigned as problem-based. For example, we are now designing and testing a problem-based version of a materials science course, which is the basis for the mechanical engineering curriculum. Regardless of where the problem-based activity falls along the continuum, all problem-based learning is case-based. That is, problems are represented as cases, and cases are used in various ways to support learning how to solve problems. Problems that students solve are represented as cases, and ancillary support problems are also represented as cases. The primary medium of problem-based learning is the case. Like different kinds of problems, though, I argue that there are different kinds of cases, and the ways that cases are represented and interrelated by the learners depend on the functions of the cases, which I describe next.

Functions of Cases

Rather than teaching students theoretical abstractions to describe a field of study, cases typically describe situations or scenarios wherein students attempt to apply theories to the cases. That is, the cases are used by students to learn how to solve problems. Merseth (1994) identified three categories of cases: cases as exemplars, cases for analysis, and cases for personal reflection. In this chapter, I suggest a typology of case applications to instruction that expands on Merseth's categories. This typology specifically addresses the function of cases, that is, the ways that students use cases in order to learn to solve problems. Cases may be used as exemplars or models, as analogues, as alternative perspectives, as remindings, as experiences to analyze (case study method), as problems to solve, and as student constructions (see Table 11.1). It should be

Table 11.1 Functions of Cases

Cases as . . .	*Function as . . .*	*Manifested in . . .*
Exemplars (models)	Expository schemas or processes	Textbook examples; worked examples
Analogues	Analogical comparison of problem structures for schema induction	Analogical encoding
Remindings	Case-based reasoning	Case libraries of indexed stories
Alternative perspectives	Interconnectedness of multiple perspectives and themes	Cognitive flexibility hypertexts
Experiences to analyze (case study)	In-depth, longitudinal analysis of a single event analyzed by learners	Teacher education; Harvard business cases
Problems to solve	Scenario or context for which solution is needed	Problem-based learning; goal-based scenarios; anchored instruction
Student-constructed	Student interpretations of phenomena	Constructionist activities

noted that these classes describe the ways that cases are used, not the structure, form, or representation of the cases. It is probable that certain functions demand certain formats or structures, a hypothesis that demands investigation.

Cases as Examplars (Models). Typically, problem-based learning is practice-based. That is, students must practice whatever kind of problem solving that is required. That means that they need examples of problems that need solutions. Most scholars agree that you cannot effectively didactically teach students how to solve problems (although many try using procedural phase models of problem solving). That is, students cannot learn to solve problems when they are taught about problem solving. Rather, instruction needs to focus on ''how'' to solve problems. In order to show students how to solve a problem, then you must use cases as examples of problem solving. All models of instructional design insist on the inclusion of examples in instruction. After defining an entity, examples should be presented to learners. The role of examples is to serve as models or applications of ideas being represented abstractly. Their purpose is to help learners to construct schemas for the ideas being presented. A schema for a problem consists of the kind of problem it is, the structural elements of the problem (for example, acceleration, distance, and velocity in a physics problem), situations in which such problems occur (for example, inclined planes, automobiles), and the processing operations required to solve that problem (Jonassen, 2003).

The most common form of cases as exemplars in problem solving is the worked example. When learning to solve problems, cases in the form of worked examples may be provided as a primary form of instruction. Worked examples are instructional devices that typically include the problem statement and a procedure for solving the problem for showing how other problems may be solved (Atkinson, Derry, Renkl, & Wortham, 2000). Worked examples that focus on problem type and sub-procedures involved in the process are cognitively very demanding. Therefore, Gerjets, Scheiter, and Catrambone (2004) break down complex solutions into smaller meaningful solution elements that can be conveyed separately. Worked examples should present multiple examples in multiple modalities for each kind of problem, emphasize the conceptual structure of the problem, vary formats within problem types, and signal the deep structure of the problem (Atkinson, Derry, Renkl, & Wortham, 2000). When learners self-explain the examples, they learn to better solve problems. Because research with worked examples has always focused on well-structured problems with convergent solutions and solution methods, no research has examined the effectiveness of worked examples with ill-structured problems. Because ill-structured problems do not possess correct solutions, solution criteria, or solution paths, they may not be amenable to worked examples. In order to use worked examples to support ill-structured problem solving, multiple examples would have to be demonstrated and the students required to evaluate and select the most appropriate, a task that would involve heavy cognitive load. Heavy cognitive load contradicts the espoused purpose of worked examples.

Cases as Analogues. One of the most effective ways to learn to solve problems is to examine similar problems for their structure and their solutions. There are two theoretical approaches to using cases as analogues: analogical encoding and case-based reasoning (see next section).

The simplest, yet the most common method for teaching problem solving is the worked example (described before). With worked examples, the instructor shows students how to solve a problem (usually by plugging variables into an equation) and then requires the students to solve a practice problem. The students are expected to construct a schema for the problem by examining the worked example, store the schemas in memory, and later analogically apply that schema in order to solve a new problem (Gick & Holyoak, 1983). Students usually construct only a process schema for the set of operations required to solve the problem but miss the underlying structure of the problem. When learners examine worked examples that illustrate certain rules and then apply lessons learned from that example to solving a new problem, they tend to apply only those examples that are most similar to the target problem. Mapping examples to problems is affected by the similarity of objects between the examples and problems being solved, especially story lines and object correspondences (that is, whether similar objects filled similar roles) (Ross, 1984, 1987, 1989a). That is, learners often fail to recall or reuse examples appropriately because their retrieval is based on a comparison of the surface features of the examples with the target problem, not their structural features. When the target problems emphasize structural features that are shared with the example, generalization improves (Catrambone & Holyoak, 1989; Reed, Ackinclose, & Voss, 1990).

The theory that best describes the required analogical reasoning is structure mapping theory (Gentner, 1983), where mapping the analogue to the problem requires relating the structure of the analogue to the structure of the problem independent of the surface objects in either. In order to do so, those surface features (which attract the attention of poor problem solvers) must be discarded. Learners must focus on the higher-order, structural relations among elements within the problem, comparing the analogue with the problem being solved, a process known as analogical encoding.

Analogical encoding is the process of mapping structural properties between multiple analogues. Rather than attempting to induce and transfer a schema based on a single example, comprehension, schema inducement, and far transfer across contexts can be greatly facilitated by analogical encoding, the comparison of two analogues for structural alignment (Catrambone & Holyoak, 1989; Gentner & Markman, 1997; Gentner & Markman, 2005; Loewenstein, Thompson, & Gentner, 2003). If presented with just one example, students are far more likely to recall and apply problems that have similar surface features. Analogical encoding fosters learning because analogies promote attention to commonalities, including common principles and schemas (Gick & Holyoak, 1983). During analogical encoding, students must compare analogous problems for their structural alignment. Problems are structurally aligned when the relationships (arguments) among problem elements match (Gentner & Markman, 1997). It is this abstraction of higher-order structural relationships that facilitates far transfer in problem solving. Analogical encoding also contrasts with the conventional behaviorist practice of having the learner practice problems as close to the target problem as possible, without much hope of far transfer.

Cases as Remindings. Another way of using cases to support problem solving is by analogy directly with the source problem without attempting to construct a schema. A common approach to problem solving consists of finding the nearest case in an organized library of annotated problem cases and reusing or adapting it. When a new problem is encountered, most humans attempt to retrieve cases of previously solved problems from memory in order to reuse the old case. If the solution suggested from the previous case does not work, then the old case must be revised (Jonassen & Hernandez-Serrano, 2002). When either solution is confirmed, the learned case is retained for later use. Case-based reasoning is based on a theory of memory in which episodic or experiential memories in the form of scripts (Schank & Abelson, 1977) are encoded in memory and retrieved and reused when needed (Schank, 1990; Kolodner, 1993).

Case-based reasoning is applied to instruction in the form of case libraries of stories that are made available to learners. The stories in the library are indexed in order to make them accessible to learners when they encounter a problem. Those indexes may identify common contextual elements, solutions tried, expectations violated, or lessons learned. Case libraries are included in goal-based scenarios (described later) as a primary form of instructional advice. Rather than teaching students theory, those environments retrieve relevant cases for the students to learn from. When allowed to retrieve previous relevant cases while solving food product development processes, Hernandez-Serrano and Jonassen (2003) found that on tests assessing problem-solving skills, such as reminding, identifying, and recognizing the problem, identifying and explaining failure, selecting solutions, adapting solutions, explaining success or alternate strategies, and identifying needed information, students who accessed stories outperformed students who reviewed expository help in lieu of the stories.

Cases as Alternative Perspectives. Ill-structured problems tend to be more complex than well-structured problems. In complex knowledge domains or problems, the underlying complexity should be signaled to the learner, who considers alternative perspectives on the problem in order to construct personal meaning for the problem (Spiro, Coulson, Feltovitch, & Anderson, 1988). Cognitive flexibility theory prescribes the use of hypertexts to provide random access to cases representing multiple perspectives and thematic representations of content, enabling students to crisscross the cases that they are studying through the use of multiple conceptual representations, linking abstract concepts to different cases, highlighting the interrelated nature of knowledge via thematic relations among the cases, and encouraging learners to integrate all the cases as well as their related information into a coherent knowledge base (Jacobson & Spiro, 1995). The interlinkage of concrete cases and perspectives with abstract themes allows students to develop a much more complex and coherent knowledge base (Jacobson, Maouri, Mishra, & Kolar, 1995). Most ill-structured problems demand the use of cases as alternative perspectives. For example, in a problem-based learning environment on engineering ethics that we are conducting research on, students solve three everyday ethics cases (cases as problems to solve). The first case is a story about a manufacturing engineer who becomes socially involved with a vendor through playing and betting on golf matches. Each case problem to solve is supported by multiple interpretations or perspectives on the case, including the engineer's perspective, the company's perspective, and the vendor's perspective. Additionally, each case is interpreted in terms of different theoretical approaches, including utilitarian, rights and duty, and virtue, as well as ethical canons from the National Society of Professional Engineers. Students are required to examine each interpretation of the case and develop an argument in support of their solution. It is important to emphasize that all perspectives were theoretical and canonical interpretations of the case events, not descriptions of perspectives, theories, or canons. A theory has no meaning unless the learners can apply it.

Cases as Experiences to Analyze: Case Study Method. Perhaps the most common conception of case-based learning is the case study. In case studies, students study an account (usually narratives from one to thirty pages) of a problem that was previously experienced. Frequently guided by questions, students analyze the situation and processes and evaluate the methods and solutions. This analysis is usually ex post facto. In most case studies, students are not responsible for solving the problems, only analyzing how others solved the problems and engaging in what-if thinking. Case studies are stimuli for discussions. The goals of the case study method are to embed learning in authentic contexts that require students to apply knowledge rather than acquire it. Mayo (2002) found that students in an introductory psychology class studying case narratives outperformed students enrolled in a lecture-based class in terms of theoretical comprehension and application. Examples of the case study method include the Harvard business cases (Barnes, Christiansen, & Moore, 1994) and case studies in teacher education (Schulman, 1992; Sudzina, 1999). Because this method was developed before an adequate theory of case-based learning, the cases and surrounding activities often lack the desired features and learner activities necessary for learning to transfer problem-solving skills.

Cases as Problems to Solve. Although analysis of cases may be the most common application of cases, cases are also used quite commonly to represent problems that students solve, rather than analyze. In that research, cases provide background information, contextual information, and instructional supports to help students generate and test different solutions to the problem presented. Cases as problems to solve have many examples in practice, including anchored instruction and goal-based scenarios.

One of the best known and most effective innovations in instructional design is anchored instruction. Based on situated learning theory and cognitive apprenticeships, anchored instruction embeds problems into complex and realistic scenarios, called macrocontexts. Developed by the Cognition and Technology Group at Vanderbilt (1991, 1993), anchored instruction uses high-quality video scenarios to introduce the problem and engage learners in identifying and solving the problem. The video is used to present a story narrative that requires the learners to generate the problem to be solved, rather than having the entire problem circumscribed by the instruction. All of the data needed to solve the math and science problems are embedded in the story, requiring students to make decisions about what data is important. The problems that students generate and solve are complex, often requiring more than twenty steps to solve, rather than simple story problems.

In goal-based scenarios (GBSs), students become active participants in a scenario (as compared with anchored instruction, in which learners only observe the scenario). GBSs teach complex systems by identifying a goal to be achieved and a set of skills the student can learn and apply in the context of the system in question. They employ a ''learning by doing'' architecture (Schank & Cleary, 1995) in which learners are immersed in a focused, goal-oriented situation (for example, selling Yellow Pages advertising, accommodating new business practices); required to perform authentic, real-world activities; and supported with advice in the form of stores that are indexed and accessed using case-based reasoning formulae. The situatedness of the instruction facilitates comprehension, retention, and recognition of the conditions in which learning may be applied and therefore transferred. Skills are developed through practice in an authentic environment, so the scenario must be fairly realistic, feedback continuously provided, and the action-outcomes plausible. Learning is driven by acceptance of a meaningful goal.

Student-Constructed Cases. Strobel, Jonassen, and Ionas (2008) engaged in a three-year, design-based research study of case-based learning. Beginning with a cognitive flexibility hypertext (described earlier), they found that students slowly adapted to the non-linear interconnections of the different content-types; however, students experienced difficulty in making comparisons, because the environments did not provide space for annotating one's thoughts. Flexibility hypertexts are static, providing a definitive body of material that is difficult for users to elaborate. Users of the system were unable to contribute their own perspectives, links, or connections, so they were passive consumers of information stored in the environment.

In the second and third iterations, the system shifted from navigation to a student authoring environment, because authoring hypertext requires deeper understanding of the domain; identification of core concepts, cases, themes; and careful selection of new cases to represent the content. We incorporated authoring functions that gave students more control of the environment, so that the focus of designing the hypertext system shifted from content and relationship development to providing support structures and guidance to the end-users as the instructional designer of their own learning experience. When students construct and elaborate their own cases, they are more deeply engaged in learning than when interpreting someone else's cases. That is the principle of constructionism.

Summary: Role of Cases

As indicated before, Jonassen (2000) articulated a typology of problems that varies from well-structured to ill-structured problems. The original list including puzzles, algorithms, story problems, rule-using problems, decision making, troubleshooting, diagnosis-solution problems, strategic performance, systems analysis, design problems, and dilemmas. In this evolving theory of problem solving, I am now associating the kinds of cases that would be essential in teaching students how to solve these problems. Table 11.2 proposes the kinds of cases that are essential to problem-based learning environments for each kind of problem; however, this is currently only a prediction. As with any evolving theory, the associations in Table 11.2 require empirical validation through learning research as well as development research. Table 11.2 is proffered as a heuristic for selecting cases to be included in problem-based learning environments.

Table 11.2 Essential Uses of Cases to Support Learning Different Kinds of Problems

Problem Type	*Uses of Cases*
Algorithms	Cases as Problems; Cases as Examples; Cases as Analogues
Story (word) problems	Cases as Examples; Cases as Problems; Cases as Analogues; Student-Produced Cases (Problem Posing)
Rule (using/rule induction)	Cases as Examples; Cases as Problems; Cases as Analogues
Decision making	Case Studies; Cases as Analogues; Cases as Alternative Perspectives
Troubleshooting	Cases as Problems; Cases as Analogues; Cases as Reminding; Strategic Performance
Policy analysis	Case Studies; Case as Alternative Perspectives; Student-Constructed Cases
Design problems	Cases as Problems; Cases as Analogues; Student-Produced Cases
Dilemmas	Case Studies; Cases as Alternative Perspectives

HOW CASES CAN BE INTERRELATED: COGNITIVE SKILLS OF PROBLEM SOLVING

Earlier, for an institution responsible for designing problem-solving instruction, that may be problematic unless the nature of the problems in their institution are constrained (which is often the case). For training automobile mechanics, physicians, and a host of trades involved in troubleshooting systems, the troubleshooting model should be sufficient. However, for solving school-oriented science and math problems, that model is likely to be ineffective, just the same as it would be for teaching design skills in an engineering class. So, for reasons of generalizability and scalability, I have begun to examine the degree to which problems and problem solving call on similar cognitive skills.

I believe that there are three cognitive skills that are essential to solving problems: analogical reasoning, casual reasoning, and argumentation. I describe these skills next and later speculate on how they can interrelate cases in problem-based learning environments.

Analogical Comparison of Problems

The goal of problem solving includes not only a correct answer (should one exist) but also the construction of a problem schema, knowing what kind of problem it is. A problem schema consists of the kind of problem it is (such as a work-energy problem), the structural elements of the problem (see Figure 11.2), situations in which such problems occur (inclined planes, automobiles, and so forth), and the processing operations (for example, equations) required to solve that problem (Jonassen, 2003). In physics problems, students learn the processing operations and focus on the surface-level, situational characteristics of the problem, but they do not understand what kind of problem it is based on its structural characteristics, impeding far transfer. Learners are expected to construct (induce) schemas from the examples, analogues, alternative perspectives, or student-constructed cases; store the schemas in memory; and later analogically transfer to them when solving new problems (Gick & Holyoak, 1983). Analogical transfer is a form of generalization, that is, transferring one problem schema onto another. Generalization from problem-solving examples may automatically occur in limited ways (Catrambone & Holyoak, 1983), however, knowledge generalization is not a natural consequence of reasoning by analogy (Didierjean, 2003). Learners sometimes adapt highly specific, contextualized knowledge from analysis of examples without engaging in the reasoning by analogy process that leads to generalization. Attracting attention to the structural similarity between problems improves generalization during problem solving (Didierjean, 2003). Failing to do so can inhibit far transfer/generalization.

In order to ensure that learners are transferring appropriate problem schemas, it is necessary to analyze the underlying structure of each example, experience, analogue, or perspective to ensure that they are structurally congruent. The theory that best describes the required analogical reasoning is structure-mapping theory (Gentner, 1983), where mapping the analogue to the problem requires relating the structure of the example, experience, analogue, or alternative perspective to the structure of the problem independent of the surface objects in either. In order to do so, those surface features (which attract the attention of poor problem solvers) must be set aside. Then the higher-order, systemic relations must be compared on a one-to-one basis in the example and the problem.

Consistent with semantic network theory (Quillian, 1968), knowledge in human memory can be represented as propositional networks of nodes and predicates. According to structure-mapping theory (Gentner, 1983), attributes are predicates taking one argument (a concept), and relationships are predicates taking two or more arguments. The rules that define those relationships depend only on syntactic properties of knowledge representation and not on specific content of domains, allowing analogies to be distinguished clearly from literal similarity (Gentner, 1983). So it is important for the designer to certify that the structures of problems consist of propositions that are congruent across analogues. To do so, the designer must compare the relational predicates, but few or

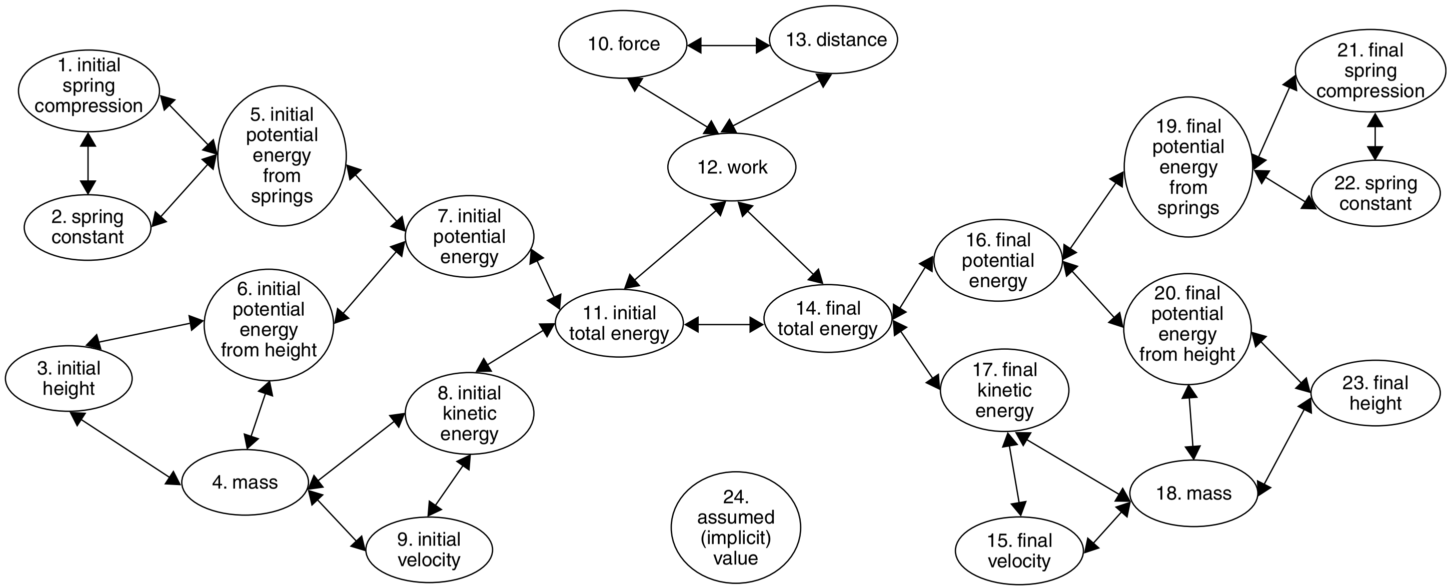


Figure 11.2 Structure Map of Work-Energy Problems.

no object attributes, between base to target. That is, analogous problems are structurally equivalent but may look quite different. Figure 11.2 illustrates all of the propositions commonly found in work-energy problems in physics. We have used this structure to structurally compare problem sets that we have used in physics classes. We, the designers, compared analogues for their relational (structural) similarity. All work-energy problems contain some of these predicates, so the job of the student and the designer is to compare problems to see which elements are included. The problem in Figure 11.3 calls on the learners to calculate how far the block will return up the plane. They need to highlight the components of Figure 11.2 that are present in the problem and show the connections that are implied by the problem (the one on Figure 11.3). In order for analogues to be structurally aligned, relations must have matching arguments, the relational focus must involve common relations but not necessarily common objects, and analogies must match connected systems of relations (Gentner & Markman, 1997). Being able to analogically compare problems is essential to learning how to solve problems qualitatively, that is, to understand the problems conceptually.

Causal Reasoning

As described before, propositions describe relationships among the elements stated in any problem. The relationships (second order predicates) that define problems are universally causal (Jonassen & Ionas, 2008). So, when comparing the structures of cases, the learner must examine the underlying causal relationships in order to ensure that the same kinds of causal reasoning are called by different cases. For example, in the work-energy problems, the learners must identify the circumstances under which potential energy is converted to kinetic and how potential energy is conserved. If the same problem elements are causally related to each other in the same ways, the problems are structurally isomorphic. Causal relationships enable learners to perform cognitive activities that are required for solving all problems, including predictions, inferences, implications, and explanations.

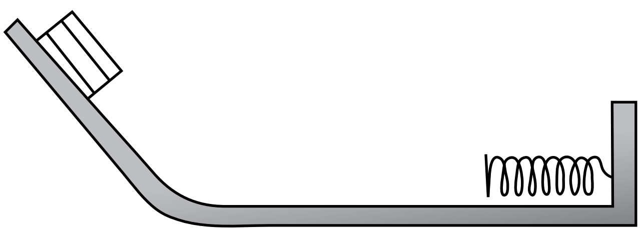


Figure 11.3 Work-Energy Problem.

Reasoning from a description of a condition or set of conditions or states of an event to inferring the possible effect(s) that may result from those states is called prediction. Prediction assumes a deterministic relationship between cause and effect. The two primary functions of prediction are forecasting an event (for example, economic or meteorological forecasts) and testing of hypotheses to confirm or disconfirm scientific assumptions (such as predicting the effects of a hormone on an animal's growth rate). Predictions are the basis of experimentation; they are the hypotheses of experiments. A physicist, for example, predicts (hypothesizes) that the application of a force to an object will result in a change of state in that object. Scientific predictions are tested for their empirical validity. A psychologist may predict that changes in environmental conditions, stress, for example, will affect or change a person's behavior. When causal agents are known and used to predict the effects on another agent or entity, casual reasoning is required.

A non-deterministic form of prediction is to draw implications from a set of conditions or states or from a cause-effect relationship. To imply is to entail or entangle events or to involve an effect as a necessary consequence of some cause without necessarily knowing what the effect will be. Drawing implications involves identifying potential effects from a causal antecedent. For example, the implications of any new law are potentially (directly or indirectly) complex. Therefore, implications of any event are often not known or articulated or could not have been hypothesized. As such, implications represent a conditional form of prediction that is less deterministic than a prediction. Implications have received very little research or analysis in psychology or philosophy, so little is known about implicational reasoning.

When an outcome or state exists for which the causal agent is unknown, then an inference is required. That is, reasoning backward from effect to cause requires the process of inference. A primary function of inferences is diagnosis. Etiology is the identification of a cause, an origin, or a reason for something that has occurred. In medicine, etiology seeks to identify the cause or origin of a disease or disorder as determined by medical diagnosis. For example, based on symptoms, historical factors, and test results of patients that are thought to be abnormal, a physician attempts to infer the cause(s) of that illness state. Medical specialties are based in part on an understanding of different subsystems in the human body and the cause-effect relationships that determine various states. In the same way, automobile mechanics specialize in certain makes of cars because of an increased awareness of causal (functional) connections in different makes of automobiles. Inferences are required to associate effects with possible causes.

Causal reasoning also supports explanations. Explaining any entity requires more than an awareness of the parts of that entity. Explanations require functional knowledge of the entity or system being explained. Functional knowledge includes the comprehension of the function and structure of the interrelationships among the components in any system and the causal relationships between them (Sembugamoorthy & Chandrasekaran, 1986). You cannot fully explain any entity or event without understanding it causally. For example, the process of diagnosing disease states requires that physicians' explanations of diseases include causal networks that depict the combination of inferences needed to reach a diagnosis (Thagard, 2000). The abilities to predict, implicate, infer, and explain all require causal reasoning.

Although causality is most commonly induced empirically, empirical descriptions are insufficient for understanding causality. Contemporary accounts of causality emphasize three main principles that validate a causal relationship, including covariation (co-occurrence) principle, priority principle, and mechanism principle (Bullock, Gelman, & Baillargeon, 1982). Covariation is the degree or extent to which one element consistently affects another, which is expressed quantitatively in terms of probabilities and covariance. The mechanism principle implies a qualitative conceptual understanding of the mechanisms that mediate causal relationships. The covariational and the mechanism principles are the two most common conceptual frameworks for studying causal reasoning (Ahn, Kalish, Medin, & Gelman, 1995). However, all three principles are necessary for completely describing causal relationships. Understanding causal relationships necessitates analysis of all of these characteristics. Both covariation and mechanisms are manifest in different ways and are described by several different attributes. In this chapter, I will explicate these attributes of covariation and mechanisms and show how understanding causality in terms of all of these attributes is necessary for scientific reasoning and explanation.

In order to transfer problem-solving abilities, learners must comprehend both covariational and mechanistic attributes of the causal relationships. Covariational attributes of causal relationships are quantitative descriptions of causal relationships. Most science problems are solved by associating problem elements with variables in an equation and then solving the equation for the correct answer. Every equation is a quantitative (covariational) description of the causal relationships in the problem. Equations describe the direction (positive, negative), the strength or valency (slight, moderate, significant, strong), the probability, the immediacy (immediate, delayed), and the duration (immediate, short term, long term) of the causal relationships in any problem (Jonassen & Ionas, 2008). In different problems, these characteristics have differential importance. In order to be able to make predictions and inferences using causal relationships, students must be able to comprehend these characteristics.

Qualitative understanding of causality is necessary for transferring problem-solving skills. Mechanisms are theoretical entities, theories, or processes that underlie the relationships. They specify the way that something works, answering ''why'' questions in order to specify ''how'' the event occurred. For example, when the Federal Reserve cuts interest rates, why or how does that increase money supply, and what effect does increased money supply have on borrowing, lending, gross domestic product, and so on? Causal mechanisms explain the empirical relationships in any event (Salmon, 1984).

So understanding of causal relationships also requires qualitative, mechanistic explanations. ''To describe the cause or causes of an event is to explain why it occurred'' (Kuhn, 1977, p. 23). The mechanism is the force or motive that changes the effect, according to Aristotle. The mechanism principle describes the beliefs that people construct for explaining the relationship between cause and effect. The causal mechanism is the causal chain of intermediary events that connects a cause and an effect. The mechanism-based causal reasoning approach understands causal reasoning qualitatively because it explains ''how'' and ''why'' the cause(s) produces the effect.

Causal mechanisms include conjunction/disjunction, causal process, and necessity/sufficiency (Jonassen & Ionas, 2008). Although we too often associate a specific cause with a specific effect, most causal relationships result from a conjunction of different types of causes. Conjunctive plurality occurs when two or more causes, C_1 and C_x, must be jointly present in order to produce the effect E, and no subset of causes will produce the same effect E. For example, many people believe that terrorism results from overzealous adherence to religious dogma. However, effects are almost invariably produced by multiple factors, including poverty, lack of economic opportunity, and cultural mores, that are individually necessary and jointly sufficient to produce the effect (Cheng, 1997).

In order to understand the role of different conjunctive causes, it is necessary to examine the roles that each plays. Cheng and Nisbett (1993) proposed that causal relationships be represented in terms of whether the causal factor is a necessary or a sufficient condition for an effect to occur. Necessity/sufficiency is a difficult but essential attribute of causality. Necessity is more a complex concept than sufficiency. For sufficiency, people only verify whether the cause is always followed by the effect, whereas for necessity, there are two possibilities that can be verified: ''Does the cause always precede the effect?'' and ''Can the effect occur without the cause?''

More importantly, causation is commonly conceived on a molar level where we attribute a relationship without understanding the underlying process. For example, most of us attribute the contraction of a common cold to someone sneezing near us. While the sneeze may be a key causal agent, the process of viral transmission is much more complex than that. So students of medicine, microbiology, or other related fields must be able to explicate that causal process in a more detailed way. Their models must explicate the numerous casual factors that mediate that relationship. Germs are dispersed through the air by the sneeze, and some attach to host cells. The virus injects its genetic material into the host cell. That genetic code is copied into the host cell, breaking out of it and invading other cells, all of which sets off complex immunological reactions, including the distribution of mast cells to the site of the infection, the release of histamines, causing inflammation of the tissue, causing more immune cells to be delivered to fight off the infection. If learners cannot adequately articulate these complex causal processes, their mechanistic understanding is overly simplified. Likewise, they are unable to articulate the complexity of individual causal relationships; they do not possess sufficient prior knowledge to solve problems.

Because most learners have grossly inadequate understanding of the causal relationships that are required for problem solving, it is important that the designer identify all of the causal relationships implied by any problem. For example, understanding of economics includes causal relationships such as those among the supply, demand, and price of any product. Unless these causal relationships are understood more completely, inappropriate economic responses may be recommended. Our claim is that all students learning important causal relationships in whatever discipline is being studied must be able to provide both covariational and mechanistic kinds of explanations of those relationships. These explanations call on distinctly different kinds of reasoning. For example, in showing students how to solve science problems, in addition to plugging problem values into a formula, use a mechanistic map, such as Figure 11.2, to explain how the various problem elements act on each other.

Test items requiring learners to apply those relationships may be used as a pretest of prior knowledge. Those causal relationship may also provide the objectives for prerequisite or co-requisite instruction. For example, economics problems may be represented in terms of their causal relationships. Figure 11.4 represents the underlying mechanisms of economics problems related to supply and demand.

Argumentation

The third and final cognitive skill required to learn how to solve problems is argumentation. Argumentation is an essential critical thinking skill that is consistently applied in formal and informal (everyday) reasoning situations. Argumentation is the means by which we rationally resolve questions, make decisions, and solve problems (Nussbaum & Sinatra, 2003). Although the implications for argumentation are more obvious for ill-structured problems that have no correct answer and multiple potential solution criteria, argumentation has also been shown to be effective with well-structured problems. When students answered physics story problems incorrectly, they were required to construct an argument for the scientifically correct answer. Nussbaum and Sinatra (2003) prompted the students with ‘‘Some people actually pick [choice X]. Can you think of a reason why it may be [choice x]? They found that students

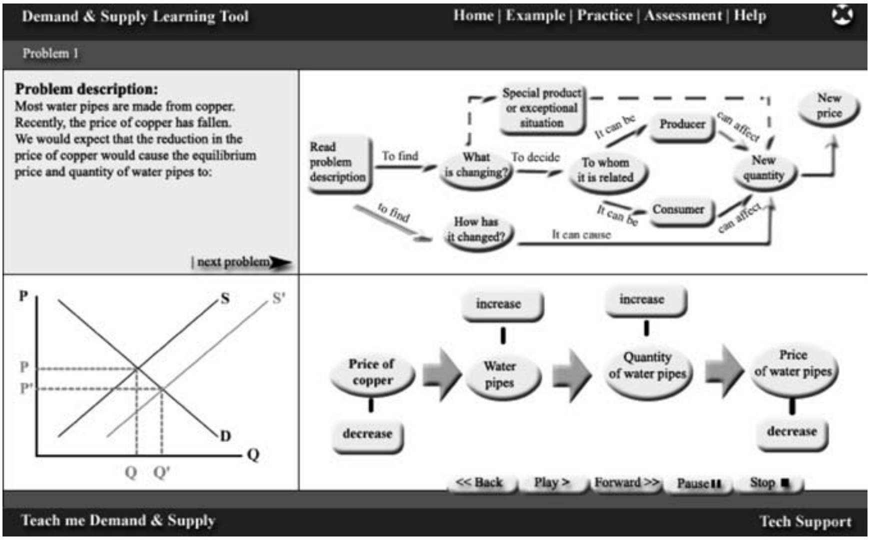


Figure 11.4 Mechanisms in Supply-and-Demand Problem.

who argued in favor of an alternative (correct) explanation of a physics problem showed improved reasoning on problems, and when they retested participants from the experimental treatment a year later, the quality of their reasoning remained strong.

Unfortunately, most American students do not understand argumentative discourse (Reznitskya, Anderson, McNurlin, Nguyen-Jahiel, Archodidou, & Kim, 2001). They experience difficulty writing persuasive essays; comprehending written arguments; differentiating between theory and evidence; generating genuine evidence, alternative theories, counterarguments, or rebuttals (Kuhn, 1991; Means & Voss, 1996). There are serious weaknesses in the arguments of adolescents and young adults. They are unlikely to construct two-sided arguments or distinguish evidence from explanation in support of a claim (Kuhn, 1991; Kuhn, Shaw, & Felton, 1997; Voss & Means, 1991). So regardless of how important argumentation is to problem solving, problem-based learning environments must scaffold argumentation skills.

There are two primary forms of argument that may be supported in problem-based learning environments: rhetorical and dialectical. The goal of rhetorical arguments is to persuade or convince an audience. Thus, rhetorical argumentation models, for example, focus on how to persuade or convince a target audience of a claim or proposition. So an argument is considered successful if it gains the approval of the target audience (van Eemeren & Grootendorst, 1992). So learning to argue is a matter of developing effective persuasive argumentation techniques.

Rather than a dialogue between arguer and an audience, dialectical argumentation represents a dialogue between a proponent and an opponent of a controversial issue during a discussion. Dialectical arguments attempt to resolve differences of opinions (Barth & Krabbe, 1982; van Eemeren & Grootendorst, 1992), so a good dialectical argument is able to resolve differences of opinion (van Eemeren & Grootendorst, 1992). Among the essential skills of argumentation, Kuhn (1991) included the abilities to generate alternative theories, envision conditions that would undermine those theories, and rebut alternative theories, all of which are essential to counter-argumentation. Counter-argumentation is a defining attribute of good arguments (Andriessen, Baker & Suthers, 2003; Voss, Perkins, & Segel, 1991) and a standard for assessing arguments (Kuhn, 1991). For ill-structured problems that may have multiple solutions, dialectical arguments seem more appropriate. Unfortunately, one of the most common weaknesses in argumentation is the lack of ability to generate counterarguments. When a person is asked to generate arguments for or against his or her own position, typically more reasons are stated supporting one's own position (Stein & Bernass, 1999). Students believe that if they identify counter-arguments, it would make their own argument less persuasive (Nussbaum & Kardash, 2005).

In order to support dialectical argument construction, Nussbaum and his colleagues have examined a variety of strategies, including the refutation strategy, the synthesizing strategy, and the weighing strategy (Nussbaum & Schraw, 2007). In the refutation strategy, students learn to recognize alternative solutions and to rebut those arguments. In the synthesizing strategy, students try to develop a compromise position that combines merits of both sides (''Is there a compromise or creative solution?''). In the weighing strategy, students must learn to evaluate alternative arguments and support the stronger argument based on the weight of evidence on that side of the issue (''Which side is stronger and why?''). When writing opinion essays, students need to integrate both argument and counterarguments in order to develop a conclusion (Nussbaum & Schraw, 2007).

There are three methods for scaffolding argumentation that have been most commonly investigated. The first includes prompts for constructing arguments. In a recent study (Jonassen, Shen, Marra, Cho, Lo, & Lohani, 2009), we contrasted evaluate and construct approaches. In response to engineering ethics problems that presented a variety of perspectives, theories, and canons related to those problems, students in the evaluate treatment were required to answer a series of questions, including:

- Which solution is better, solution 1 or solution 2?
- Whose perspective(s) support(s) your selection?

- Which theoretical approach(es) support(s) your selection?
- Which ethical codes support your selection?
- How might someone supporting the other solution disagree with your preferred solution?

Participants in the construct treatment were asked:

- What should you, as the engineer, do? What is your solution to this ethical problem?
- Whose perspective(s) support(s) this solution?
- Which theoretical approach(es) support your solution?
- Which ethical codes support your solution?
- What might someone else do? What alternative solution might someone recommend?
- What reasons would someone provide to support this solution?

Students who evaluated alternative arguments better supported their arguments on the immediate transfer task. They provided more elaborate discussions and justifications for their solutions to ethics problems.

Another approach to scaffolding argumentation is the use of note starters in essays or online discussions. Note starters consist of a menu of phrases, from which students begin the first sentence of discussion notes in an online discussion board. Nussbaum, Hartley, Sinatra, Reynolds, and Bendixen (2004) found that note starters encouraged students to consider other points of view. In a constrained online discussion, Oh and Jonassen (2007) found that students using note starters generated more evidence notes and also generated more hypothesis messages and hypothesis testing messages as well as problem space construction messages. Note starters provide prompts to consider alternative perspectives and can easily be implemented in problem-based learning environments.

The third approach to scaffolding counter-argumentation is the use of various graphic organizers or visualizations. The simplest (but perhaps most effective) may be Vee diagrams (Nussbaum & Schraw, 2007). The diagram illustrates the argument and counter-argument (see Figure 11.5). There are a number of computer-supported argumentation environments that help learners to visualize their arguments. Veerman and Treasure-Jones (1999) compared five systems for provoking and supporting collaborative argumentation during problem solving, including Dialab, Conference MOO, CTP (collaborative text processing), CLARE, and Belvedere. Cho and Jonassen (2003) found that students using Belvedere generated more coherent arguments and resulted in significantly more problem-solving actions during

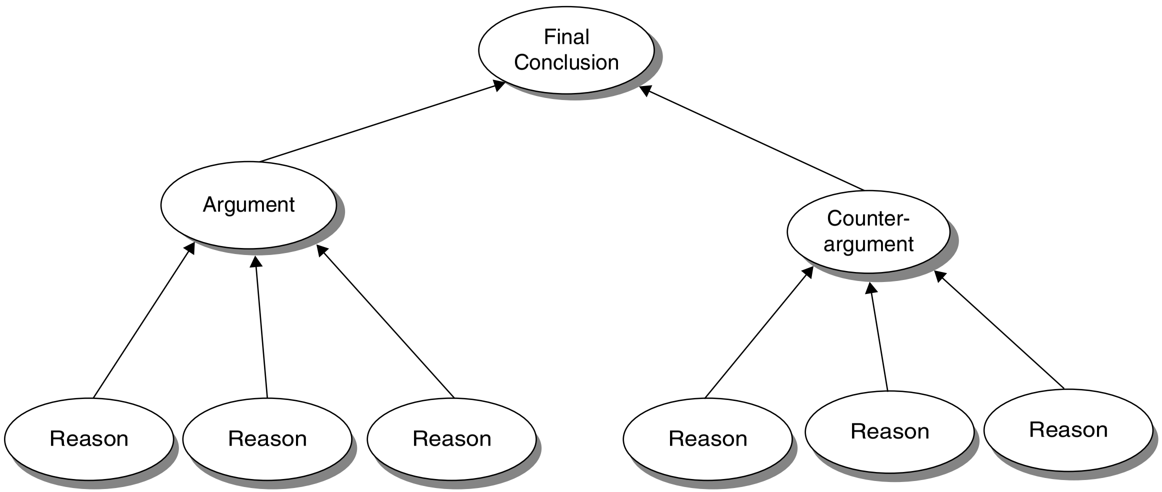


Figure 11.5 Argument Vee Diagram.

collaborative group discussions. The effects of the Belvedere scaffold varied for problem type. Groups who solved ill-structured problems produced more extensive arguments. Again, when solving ill-structured problems, students need more argumentation support because of the importance of generating and supporting alternative solutions. This study documented the close relationship between argumentation and problem solving, especially ill-structured problem solving. Although argumentation has not been empirically validated in many problem-based learning environments, argumentation is essential to problem solving and also provides an excellent source of assessment. That is, if learners can effectively argue in support of their solutions to problems, they are providing much richer information to the instructor about their comprehension of the principles involved in the problem.

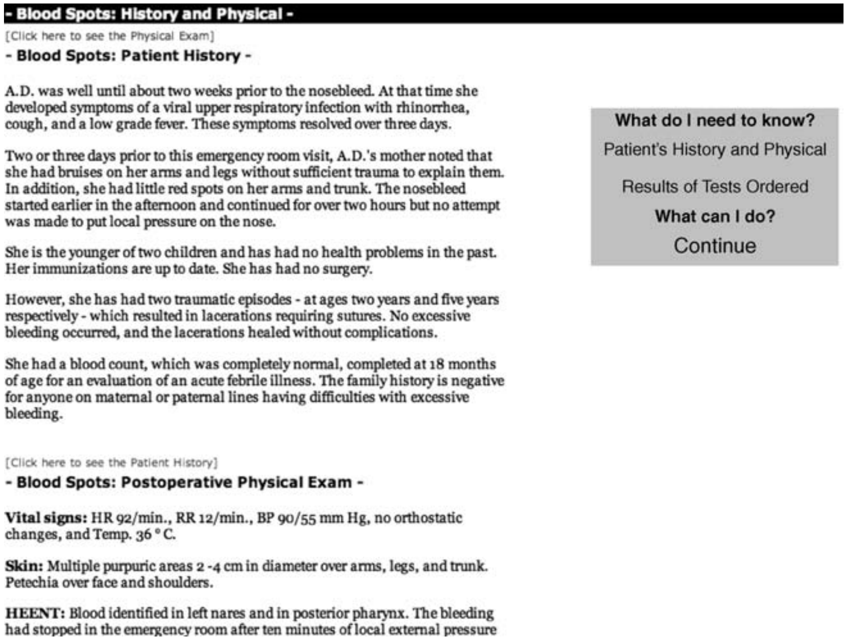
Relating Cases Cognitively

At this point, we suggest that the whole cases as exemplars, analogues, remindings, problems to solve, alternate perspectives, student constructions can be compared analogically, causally, or argumentatively. Cases as analogues and remindings should be compared analogically by comparing the structural relationship among its elements. In order to be able to transfer their problem solving, students must be able to articulate the causal relationships among the elements within and between problems. Students should also be required to develop arguments that compare the elements within and between cases. Although these recommendations are abstract and not fully tested, they comprise the best that we know now about problem solving.

SUMMARY

So little is known about designing instruction for problem solving. In over a decade's work, I have identified a dozen different kinds of problems that are important for people to learn how to solve. Most of my work has focused on the differences among problems and the resulting differences required in the instruction of those problems. While emphasizing differences among problems, it is also important to consider the similarities.

In this chapter, I have argued that the most important similarity in problem-solving instruction is its reliance on cases. Cases play a variety of different roles, including cases as examples, cases for analysis, cases as problems to solve, cases as analogues, cases as alternative perspectives, and student constructed cases. That is, cases form the foundation of problem-solving instruction. By way of illustration, we built a problem-based learning environment for diagnosing platelet problems in hematology. Figure 11.6 shows the results of the history and physical examination data, representing the problem to be solved. We also


- Blood Spots: History and Physical -

[Click here to see the Physical Exam]

- Blood Spots: Patient History -

A.D. was well until about two weeks prior to the nosebleed. At that time she developed symptoms of a viral upper respiratory infection with rhinorrhea, cough, and a low grade fever. These symptoms resolved over three days.

Two or three days prior to this emergency room visit, A.D.'s mother noted that she had bruises on her arms and legs without sufficient trauma to explain them. In addition, she had little red spots on her arms and trunk. The nosebleed started earlier in the afternoon and continued for over two hours but no attempt was made to put local pressure on the nose.

She is the younger of two children and has had no health problems in the past. Her immunizations are up to date. She has had no surgery.

However, she has had two traumatic episodes - at ages two years and five years respectively - which resulted in lacerations requiring sutures. No excessive bleeding occurred, and the lacerations healed without complications.

She had a blood count, which was completely normal, completed at 18 months of age for an evaluation of an acute febrile illness. The family history is negative for anyone on maternal or paternal lines having difficulties with excessive bleeding.

[Click here to see the Patient History]

- Blood Spots: Postoperative Physical Exam -

Vital signs: HR 92/min., RR 12/min., BP 90/55 mm Hg, no orthostatic changes, and Temp. 36 ° C.

Skin: Multiple purpuric areas 2 -4 cm in diameter over arms, legs, and trunk. Petechia over face and shoulders.

HEENT: Blood identified in left nares and in posterior pharynx. The bleeding had stopped in the emergency room after ten minutes of local external pressure

What do I need to know?

Patient's History and Physical

Results of Tests Ordered

What can I do?

Continue

Figure 11.6 History and Physical Exam as Problems to Solve.

provided analogous cases (Figure 11.7) for comparison to the existing case. Students were also allowed to order hematology, coagulation, chemistry, immunological, blood smear, and bone marrow tests at any time. Working through different analogous problems, students were encouraged to compare the cases for their structural similarities. Causality and argumentation were both engaged and supported during initial diagnosis, etiology, differential diagnosis,

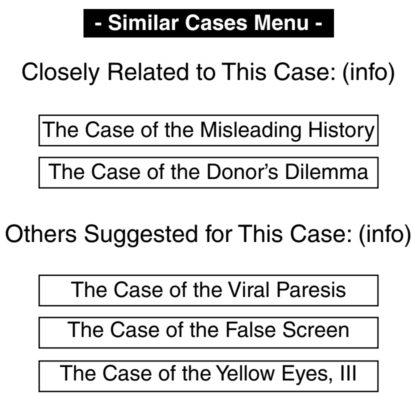


Figure 11.7 Analogous Cases.

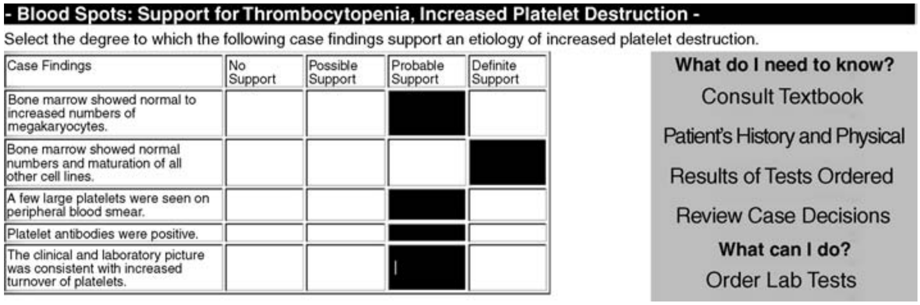


Figure 11.8 Argumentation Based on Causal Agents in the Case.

and case management prescriptions. For example, after determining etiology (an inference about causes), learners are required to justify their diagnosis based on case evidence (Figure 11.8).

The design prescriptions in this chapter are based on solid empirical evidence. However, no research has systematically examined the roles of these elements (functions of cases and kinds of reasoning) in supporting different kinds of problem solving. This chapter represents the state-of-the-art of my research at this moment in time, but it is an evolving theory. These components of problem-based learning environments could comprise a research agenda consisting of hundreds of studies. Those studies presuppose an understanding of and commitment to problem-solving instruction. In order for that to occur in the corporate culture, trainers must recognize that their organizations have problems (not just opportunities) and that employees should learn to solve them. Given that assumption, then they will need to begin constructing problem-based learning environments using the building blocks described in this chapter.

References

Ahn, W., Kalish, C. W., Medin, D. L., & Gelman, S. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition*, *54*, 299–352.

Andriessen, J., Baker, M., & Suthers, D. (2003). Argumentation, computer support, and the educational context of confronting cognitions. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 1–25). Boston: Kluwer.

Atkinson, R., Derry, S. J., Renkl, A., & Wortham, D. (2001). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research*, *70*, 181–215.

Barnes, L. B., Christiansen, C. R., & Moore, J. F. (1994). *Teaching and the case method.* Cambridge, MA: Harvard Business School.

Barrows, H. S., & Tamblyn, R. M. (1980). *Problem-based learning: An approach to medical education*. New York: Springer.

Barth, E. M., & Krabbe, E. C. W. (1982). *From axiom to dialogue: A philosophical study of logics and argumentation*. Berlin: W. de Gruyter.

Bransford, J. D., & Stein, B. S. (1984). *The IDEAL problem solver*. New York: W.H. Freeman.

Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In W. Friedman (Ed.), *The developmental psychology of time* (pp. 209–254). New York: Academic Press.

Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1147–1156.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*(2), 367–405.

Cheng, P. W., & Nisbett, R. E. (1993). Pragmatic constraints on causal deduction. In R. E. Nisbett (Ed.), *Rules for reasoning* (pp. 207–227). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cho, K. L., & Jonassen, D. H. (2003). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology: Research & Development*, *50*(3), 5–22.

Cognition and Technology Group at Vanderbilt. (1991). Anchored instruction and situated cognition revisited. *Educational Technology*, *33*, 52–70.

Cognition and Technology Group at Vanderbilt (CTGV). (1993). Designing learning environments that support thinking. In T. M. Duffy, J. Lowyck, & D. H. Jonassen (Eds.), *Designing environments for constructive learning* (pp. 9–36). New York: Springer-Verlag.

Didierjean, A. (2003). Is case-based reasoning a source of knowledge generalization? *European Journal of Cognitive Psychology*, *15*(3), 435–453.

Gagné, R. M. (1965). *The conditions of learning*. New York: Rinehart and Winston.

Gentner, D. (1983). Structure mapping: A theoretical framework for analogy. *Cognitive Science*, *77*, 155–170.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, *52*(1), 45–56.

Gentner, D., & Markman, A. B. (2005). Defining structural similarity. *Journal of Cognitive Science*, *6*, 1–20.

Gerjets, P., Scheiter, K., & Catrambone, R. (2004). Designing instructional examples to reduce intrinsic cognitive load: Molar versus modular presentation of solution procedures. *Instructional Science*, *32*(1–2), 33–58.

Gick, M. L, & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*(1), 1–38.

Hall, E. P., Gott, S. P., & Pokorny, R. A. (1995). *A procedural guide to cognitive task analysis: The PARI methodology* (Tech. Report AL/HR-TR-1995-0108). Brooks Air Force Base, TX: Human Resources Directorate.

Hernandez-Serrano, J., & Jonassen, D. H. (2003). The effects of case libraries on problem solving. *Journal of Computer-Assisted Learning*, *19*, 103–114.

Jacobson, M. J., Maouri, C., Mishra, P., & Kolar, C. (1995). Learning with hypertext learning environments: Theory, design, and research. *Journal of Educational Multimedia and Hypermedia*, *4*, 321–364.

Jacobson, M. J., & Spiro, R. J. (1995). Hypertext learning environments, cognitive flexibility, and the transfer of complex knowledge: An empirical investigation. *Journal of Educational Computing Research*, *12*(4), 301–333.

Jonassen, D. H. (1997). Instructional design model for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology: Research and Development, 45*(1), 65–95.

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology: Research & Development*, *48*(4), 63–85.

Jonassen, D. H. (2001). Can you train your employees to solve problems: If so, *what kind? Performance Improvement*, *40*(9), 16–22.

Jonassen, D. H. (2002). Engaging and supporting problem solving in online learning. *Quarterly Review of Distance Education*, *3*(1), 1–13.

Jonassen, D. H. (2003a). Designing research-based instruction for story problems. *Educational Psychology Review*, *15*(3), 267–296.

Jonassen, D. H. (2003b). Using cognitive tools to represent problems. *Journal of Research in Technology in Education*, *35*(3), 362–381.

Jonassen, D. H. (2004). *Learning to solve problems: An instructional design guide*. San Francisco: Pfeiffer.

Jonassen, D. H. (2006). Typology of case-based learning: The content, form, and function of cases. *Educational Technology*, *6*(4), 11–15.

Jonassen, D. H. (2007a). A taxonomy of meaningful learning. *Educational Technology*, *47*(5), 30–35.

Jonassen, D. H. (Ed.). (2007b). *Learning to solve complex, scientific problems*. Mahwah, NJ: Lawrence Erlbaum Associates.

Jonassen, D. H. (2008, May/June). Instructional design as design problem solving: An iterative process. *Educational Technology*, *48*(3), 21–26.

Jonassen, D. H., & Hernandez-Serrano, J. (2002). Case-based reasoning and instructional design: Using stories to support problem solving. *Educational Technology: Research and Development*, *50*(2), 65–77.

Jonassen, D. H., & Hung, W. (2006). Learning to troubleshoot: A new theory-based design architecture. *Educational Psychology Review*, *18*, 77–114.

Jonassen, D. H., & Hung, W. (2008) All problems are not equal: Implications for PBL. *Interdisciplinary Journal of Problem-Based Learning*, *2*(2), 6–28.

Jonassen, D. H., & Hung, W. (in press) All problems are not equal: Implications for PBL. *International Journal of Problem-Based Learning*.

Jonassen, D. H., & Ionas, I. G. (2008). Designing effective supports for reasoning causally. *Educational Technology: Research & Development*, *56*(3), 287–308.

Jonassen, D. H., Shen, D., Marra, R. M., Cho, Y. H., Lo, J. L., & Lohani, V. K. (2009). Engaging and supporting problem solving in engineering ethics. *Journal of Engineering Education*.

Klein, G. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.

Kolodner, J. (1993). *Case-based reasoning*. New York: Morgan Kaufman.

Kuhn, D. (1991). *The skills of argument*. Cambridge, UK: Cambridge University Press.

Kuhn, T. (1977). *The essential tension*. Chicago: University of Chicago Press.

Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentative reasoning. *Cognition and Instruction*, *15*(3), 287–315.

Lehman, D., Lempert, R., & Nisbett, R. E. (1988). The effects of graduate training on reasoning: Formal discipline and thinking about everyday-life events. *Educational Psychologist*, *43*, 431–442.

Loewenstein, J., Thompson, L., & Gentner, D. (2003). Analogical learning in negotiation teams: Comparing cases promotes learning and transfer. *Academy of Management Learning and Education*, *2*(2), 119–127.

Mayo, J. A. (2002). Case-based instruction: A technique for increasing conceptual application in introductory psychology. *Journal of Constructivist Psychology*, *15*(1), 65–74.

Means, M. L., & Voss, J. F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability, and knowledge levels. *Cognition and Instruction*, *14*, 139–179.

Merseth, K. K. (1994). *Cases, case methods, and the professional development of teachers*. Washington, DC: ERIC Clearinghouse on Teaching and Teacher Education.

Nussbaum, E. M., Hartley, K., Sinatra, G. M., Reynolds, R. E., & Bendixen, L. D. (2004). Personality interactions and scaffolding in on-line discussions. *Journal of Educational Computing Research*, *30*(1–2), 113–137.

Nussbaum, E. M., & Kardash, C. M. (2005). The effects of goal instructions and text on the generation of counterarguments during writing. *Journal of Educational Psychology*, *97*(2), 157–169.

Nussbaum, E. M., & Schraw, G. (2007). Promoting argument-counterargument integration in students' writing. *Journal of Experimental Education*, *76*(1), 59–92.

Nussbaum, E. M., & Sinatra, G. M. (2003). Argument and conceptual engagement. *Contemporary Educational Psychology*, *28*, 384–395.

Oh, S. C., & Jonassen, D. H. (2007). Scaffolding argumentation during problem solving. *Journal of Computer Assisted Learning*, *23*(2), 95–105.

Pokorny, R. A., Hall, E. P., Gallaway, M. A., & Dibble, E. (1996). Analyzing components of work samples to evaluate performance. *Military Psychology*, *8*(3), 161–177.

Popper, K. (1999). *All life is problem solving*. London: Routledge.

Quillian, M. R. (1968). Semantic memory. In M. Minsky (Ed.), *Semantic information processing*. Cambridge, MA: MIT Press.

Reed, S. K., Ackinclose, C. C., & Voss, A. A. (1990). Selecting analogous problems: Similarity versus inclusiveness. *Memory & Cognition*, *18*(1), 83–98.

Reznitskya, A., Anderson, R. C., McNurlin, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. Y. (2001). Influence of oral discussion on written argumentation. *Discourse Processes*, *32*(2–3), 155–175.

Ross, B. H. (1984). Reminding and their effects in learning a cognitive skill. *Cognitive Psychology*, *16*(3), 371–416.

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(4), 629–639.

Ross, B. H. (1989a). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(3), 456–468.

Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton, NJ: Princeton University Press.

Schank, R. C. (1990). *Tell me a story: Narrative and intelligence*. Evanston, IL: Northwestern University Press.

Schank, R. C., & Abelson, R. (1977). Script, plans, goals and understanding, Hillsdale, NJ: Lawrence Erlbaum Associates.

Schank, R. C., & Cleary, C. (1995). *Engines for education*. Mahwah, NJ: Lawrence Erlbaum Associates.

Schulman, J. H. (1992). *Case methods in teacher education*. New York: Teachers College Press.

Sembugamoorthy, V., & Chandrasekeran, B. (1986). Functional representations of devices and compilation of diagnostic problem-solving systems. In J. Kolodner & C. K. Riesbeck (Eds.), *Experience, memory, and reasoning* (pp. 47–53). Hillsdale, NJ: Lawrence Erlbaum Associates.

Simon, H. A. (1969). *The sciences of the artificial*. Cambridge, MA: MIT Press.

Smith, M. U. (1991). A view from biology. In M. U. Smith (Ed.), *Toward a unified theory of problem solving*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Spiro, R., Coulson, R., Feltovitch, P., & Anderson, D. (1988). Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains. Paper presented at the Tenth Annual Conference of the Cognitive Science Society, Hillsdale, NJ.

Stein, N. L., & Bernass, R. (1999). The early emergence of argumentative knowledge and skill. In J. Andriessen & P. Corrier (Eds.), *Foundations of argumentative text processing* (pp. 97–116). Amsterdam: Amsterdam University Press.

Sternberg, R. J., & Frensch, P. A. (1991). *Complex problem solving: Principles and mechanisms*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Strobel, J., Jonassen, D. H., & Ionas, I. G. (2008). The evolution of a collaborative authoring system for non-linear hypertext: A design-based research study. *Computers and Education*, *51*, 67–85.

Sudzina, M. R. (1999). *Case study application for teacher education*. Boston: Allyn & Bacon.

Thagard, P. (2000b). Explaining disease: Correlations, causes, and mechanisms. In F. C. Keil & R. A. Wilson (Eds.), *Explanation and cognition* (pp. 254–276). Cambridge, MA: MIT Press.

van Eemeren, F. H., & Grootendorst, R. (1992). *Argumentation, communication, and fallacies: A pragma-dialectical perspective*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Veerman, A., & Treasure-Jones, T. (1999). Software from problem solving through collaborative argumentation. In P. Coirier & J.E.B. Andriessen (Eds.), *Foundations of argumentative text processing* (pp. 203–220). Amsterdam: Amsterdam University Press.

Voss, J. F., & Means, M. L. (1991). Learning to reason via instruction in argumentation. *Learning and Instruction*, *1*, 337–350.

Voss, J. F., Perkins, D. N., & Segal, J. W. (1991). Preface. In F. Voss, D. N. Perkins, & Segal (Eds.), *Informal reasoning in education*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Voss, J. F., & Post, T. A. (1989). On the solving of ill-structured problems. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Voss, J. F., Wiley, J., & Sandak, R. (1999). Reasoning in the construction of argumentative texts. In J. Andriessen & P. Coirier (Eds.), *Foundations of argumentative text processing* (pp. 29–41). Amsterdam: University of Amsterdam Press.

CHAPTER TWELVE

High Engagement Strategies in Simulation and Gaming

Conrad G. Bills

INTRODUCTION

In this chapter we look at high engagement strategies in simulation and gaming. For the purpose of this discussion, simulation is considered as the acceptable representation of real situations or processes, used to provide instruction. Gaming in this context is a form of simulation used to provide instruction. Serious games use gaming technology and design principles for purposes other than entertainment. In this context, serious games are used for education and training. Simulation and serious games are selected as instructional media when they are less costly or less dangerous, when they are more convenient or more flexible, when they focus on important aspects of a training problem, and when they are designed to support and facilitate learning. Simulation and serious games can also be used in solving design and research problems.

Clark and Mayer (2008) found in their review of simulation and games research that there are challenges facing the instructional designer who wants to use simulation and serious games for instruction. A game or simulation can bring about cognitive overloading in ways that are counterproductive to learning. Simulations or games can lead to a great deal of activity but little learning. Since the acquisition of different types of knowledge and skill require different conditions of learning (Gagné, 1985), then an instructional experience or environment that does not include the instructional strategies required for the acquisition of the desired knowledge or skill will not result in learning. The desired outcome will not occur.

When high engagement strategies in simulation and gaming include instructional strategies, the focus should be on the development of tacit knowledge and integrated metaskills, thus capitalizing on the training capabilities for evolving novice learners into experts (Bills & Wood, 1999). This focus is accomplished by advancing learners to more complex, integrated learning activities in simulation and gaming so tacit knowledge develops and a higher-order skill is acquired, called a metaskill (Spears, 1985). Metaskills become the complex skills of adapting, monitoring, and correcting the use of individual skills in complex performances; the skills that integrate cognitive, perceptual, and motor processes. In other words, the system of instruction needs to not only include mastery of concepts and procedures, but also achievement of metaskills and the transformation to tacit knowledge.

In this chapter, learning theory and instructional design theory are applied to high engagement strategies in complex integrated education and training. In the military, this context is particularly applied in full mission simulation. The principles discussed come from several decades of training research and applied theory, particularly military flight training. The concepts, though abstract, attempt to explain the process of learning as a new learner takes on individual tasks and begins to integrate them into a complex system of operation such as an aircraft or an organization, becoming qualified to perform a mission, and then maintaining the high level of mission-essential competency (Colegrove & Bennett, 2006). Complementary research on serious games is still in its infancy, and a parallel set of design principles is emerging (see Chapter Thirteen on video game-based learning).

Instructional design for high engagement strategies applies basic concepts that lead to integrated skill performance and the design of training media within the context of a total instructional or training system. The Instructional System Development (ISD) process should identify the integrated activities and the desired instructional media most appropriate for the planned education and training activities. The instructional strategy for simulation and gaming should be designed to effectively achieve those integrated activities. The functions covered by the design should trace back to the education and training, support, and operations requirements. The concept of design used here is for instruction rather than engineering simulation. ISD in its broadest application is the systematic process for acquisition and implementation of a total education and training system (Bills & Wood, 1999).

Immersion creates a visual image of total mental and physical involvement in a simulated environment. The degree to which a learner experiences immersion is associated with how physical and cognitive reality is simulated. In order to understand the nature of cognitive reality, training development must also include cognitive analysis. High engagement strategies require a more complete understanding of the cognitive world. This includes understanding the cognitive aspects of problem solving, decision making, developing situational awareness, and other mental activities. Examples in which these high engagement strategies are applied are immersion in virtual, constructive, and live simulation; serious games, augmented reality, and Second Life.

Virtual, Live, and Constructive Simulation

Virtual simulation is the presentation in an artificial environment of the represented situation or process. The learner is the performer. Constructive simulation is a fully contained entity or set of entities that have been programmed with their own behavior and performance within the artificial environment. Constructive simulation operates in the virtual space and can perform in conjunction with or in opposition to the learner. Live simulation is presentation in a real-world environment in which artificially constructed situations are provided for the learner or team of learners. Current technology is providing simultaneous virtual, constructive, and live simulation to achieve a desired instructional outcome. For example, two fighter aircraft take off from the air force base. The virtual simulation represents each of the aircraft in the virtual simulation visual scene. Two virtual aircraft linked through simulation join the live fighters in a live/virtual mission activity. Constructive simulation is added to the virtual scene and the live aircraft sensor presenting four aggressive enemy aircraft. The two live aircraft work with the two virtual aircraft in an air defense scenario to defeat the aggressive progression of the constructive enemy aircraft. From the pilot's perspective, the same training has been accomplished from the simulator as from the in-flight aircraft.

Serious Games

Serious games can be defined as simulations that have the look and feel of a game, but correspond to non-game events or processes such as business or military operations. Serious games are used in instruction to engage the learner in a self-reinforcing context in which the learner becomes a player. Non-instruction purposes for serious games include marketing and advertisement (Serious Games,[1] 2008a). One serious games initiative is focused on using games in exploring management and leadership challenges facing the public sector. Initiatives like this provide productive links between the electronic game industry and the use of serious games in education and training (Serious Games,[2] 2008b).

Augmented Reality

Augmented reality (AR), also referred to as mixed reality (MR), is a combination of the real-world environment and virtual simulation that is typically interactive in real time. One example of AR was a real helicopter cockpit with instruments and windscreen overlaid by blue cutouts. The pilot wearing a helmet-mounted projected display was presented a full fidelity visual scene with active cockpit instruments replacing the blue cutout areas, the same technology used by the weather commentator for television (Morrish, Stoner, & Gurcak, 1997). According to IEEE and ACM, the field of AR/MR is highly interdisciplinary, bringing together technologies such as signal processing, computer vision, computer graphics, user interfaces, human factors, wearable computing, mobile computing, computer networks, displays, and sensors (Augmented Reality, 2008).

Second Life

Second Life is a 3D virtual world entirely created by its residents (Second Life, 2008). Second Life is the brain child of Philip Rosedale's Linden Lab, San Francisco, founded in 1999. Second Life resides on the Second Life Grid. For instructional applications, this platform offers the tools for business, educators, nonprofit organizations, and entrepreneurs to develop a virtual world of their own making. There are outside agencies that have services providing simulation development and content creation. Second Life differs from massively multiplayer games (MMORPGs) in two respects: (1) creativity: residents have near unlimited freedom to make this world whatever they want to make out of it; and (2) ownership: residents retain intellectual property rights over their in-world virtual creations.

Dimensions of High Engagement Strategies

Another way of looking at high engagement strategies is by the dimensions of live reality, augmented reality, and virtual reality. In any of these dimensions, the participant can become fully immersed and highly engaged in the scenario. When we design a high engagement strategy for instructional purposes, then we need to look at the principles to be applied in order for learning to take place.

Dr. Allison Rossett, in a discussion about the strategic value of simulations and games (Rossett, 2004), focused on two elements: (1) authenticity [fidelity] and (2) gaminess [immersion]. She depicted these elements in four quadrants, as shown in Figure 12.1. Authenticity relates to perceived parallels between the learning experience and the learner's life and real-world application. At one end of the continuum is an aircraft flight or a fire training simulator. At the other end is an in-basket exercise in leadership training during which the learner deals with multiple outputs, memos, policies, and requests in a timed exercise. Gaminess is about engagement, tracking, competition, and measurement against a standard. The combination of simulations and games shown in Quadrant 3 is the focus of this chapter. This combination of authenticity and

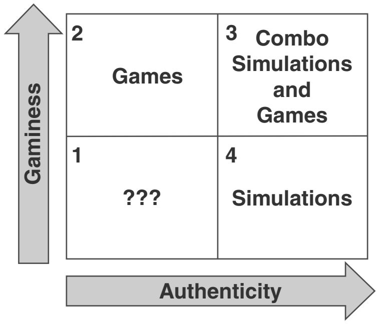


Figure 12.1 Elements of Authenticity and Gaminess with Respect to Simulations and Games.

Based on Rossett, 2004

gaminess wins the hearts and minds of the learner in achieving the desired instructional outcomes. Rossett concluded, ''In a world of knowledge work, global threats, regulations, compliance, and fierce competition, there are many good reasons to make tough choices in favor of the strategies associated with simulations and games.''

INSTRUCTIONAL DESIGN

This section addresses instructional design for high engagement strategies. The basic concepts presented here lead to integrated skill performance and the design of training devices within a total instructional or training system. The guidance on developing instruction offers suggested procedures to follow when high engagement simulation and gaming strategies are selected as media for training integrated activities.

The use of high engagement simulation is different from the past concept of replicating real-world equipment such as a vehicle or aircraft. Simulation has known capabilities that only a training device can support, such as repeated reset and predetermined conditions. Simulation can be used to regulate complexity or cognitive load by applying a simple-to-complex building block strategy. Simulation can be paused for providing feedback and then re-entered to apply this redirection in the situation. Simulation can also be used for training rarely occurring events such as emergencies or abnormal procedures that might take years to experience similar situations.

The context or ''big picture'' in which the high engagement simulation is used should also guide the design of the instruction. Long-term schema development is facilitated by this approach. Here the context of the real-world, operational application is cognitively connected to the patterns of instructional events much like a hat rack connects each different hat to the center pole. The functions covered by the total training system design should trace back to the instructional, support, and operational requirements. This concept of total training system design is for achieving instructional outcomes, rather than for engineering a high engagement simulation or game. This level of design will require a multi-disciplinary team approach.

Total Training Approach

The total training approach recognizes that the most profound facilitator of learning is building learning on previous learning. As a skill is learned in simulation, it transfers positively to performance in the vehicle or aircraft. The task-centered principle includes training a new skill in context of the ''big picture.'' As much as possible, meaning gained by the operator or pilot comes in a natural way. Situations are presented to facilitate cue discrimination for goal achievement. Those situations are presented in a broad context so that skills are more adaptable and become more robust with experience. The total training approach enhances the retention of skill, improves motivation for continued advancement, and leads to development of expertise.

Instructional design for a total training system is a systematic building block approach using a suite of media to achieve the terminal proficiency objectives (Bills & Burkley, 2002). This sequence of instruction is depicted as a development continuum depicted in Figure 12.2. The instructional requirements are appropriately allocated to each medium in the suite. The suite of media is designed to systematically build the foundation of entry-level learners, preparing them to perform complex integrated tasks in high engagement simulation and finally completed in the real-world vehicle or aircraft. When entry-level tasks become more complex, a shift occurs from the paradigm of a specific objective for each task to one of combining multiple objectives to achieve a common goal or enterprise. The traditional structural analysis approach needs a new dimension to capture enterprises. Dynamics such as those in crew interactions or interactions with systems outside the vehicle or aircraft (especially if hostile or unknown) should not be shortchanged. Critical enterprise requirements should be identified, particularly for requirements beyond initial qualification training.

Training Effectiveness Studies

Lessons on how to integrate high engagement strategies come out of several years of observation and study on how aircrew training simulation into a total training or instructional system (Bills, 1987; Nullmeyer & Rockway, 1984). The training effectiveness studies of high engagement simulation led to reviews of entire training systems. From one of these comprehensive studies

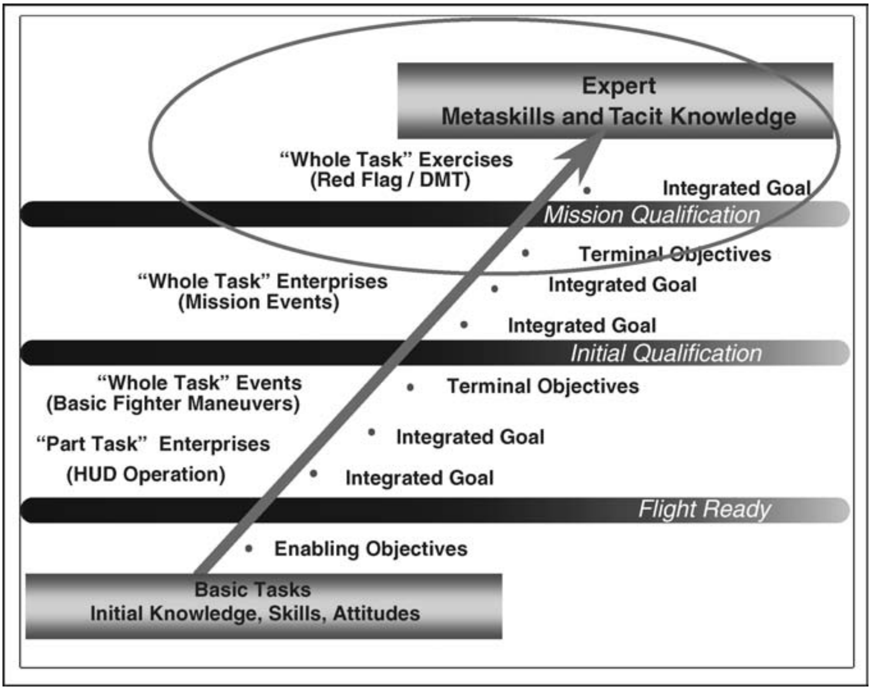


Figure 12.2 Development of Metaskills Through a Simple to Complex Training Continuum.
Based on Bills and Burkley, 2002

(Fishburne, Spears, & Williams, 1987) was derived from a model aircrew training system (MATS), which provided a method for instructional design of high engagement strategies. MATS began with a review of the total training system, keeping in mind the entire development life cycle of the operator or crewmember. Instructional capabilities of new high engagement simulations were defined and then appropriately incorporated. Through this total training system perspective, the effectiveness of the new instructional capabilities of high engagement simulation could be realized.

Train to Proficiency in Simulation. Results of training effectiveness studies identified that, when certain tasks were certified as trainable to proficiency in simulation, instructors shifted their focus from just doing events in the simulation to achieving the desired level of proficiency. Once this focus on proficiency was achieved, training effectiveness of high engagement simulation led to better use of the real-world training in the operational vehicle or aircraft. Instructors could then attend to those remaining events that could only be completed in the operational vehicle or aircraft. The conclusion from this research was that criterion proficiency of an aircrew training event first achieved in high engagement simulation results in a significant reduction in the number of attempts to attain proficiency in the aircraft (Bills, 1987; Nullmeyer & Rockway, 1984). Tasks can be certified as trainable to proficiency in simulation.

Techniques of Instruction. Techniques of instruction in simulation were shown as different from those used in real equipment or in flight. In the early years of flight simulation, Caro (1978) noted that instructors could be taught how to use these instructional techniques available to them in simulation. Examples of these techniques of instruction were adaptive training, performance playback, backward chaining, and automatic performance monitoring. Along with these techniques, instructional strategies were selected according to the objective or desired outcome of training. Methods of instruction included the ''building block'' approach and or individualized proficiency advancement. Based on this understanding, instructors should be given an opportunity to practice the selected technique(s) in high engagement simulator instruction before beginning with a new student. Even when the purpose of the simulator period is proficiency assessment by the expert judgment of instructors, the techniques used need to be standardized.

The complex training capability of high-end, full-weapon system simulation required considerable preparation of the student, first building up a skill set through a sequence of lower-level media and individual instruction before being scheduled for full crew training events. With this preparation, individual crew members were able to begin functioning as a team in high engagement simulation and benefit from the training capability available in a full-crew high engagement scenario (Bills, 1987; Nullmeyer & Rockway, 1984).

Evolution of Simulation Training. Early Weapon System Trainer (WST) implementation illustrates training evolution (Bills, 1987). The B-52 WST was three simulator stations linked together electronically for training operations as a full crew. One station was for the two pilots, another station was for the two navigators, and the third was for the two defensive systems operators. The Follow-On Operational Test and Evaluation (FOT&E) was conducted at a time when the bomb wing was tasked to produce twice the number of co-pilots the following year for the new B-1 bomber without an increase in flying hours. The only training variable available to the wing commander was the WST. The FOT&E design team formulated new training options that capitalized on the projected training capabilities of the newly operational B-52 WST as the training media for making up the difference in needed resources. Initial assessment of crewmember preparation for training in the WST revealed that when the co-pilots arrived, they lagged behind the well-prepared navigators and were not ready for crew interactions. The next iteration of the instructional design started with three independent crew station periods front-loaded for the pilots. This evolution in training ensured that prerequisites for each crew position were accomplished before full crew integration training started in the WST. With this individual preparation by crew position, when the full crew came together, they could all benefit from the complex training capability of full mission simulation.

As the evolution of WST implementation progressed, additional training options were developed using the ''building block'' approach to systematically develop crew proficiency. Each option was tested to determine degree of training effectiveness. A final training option was developed to determine the maximum number of training flights that could be reduced and still achieve a successful initial-qualification checkride. The final option demonstrated that, with select instructors and students training using a well-designed instructional strategy for integrating the WST, qualification could be achieved in almost half the number of flights. With this understanding that an option with fewer flights was workable under select conditions, the director of operations put enough flights back into the training design to be safe under variable conditions. As a result of this evolution in instructional strategy, the WST became the method used the following year for producing twice the number of co-pilots without an increase in flying hours. Subsequent to this initial study as WSTs were implemented in operational units, one commander evaluating the WST impact on his unit said, ''Before we had the WST, I thought we were mission-ready. Now we have the WST, I know we are mission-ready.''

Findings in Surgical Education. Surgical educators have taken lessons from flight training for incorporating simulation into their training system (Fried, Sadoughi, & Bills, 2006). Surgery and aviation are similar in enough ways that surgical educators developed an intense interest in the aviation paradigm of training. They found that high engagement strategies used for pilot development are mature applications that can transfer to other disciplines.

Principles for Promoting Learning. Clark and Mayer (2008) reviewed the research on simulation and games in e-learning, looking at what successfully brought about learning. They summarized their findings in the following games and simulations principles. They noted that feedback was the single most commonly mentioned success factor.

1. *Principle 1: Match game types to learning goals.* Align goals, activities, feedback, and interfaces of simulation and games with the desired instructional outcomes.
2. *Principle 2: Make learning essential to progress.* Ensure game progression and success translate into learning. More specifically,

promote generative processing aligned with instructional goals. In concert with this principle is the fundamental understanding that the design does not overload working memory at any given point during the progression sequence. In contrast, discovery learning does not pay off in learning.

3. *Principle 3: Build in guidance.* Build in guidance in ways that offer structure and learning support. Incorporate instructional explanations. This includes encouraging reflection on instructional content, managing complexity, and providing instructional support.
4. *Principle 4: Promote reflection of correct answers.* Learning from simulation and games improves with opportunities for the learner to explain the right answers. Reflection is needed for the learner to step back and abstract lessons learned from the game or simulation experience.
5. *Principle 5: Manage complexity.* Again, avoid mental overload; segment and sequence instruction. Methods for managing complexity include:
   a. Moving from simple to complex;
   b. Providing ''training wheels'' at first, then adding functionality until the desired level of performance is achieved;
   c. Start with a complete demonstration where the learner is the observer and the instructor is the performer. Next the instructor gradually allows the learner to assume task responsibility and provides correction as needed, Finally, the learner takes over fully as the performer, with the instructor as observer;
   d. Pace the presentation or scenario to avoid cognitive overload;
   e. Optimize interface fidelity. For the novice learner, remove irrelevant details that may deteriorate learning; and
   f. Provide instructional support. This is accomplished by first teaching how the game or simulation works and how to use it. Instructional support includes providing memory support such as a method for accumulating facts. Also include guidance on process or strategy and include visualization support such as adding an overlay graphic to focus attention or including a guiding line for showing direction.

Simulation Benefits for Advanced Training

Before high engagement fighter aircraft simulation, graduate-level training for fighter pilots had been primarily accomplished in the aircraft during exercises such as Red Flag. Red Flag is a simulated combat training exercise conducted over a huge flying range north of Nellis AFB, Nevada. Fighter aircraft and personnel come from all parts of the free world to make up the ''Blue Forces.'' These forces use a variety of tactics to attack targets such as airfields, missile sites, and tanks. The targets are defended by the enemy ''Red Force,'' which electronically simulates anti-aircraft artillery, surface-to-air missiles, and electronic jamming equipment. In addition, Red Force ''Aggressor'' pilots fly fighter aircraft using known enemy tactics. Thorough mission debriefings are based on the Red Flag Measurement and Debriefing System, along with TV ordnance scoring and threat video. Participants can replay the mission and learn what was done correctly and what needed work.

Training in the operational aircraft brought with it all the limitations of flight in the real world, including weather, lost time en route, fuel consumption, and environmental restrictions. In contrast, simulation put the safety of the pilot first and protected the aircraft as an increasingly costly asset. Accomplishing the same level of combat mission events in simulation had the potential of increasing the number of training events in the same timeframe by three to four times over aircraft in-flight training. A review of actual in-flight training in fighter aircraft showed that training benefit was often limited to ten minutes here and ten minutes there, with a total of twenty or thirty minutes of actual training over a two-hour flight. Nullmeyer and Rockway (1984), studying simulator training compared to in-flight training, showed that during a two-hour period in simulation there was an increase of 20 to 100 percent more opportunities than in the aircraft to experience critical events, which could be sequenced repetitively until proficiency was achieved.

High engagement strategies for simulation and gaming have a vision of immersion for the ultimate simulation capability. Immersion creates a visual image of total mental and physical involvement in a simulated environment. In order to achieve immersion, the whole of reality is simulated. Conceptually, this includes physical reality as well as cognitive reality (Bills & Wood, 1999).

Cognitive Reality

The more typical focus of instructional development has been on the simulation of the physical reality that supports cognitive processes. In contrast, the cognitive reality design approach articulates the knowledge and skills required for training and leads to performance that can demonstrate the accomplishment of the desired learning outcomes.

In order to understand the nature of cognitive reality, training development must include some form of cognitive analysis. Cognitive task analysis techniques are presented in Chapter Seven, Cognitive Task Analysis (Villachica & Stone). Advanced training requires a more complete understanding of the cognitive world in which the experienced learner operates. This includes understanding the cognitive aspects of problem solving, decision making, and acquiring situational awareness.

Another aspect of cognitive reality is understanding the learner's development of tacit knowledge and metaskills. Note that metaskills are closely related to cognitive strategies and metacognition. The novice learner is provided explicit methods for chunking task and knowledge elements. The next stage is providing methods for economy of thought used in organizing with schemas and models. High physical fidelity does not necessarily correspond to high cognitive fidelity. For novices, high physical fidelity may obscure high cognitive fidelity. This is due to excessive cognitive load as well as failure to discriminate cues. The novice learner needs preparatory training in which physical complexity is reduced and focus is on cognitive development of the individual role through a building-block approach to instruction. With this preparation, entering high engagement simulation can become immersive. Over time, aging and experience result in a transformation to expert tacit knowledge.

Metaskill

Spears (1983) defined a metaskill as the complex skill of adapting, monitoring, and correcting the use of individual skills in complex performances that integrate all learning processes. The person with metaskills can deal with even novel situations successfully. A metaskill is the skill required to adapt the specific skills it affects to the requirements of a situation. The more developed the metaskill, the greater the variation among situations that can be accommodated. The more discriminative the metaskill, the more likely that given adaptations will be appropriate and precise. A metaskill thus incorporates schemas for performance, becoming a transfer system that can serve as both a process and a product of training. It should be noted at the outset, however, that metaskills are typically complex systems or composites. They involve hierarchies of components and ranges of complex enterprises. The components will normally vary in nature, ranging from verbal systems to kinesthetic systems.

Metaskills become the complex skills of adapting, monitoring, and correcting the use of individual skills in complex performances; the skills that integrate cognitive, perceptual, and motor processes (Spears, 1985). The simple-to-complex training continuum shown in Figure 12.2 illustrates systematic metaskill development. Scenarios are developed that complement one another as well as provide a systematic increase in complexity. In fighter pilot simulation, this progression can go from single-ship to two-ship to four-ship then composite forces. Variations are introduced (both known and unexpected), keeping experiences within the realm of real-world applications. This manages cognitive loading during the development continuum. Advanced training is at the top of the continuum. By shifting from "filling squares" for continuation training to simulation designed for purposeful development, advanced training becomes a means to achieve the desired metaskill outcomes. With the acquisition of metaskills, a learner who is placed in a novel situation uses tacit knowledge to succeed. Therefore, in order for instruction to promote the development of advanced expertise, external instructional structure must provide experience in multiple situations over an extended time (Bills, 1997).

Tacit Knowledge

Acquisition of tacit knowledge is a developmental process. As knowledge develops, a change or transformation occurs that makes it more than the sum of its parts. This transformed representation called tacit knowledge takes the form in which it is applied in everyday life experiences (Polanyi & Prosch, 1975). This tacit knowledge now serves problem solving, situational awareness, prediction, and so on. The expert is therefore able to quickly get to the root of the problem, assess the situation, predict the outcome, and get it right. The development of tacit knowledge by learners progressing from novice to expert can be described as a shift from parts and procedures to a new meaning of the integrated whole.

An important facet of well-designed instruction is using a lesson structure that parallels or facilitates the development of the intrinsic, internal structures (Bills, 1997). The strategy of structure for developing tacit knowledge is the task-centered principle or a building-block approach. Instruction is planned to integrate some combination of verbal information, intellectual skills, and cognitive strategies, all related by their common goal. As learners advance to more complex, integrated learning activities, tacit knowledge develops and a higher-order skill is acquired, the metaskill (Spears, 1985).

Tacit Knowledge/Metaskill Analysis

Cognitive task analysis techniques discussed in a previous chapter can be expanded to include tacit knowledge and metaskills. This analysis includes: (1) the identification and description of tacit knowledge involved in expert performance and (2) the identification and description of the expert metaskills. Some may argue that traditional ISD procedures that specify enabling skills and knowledge tap these unseen variables. Although this may be true to some degree, it seems apparent that there is much to gain from a more thorough understanding and greater control of tacit knowledge and metaskills development.

The front-end analysis for the B-52 crew did identify the enabling skills and knowledge for each crew position. The combat crew training classroom instruction or academic phase was restructured by crew position. Sound principles of instructional design were applied throughout the process. When students came out of the academic phase for flight training, they still needed to learn to function as a crew. For example, a seemingly simple process of communicating over the aircraft intercom during a high engagement event would start out with crew members ‘‘stepping’’ on each other. As crew development progressed, each crew member transformed to an almost automatic set of ''my turn'' communications. Achievement of this outcome was more than application of knowledge and individual skills. There was an element of tacit knowledge developed as part of a metaskill.

Graduate-level training for fighter pilots has been primarily accomplished in the aircraft during exercises such as Red Flag. Accomplishing these complex mission events in high engagement simulation has increased the number of training events accomplished by three to four times over aircraft training. Traditional ISD procedures are effective in identifying critical events and then design opportunities in high engagement simulation to increase the experiencing of these critical events by focused repetitive instruction. Cognitive analysis of expert pilot tacit knowledge and metaskills associated with these critical events expands the breadth of understanding to mission-essential competencies. The benefit of training to proficiency in simulation before going to the aircraft is the reduction in follow-on time to proficiency in the actual aircraft. This understood, simulation scenarios can be designed to build not only proficiency in performing critical events, but also the more comprehensive development of the mission-essential competencies. Implementation exploiting training capabilities of high engagement simulation for tacit knowledge and metaskill development is a natural progression, as shown in Figure 12.2.

The F-16 Mission Training Center (MTC) was developed by Lockheed Martin for services to the United States Air Force Air Combat Command (USAF ACC) to allow pilots to train as teams, the same way they fight in combat. The F-16 MTC is a four-cockpit, high-end pilot training system that employs computer-generated imagery for a full field-of-view visual environment. The inauguration of the F-16 MTC at Shaw AFB, South Carolina (USA), opened a new era of advanced warfighter training. Pilots fly together from each of the full-fidelity cockpits connected through the local area network. The synthetic combat environment allows insertion of simulated friendly and opposing forces in accurate air-to-air and air-to-ground operations, embracing over one thousand training tasks, including abnormal and emergency procedures. Scenarios may be designed to a level of complexity consistent with the spectrum of assigned F-16 missions (Roberson & Bills, 2006). This level of training capability provided the opportunity for accelerating metaskill development and the transformation to tacit knowledge.

Integrated Activities

Instructional design for complex mission tactics and team processes in high engagement must include approaches for integrating participant activities. Recall that the B-52 WST used for crew training initially had different crew-members arriving at training with varying levels of preparation for the planned scenario. Co-pilots lagged behind the well-prepared navigators and were not ready for crew interactions. The training was changed to ensure that prerequisites for each crew position were accomplished before full crew integration training started in the WST. With this individual preparation by crew position, when the full crew came together they could all benefit from the complex training capability of full mission simulation. In a short number of scenarios designed in a building-block approach, the crew was able to successfully accomplish the composite graduate-level mission.

Employing these integrating strategies will build individual cognitive schema. Combining activities into an event with a specific goal is employing enterprises. Progression through training, aging, and experiencing will achieve tacit knowledge and metaskills (Fishburne, Spears, & Williams, 1987). Seasoned instructors know that achievement of initial proficiency during training is only the beginning of expertise development. Placing the student in broadening and systematically challenging scenarios in simulation concurrent with real-world experiences can hasten the process of expertise development.

Schema. A schema governs cue and response selectivity so the learner is attuned to goals and conditions of performance. Training begins with establishing schema and then uses these schema in progressively more complex activities or enterprises. Schema development is a cognitive activity.

Enterprise. An enterprise is a purposeful, planned activity that combines multiple objectives in pursuit of a comprehensive purpose or goal. Progression moves from simple individual objectives to the more complex enterprises, developing along the simple-to-complex continuum. Advanced training at the upper end of the continuum has enterprise activities that become complex events requiring situational awareness and a sense of team interaction.

Metaskill continuum. The vision of advanced training is fulfilled with the outgrowth of expert metaskills that are bolstered by tacit knowledge. The training capability of integrated simulation and gaming involves the diverse qualifications of the broader team.

Understanding Cognitive ''Whole.'' Tacit knowledge and metaskills identify important basic components of the missing link for achieving the vision of advanced training. Tacit knowledge and metaskills help us understand the cognitive ''whole.''

A productive approach is through the design of enterprises or mission events that lead to desired integrated outcomes, presented in a variety of different contexts. An enterprise or mission event is formed by bringing together multiple objectives for the purpose of achieving a common goal or mission outcome. The complexity of the events progresses from simple-to-complex in a task-centered, building-block approach, building learning upon learning. Part-task training is integrated into whole-task training. For example, in-flight training cockpit procedures transition into basic flight maneuver training, first in simulation and then in the aircraft. Training progresses along the continuum toward achievement of metaskills and transformation to tacit knowledge.

FIRST PRINCIPLES OF INSTRUCTION

High engagement strategy in simulation and games requires the application of the first principles of instruction. As used here, principle is defined as a relationship that is always true under appropriate conditions, regardless of program or practice. A program is a prescribed set of instructional methods designed to teach a body of content or skill. Practices are those instructional activities implemented by an instructor or an instructional system. We can learn about the first principles of instruction from Dr. M. David Merrill (2008). He reported that an amazing number of these instructional products were surprisingly ineffective and that some did not teach at all. This is an important consideration for high engagement strategies projected for use as instruction. The First Principles of Instruction also apply to design of high engagement strategies.

- *Task-centered principle.* Learning is promoted when the strategy is in context of whole real-world tasks that progress in a building-block approach. The student driver is placed in high engagement strategies that begin with realistic road situations that have minimal traffic and unforeseen obstacles and continues in progression of complexity until the student operates successfully in heavy traffic with several unforeseen challenges.
- *Activation principle.* Learning is promoted when learners activate relevant prior knowledge or experience. Following classroom instruction on maintaining safe driving distance, the student is placed in a simulated driving scenario of a congested freeway and then suddenly another driver in front puts on the brakes such that without safe driving distance an accident will occur. Activation is enhanced when students come together after completing the same simulation and share their experiences with each other. Activation is enhanced when the instructor initiates the learning experience by building a ‘‘big picture’’ structure for organizing the new knowledge. This can be done by the instructor taking his student drivers for a ride in the local area and presenting them with the situations they will be learning to deal with and providing them the framework within which he can make references for guidance during demonstration and also while coaching during application.
- *Demonstration principle.* Learning is promoted when learners observe a demonstration consistent with the content being taught. The

demonstration is enhanced when the instructor guides the student driver to related information and the prior framework for organizing new knowledge. The demonstration is enhanced by the instructor involving the students in discussion about what they have just seen. The instructor can also enhance the demonstration by having the students observe a video clip of relevant, similar real-world situations.

- *Application principle.* Learning is promoted when learners engage in application of their newly acquired knowledge or skill that is consistent with the content being taught. The instructor puts the student driver in a simulated scenario, which integrates the knowledge and skills recently learned by applying them in the pre-planned situation. This instruction is effective only when the student receives built-in or instructor-provided corrective feedback. The instructor coaches using the principle of decreasing involvement as the student's proficiency for performing the task increases. This way the student performs correctly early and builds a habit pattern of correct response. Again, application is enhanced by guided peer-to-peer discussion and peer demonstration as well as media replay.
- *Integration principle.* Learning is promoted when learners integrate their new knowledge and skill into everyday life. The instructor directs his student drivers in a discussion of personal experiences when they used their new knowledge and skill. The discussion is enhanced by peer critique, as appropriate. The instructor can explore with his students additional ways they can use their new knowledge and skill. Integration is enhanced as students have opportunities to publicly demonstrate their new knowledge and skill.

A complete instructional strategy is problem based, involving the learner in four distinct phases of learning as shown in Figure 12.3, based on the first principles of instruction (Merrill, 2002). High engagement strategies begin with activation of prior experience. This is followed by demonstration of skills, then the application of skills, and finally integration of these skills into the real-world environment. A top-level instructional design prescription for high engagement strategies puts learners engaged in solving real-world problems. Existing knowledge is activated as a foundation for new knowledge. Learning is facilitated when new knowledge is demonstrated to the learner, applied by the learner, and then integrated into the learner's world.

Development

The training task list (TTL) is an important starting point for training development. The TTL provides for the task hierarchy required for mission success. The TTL does not account for the integration of multiple objectives to form

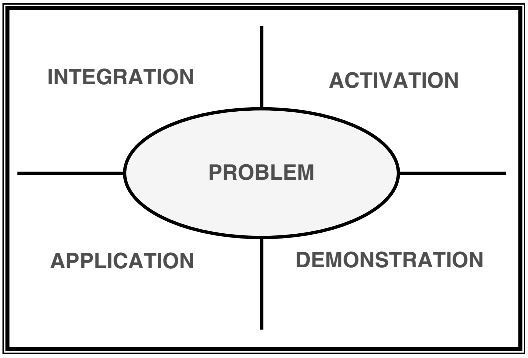


Figure 12.3 Instructional Design Prescription for High Engagement Strategies.
Based on Merrill, 2002

enterprises and develop mission-essential competencies. The variety of full-task events needs to correspond with mission requirements as well as the multiple situations in which the trained operator or organization's performer will be employed.

The goal of training development is to achieve a program that hastens the normally slow acquisition of metaskills. The program must account for the complete progression from novice to expert. Integrating multiple simulation devices linked for team training will enhance this advanced training. In addition, feedback is enhanced with multimedia briefing and debriefing stations that use timeline-based capability for achieving situational awareness and focusing on key points of interest. In time, through experiencing a wide variety of conditions, the expert emerges, having achieved the repertoire of metaskills integrated into real-world applications.

Performance Measurement. Performance measurement is crucial for verification of instructional strategies. Performance measurement supports instruction by monitoring learner performance and providing feedback for learning activities and practice performance. Kaufman (1998) suggested that in the past, the approach to dealing with the higher levels of improvement (Kirkpatrick's Levels 3 and 4; Kaufman and Keller's Level 5) has been ''dumbed down'' because of inability to measure, the feeling that it was too complex, too confusing, or too difficult to accomplish. He said, ''If performance improvement specialists only deal with that which is not 'confusing' and 'difficult,' then who takes responsibility for airplanes falling from the sky?'' (p. 23). The Air Force Performance Effectiveness and Evaluation Tracking System is an example of how to achieve measures that assess mission-essential competency development (Carolan, MacMillan, Entin, Morley, Schreiber, Portrey, Denning, & Bennett, 2003). Competency can be defined as a collection of knowledge, skills, abilities, or characteristic associated with a specified level of job performance (Brock, 2007). Competency-based training provides a seamless stream of consistent performance data, regardless of the live or virtual training methods. The ability to measure performance against standards established by the user community not only provides control but also defines consistent measurement criteria across a single operational system at multiple locations (Colegrove & Bennett, 2006). Referring back to Figure 12.3, it is of particular importance to assess integration and application of metaskills and tacit knowledge. Often assessment tasks impose an artificial structure that simplifies the task, lowers cognitive load, and does not call for either integration or tacit knowledge. This is a common error in assessment design, which violates the principle of cognitive fidelity as we say here.

Integrated Training Information Management. An important consideration for the development of an instructional system incorporating high engagement strategies is an integrated training information management system (Bills & Wood, 1999). Integrated training information management gives operators, maintainers, and evaluators the information environment they need to ensure efficient and effective training throughout the operational system lifecycle. Technology is available for providing a capability to collect, process, and manipulate information in ways that support performance improvement.

The goal of integrated training information management is to achieve a training system information highway that serves the needs of each defined user. The long-range goal is to improve efficiency and effectiveness of the training system, thus reducing the lifecycle cost.

A training database is the heart of an integrated instructional system. This database provides both horizontal and vertical traceability to all training system requirements. Horizontal traceability is consistent with the ISD process, tracking training requirements through implementation of the operational view. Changes can be traced back to requirements, as well as to the simulation variables throughout the life of a program. Vertical traceability reaches to all related components in the instructional system, particularly to the performance specification for the simulation device. Instructional decisions documented by the training database show instructional strategies, course sequence, and content selection. Technology assessment and resultant simulation fidelity requirements become the foundation for the simulation device functional description. The next generation of sinus surgery simulation was a combination of integrated imaging techniques, virtual haptics simulation, and environment for mastery of concept structure and procedures. Simply measuring performance of structured procedures is insufficient here. The challenge is to design measures of the cognitive tasks, particularly at the level of complexity performed, so that an accurate assessment of proficiency is accomplished.

Implementation

As high engagement strategies for simulation and games become a reality, learners are scheduled for instruction much the same way they are scheduled for training in the real-world vehicle or organizational environment. However, the realization of this vision will be dependent on the design of the instructional system as a whole, the application of the first principles of instruction, and the systematic implementation of high engagement strategies. The challenge is to provide advanced graduate-level instruction for transformation to tacit knowledge and development of expert metaskills. An example of graduate-level instruction is the F-16 Mission Training Center (MTC).

F-16 Mission Training Center (MTC). The F-16 Mission Training Center (MTC) was acquired for the U.S. Air Force Air Combat Command (ACC) in response to the call for a simulation that would allow pilots to train as teams the same way as they fight in combat (Bills & Devol, 2003). The F-16 MTC is a four-cockpit, high-end pilot training system that employs computer-generated imagery for a full field-of-view visual environment. Pilots fly together from each of the full-fidelity cockpits connected through the local area network. The synthetic combat environment allows insertion of simulated friendly and opposing forces. The geospecific, phototexture database supports accurate air-to-air and air-to-ground operations. As Distributed Mission Operations (DMO) expanded to a wide-area network, other remotely located players joined in to provide multiship, multiplatform warfighter interaction without the real-world environmental restrictions or range limitations. Looking at the simple to complex training continuum in Figure 12.2, the F-16 MTC was designed to train at the upper end where metaskills and tacit knowledge are developing. The F-16 MTC was for graduate-level training; therefore, the instruction design approach started with the combat mission training problem definition and then identified the activation, demonstration, application, and integration required to succeed. Rather than using the more traditional approach of delineating the TTL down to the functional object or cue, mission-training events were defined by an overall goal, the training conditions required to achieve the goal, and the performance standard to be used in determining when the goal was achieved. This approach allowed teasing out metaskill and tacit knowledge without imposing an artificial structure typical of basic initial-skills training. SME input and customer confirmation during the Training Capability Requirements Analysis process established and refined the training descriptions and simulation conditions to achieve the real-world application or operational view of the training system. This operational view definition drove the functional design for the high engagement simulation, which was in turn allocated to the physical architecture for the training devices.

Table 12.1 F-16 MTC Training Task List Training Capability Requirements

Object Identifier	*TTL_101*
Task ID Code	3.1.2.3
Training Task Description	Perform Mil Pwr takeoff as wingman
Training Description	Wingman maintains position by observing cockpit visual signals of flight lead and anticipation of power changes necessary to maintain position
Condition(s)	Aircraft control actions same as actual aircraft. Visual system will produce normal cues for maintaining formation position with other networked ownerships. Formation takeoffs will be accomplished only with locally networked ownerships.
Comments/Issues	Radio calls will replace in-cockpit visual signals in MTC (runup and brake release)
Reference Pubs	T.O. 1-F-16CJ-1, T.O. 1-F-16CG-1, AFI 11-2F16, Vol 3
Links To	SubSpec–Visual VS74; SubSpec–Level BC FFOV Runtime FFOVIG63; SubSpec–Cockpit SSCkpt_231; SubSpec–Cockpit SSCkpt_84; ProdSpec–Aural OWNA165

An example of TTL training capability requirements definition is shown in Table 12.1. The *training description* was provided by SMEs to further define the given TTL task in narrative form. The description included those behaviors unique to performance of the task at the graduate level. The *condition(s)* defined the MTC simulator setup and scenario required for accomplishing the task. These conditions were also used to define the level of fidelity required by the simulator elements in order for the training capability to be achieved. The set of conditions helped the customer understand early the intended approach to accomplishing the tasks and minimize conflict when the system was fielded. Using the example in Table 12.1, the condition description made it clear to the customer that there was no intention to supply an artificially intelligent flight lead in the synthetic environment for the purpose of training formation takeoffs.

During the process, appropriate references were identified to document USAF publications that may further define performance at the graduate level. Related comments and issues were also recorded, particularly if there was an impact on training capability for metaskill development such as expected technology limitations.

In order to achieve integration in the F-16 MTC at the same time regulating cognitive load, scenarios were created in a simple-to-complex developmental sequence. These MTC mission-scenario scripts were designed by SMEs to include all elements necessary to achieve high engagement strategies. Demonstration of F-16 MTC training capability used these mission scripts to conduct mission-oriented testing. The mission scenarios were set in typical real-world situations to which pilots may be assigned.

Assessment traces and reports progression to metaskill mastery and transformation to tacit knowledge. Grade sheets were based on the aircrew training approach of how much involvement is required by the instructor during task performance. Mission-essential competencies incorporated cognitive components in the assessment implemented by the pilot-effectiveness tracking systems. The end result was simulation mission-training design for graduate-level training in real-world operational environments rather than for conventional engineering that replicates the operational aircraft.

Command Value of Virtual Training. The 20th Fighter Wing at Shaw AFB, South Carolina, considered the value of virtual training to be close to that of the live combat training they executed on a daily basis. Colonel Joel ''Bugsy'' Malone, 20th Operations Group Commander, described their F-16 MTC as the key driver for their success in training transformation. It was considered a powerful force multiplier of limited assets in an environment of decreasing training opportunities. The MTC enabled their ability to maximize training time across the entire spectrum of F-16CJ+ operations. Colonel Philip M. Ruhlman, former 20th Fighter Wing Commander, said:

> ''We have demonstrated Combat Air Force wide value for F-16 MTC integration. Every flight in the MTC has a positive impact on combat readiness. The MTC takes simulation to a higher level, giving us a clear edge in preparing for tomorrow's battles. My wing's F-16CJ+ transformation has clearly been enabled by the training capability and capacity of our MTC.''

''William Tell'' is the name given to the Air Force's premier air-to-air weapons competition that tests aircrew performance in the air dominance and air sovereignty missions, while evaluating weapons employment and tactics. The 20th Fighter Wing Weapons and Training Flight devised a way to adapt the William Tell format to an F-16 MTC simulated environment in air-to-air competition called the Turkey Shoot. This virtual combat training competition had the look and feel of a real-world competition. The training plan incorporated at least two virtual Turkey Shoots per year. The MTC Turkey Shoot was a great example of how to incorporate advanced warfare training simulation into a composite combat training plan.

FOUNDATIONS OF LEARNING

Understanding high engagement strategies for simulation and gaming requires insight into the complex human process of learning. This understanding starts with the foundations of learning. This section applies learning theory to the design process of complex integrated instruction, particularly in high engagement strategies for simulation and gaming.

The context of this discussion comes from several decades of aircrew training research and applied learning theory documented in an unpublished manuscript of Dr. William D. Spears (1992). An abridgement of Dr. Spears' manuscript prepared by Conrad G. Bills (1993) was used as the text for *Information for Designers of Instructional Systems: Design Guide for Device-Based Aircrew Training,* Department of the Air Force (1993). The concepts, although abstract, attempted to explain the process of learning as a new learner took on individual tasks and began to integrate them into complex aircraft operation, mission qualification, and then continued combat readiness. This learning development continuum is often referred to today as life-long learning. A diagram of developmental learning is shown in Figure 12.4. In the early stages shown at the left of the diagram, the novice learner handles tasks and knowledge elements explicitly. Cue and response development occurs, along with discrimination and generalization. The identification of common relationships that can be facilitated by instruction results in chunking. Chunking is the early stage of economy

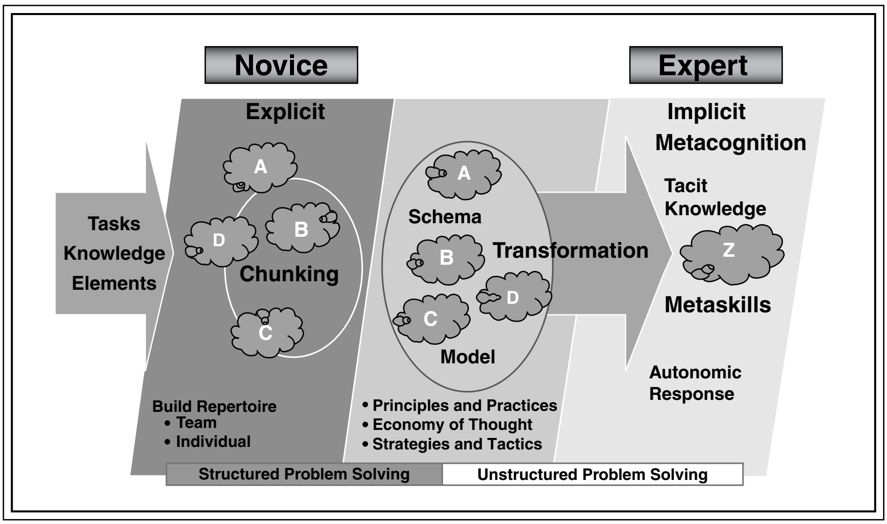


Figure 12.4 Diagram of Developmental Learning from Novice to Expert.
From Bills, 2009

of thought facilitating recall. Through individual and team interactions with the environment, the novice begins to build a repertoire for schema development. Instruction can provide an organization of cue-response patterns. Schemata form when cue and response selectively become attuned to goals and conditions for performance. Models guide the application of principles and practices, strategies and tactics. Through aging and experiencing, a transformation takes place that makes acquired knowledge more than the sum of its parts, tacit knowledge. Metacognition is possible with cognitively complex, integrated knowledge structures. Skills become more robust, adapting, monitoring, and correcting in complex performance that integrate all learning processes for metaskill. The expert is emerging, now more implicit, seemingly autonomic in responses. Successful performance is seen in even novel real-world situations able to deal with ill-structured problem solving. The expert is sought out for guidance and direction, but has difficulty providing descriptions due to the transformation.

The ultimate instructional value of high engagement strategies depends on the extent to which learning will transfer to real-world applications. In other words, the value of simulation in a driver training device is assessed against the transfer to driving a car in real-world situations. Three factors affecting this transfer to the real world are cue development, cue and response discrimination, and generalization. Associated with these factors is mediation.

Cue Development

Being able to interpret cues and respond to them appropriately is the essence of instruction for skill development. Thus, high engagement simulation used for instruction should focus on cue development. For example, a new driver education student could learn the cues provided in simulation associated with an entry-level behavior, such as controlling direction of travel, or cues associated with higher-level behavior, such as safely merging from an on-ramp into busy freeway traffic.

Cue and Response Discrimination, Generalization

Discrimination refers to the process of telling the difference between one stimulus and another. Skilled real-world performance requires innumerable discriminations of cues and cue patterns. Depending on the real-world situation and stage of an activity being performed, required discriminations may be almost entirely cognitive in nature, but no less complex. The difference, usually, is that purely cognitive discriminations by adults can be formulated according to verbal rules. Indeed, language can be described as essentially an elaborate discriminative and generalizing system.

Responding appropriately to cues requires learning to discriminate the cues and responses appropriate to a given task. Although the focus of instruction is on teaching the discriminations underlying performance such as driving a car, a driver training device has a distinct training advantage over the actual car in that discriminations can be taught in simulation at times and in ways that promote the most efficient development of skills.

It is the cue information available in a training device, rather than the stimulus fidelity of the simulation per se, that should be the criterion for deciding what skills or skill components are to be taught with high engagement simulation or gaming. Cognitive fidelity includes getting the cues right and getting the decisions right. Physical fidelity is appearance. Too much physical fidelity too early can cause excessive cognitive load and interfere with learning. The new pilot ill-prepared for crew interaction in the B-52 WST was overwhelmed by the physical fidelity on his first mission to the point of negative training.

Generalization

Once discrimination is learned in a given situation, then generalization must take place in order to apply discrimination to different but related situations. Generalization refers to the extension of cue and response discriminations across time and situation. Without generalization, all learning would be specific to the time and situation of the learning. Just responding the same way on two different days requires that minor or major differences in the stimulus be accommodated, thus deriving cues from what the situations have in common.

In complex behavior, it is usually necessary to key on cue characteristics that are often subtle, especially those arising from feedback of a kinesthetic nature (resulting from body position, presence, or movement). Hence, desirable generalizations require highly discriminated frameworks of cues and responses to be made to them. Trainees practice under conditions with sufficient variation for the trainees to learn to identify which common cues to key on and how to tailor actions to the nuances of cues and contexts. The result is a broad context for associating cue meanings and selections of responses with situational requirements. In other words, the result is a generalized discriminative system.

Because transfer is based on cue information, all valid cue interpretations learned in a training device can be generalized in both simulation and the actual equipment such as a driver performing in a trainer and then similarly performing in the car. Thus, it is important that high engagement strategies using simulation or gaming focus on the generalizable meanings of the cues they can provide.

Mediation

Mediation is the *intermediary* process that provides meaning for the situation at hand when generalizing from previously learned discriminations. Mediation is a complex psychological process important to understanding high engagement strategies. Mediation permits a person performing in a simulation or game at various levels of fidelity to acquire meanings similar to the meanings of corresponding performance in the real situation. Mediation comes between, or mediates, the acts of sensing a stimulus and responding to it. Thus, through mediation, instruction with simulation or gaming is not limited to similarity between training device and real equipment. Through mediation processes such as language and mental rehearsal, the effectiveness of the high engagement strategy is enhanced.

Principles of Learning

Associated with an understanding of developmental learning are principles of learning to be applied to the design process of complex integrated instruction, high engagement strategies for simulation and gaming. The remainder of this section covers the following principles of learning: cognitive processes, feedback, guidance, sequence, allocation, organizing practice, and duration and frequency.

Cognitive Processes. Effective and efficient instruction depends on preparation for learning. That preparation takes place in classrooms, briefings, discussions, simulations, and real situations, and in a variety of other formal and informal settings. Much of this preparation involves cognitive processes, that is, thoughts, ideas, mental images, and concepts that are both verbal and nonverbal. It basically consists of the cognitive foundations for instruction using high engagement strategies, and ultimately, real-world performance.

The principles of learning for cognitive processes are summarized as follows:

- A meaningful context should be provided for the learning of skilled performance in simulation or gaming (task-centered principle).
- Mediation should be employed systematically in teaching cue and response discriminations (activation principle).
- Briefing for an instructional simulation or gaming session should anticipate and guard against a student's relating what he or she does during instruction only to the device itself (activation principle).
- When students are practicing in the training device, they should be instructed to think about the meaning and effects of their actions with respect to requirements for performance in the real-world situation (application principle).
- Recognizing that transfer is a mediational process, a major goal for simulation or gaming for instruction should be to exploit the types and uses of mediation that maximize transfer (integration principle).

Feedback. Feedback refers to information that informs the learner about the results or effects of the learner's actions. Feedback aids in maintaining student motivation to learn and also informs the student of the appropriateness or inappropriateness of interpretations of cues and responses to those cues.

The following principles apply to the role of feedback for instruction using high engagement strategies (application principle):

- Feedback determines the nature and extent of discriminations that will be learned.
- Feedback should focus on specific aspects of cognitive, perceptual, and motor actions that must be discriminated.
- Augmented and supplemental feedback should be used only when task-intrinsic feedback cannot be discriminated by the learner, especially early in skill acquisition.
  - Its use should be specifically for the purpose of teaching cue and response discriminations intrinsic to the real-world task.
  - Timely supplemental or augmented feedback should be used to signal the availability of intrinsic feedback.
- When intrinsic feedback is needed as a signal for a subsequent action, the training device must provide it quickly enough to avoid disrupting the action.
- Feedback will help maintain student motivation if the feedback ensures progress.

Guidance. Guidance is the directing of a learner's actions toward a desired goal. Guidance may include modeling of desired behavior. The actions being guided may be thought processes, physical movements, oral communications, selection of cues, or processing of cue information. Guidance is involved every time an instructor comments to a student about what the student should do, remember, or think about. Guidance sets the stage for knowledge structure development, including chunking and schemas. Properly used guidance helps speed learning when it identifies desirable cues and responses that the student cannot recognize without guidance. Guidance sets the foundation for metaskill development and eventual transformation to tacit knowledge. By identifying correct cues and responses, guidance reduces the likelihood that inappropriate cues will be used and incorrect responses made. Thus, guidance helps prevent learning of erroneous actions that would eventually have to be unlearned.

Guidance should be used during advanced training when needed to focus learner attention on new cue and response discriminations or on previously mastered discriminations that can be generalized to new tasks. Guidance can be valuable for experienced performers when their skills have deteriorated appreciably or when the need is to define or clarify standards for performance.

Guidance should focus on aspects of skill performance in high engagement simulation or gaming that are transferable to the real world or that can promote transfer.

The following principles apply to the use of guidance in high engagement strategies:

- The purpose of guidance is to focus the student's attention on correct cue and response discriminations and to avoid incorrect cue interpretations and actions.
- To avoid dependence on guidance, it should not be used when the student is able to make the required discrimination without help.
- Contrasting, through guidance, desirable with undesirable cue interpretations and responses can highlight critical cues and responses when the discriminations to be learned are difficult.
- When feedback is used primarily in a guiding role, it should occur as soon as practical after the action, and the student should repeat the action without undue delay.
- Guidance should provide support for engaging in metacognitive activities, directly discussing, and modeling cognitive strategies to be applied.
- Use of scaffolding in providing guidance for complex cognitive skill instruction aids the learner in developing knowledge and skills beyond independent current levels. *Note:* Simply telling the learner to ''do this'' can short-circuit an ill-structured problem-solving task in simulation to making it just a lower-level procedure. Guidance should rather be ''think about this'' so learning at the higher level can be accomplished.

Sequence. The contribution of high engagement simulation or gaming strategies depends on the sequence of integration with classroom instruction and real-world exercises. The instructional sequence decisions for inclusion of simulation or gaming are guided by the following considerations.

Simulation or gaming can provide concrete meaning for theoretical knowledge. Classroom instruction can sometimes be more meaningful if it follows training device practice. Transfer of training to real-world tasks often requires practice in the real-world situation soon after simulation. Some academic and training device instruction will be more meaningful if it is preceded by exposure to the tasks in the whole, real-world situation. For example, learning to fly an airplane can be more meaningful if an orientation ride is included early in the plan of instruction.

The following principles apply to sequencing high engagement strategies relative to classroom and real-world instruction:

- Simulation and real-world instruction should be sequenced as needed to maximize the contributions of both.
- The experience level of a student should help determine the length of delays among related academic, simulation, and real-world instructional experiences.
- Simulation practice should be designed and scheduled to provide experience in the use of knowledge and concepts previously learned at the verbal level in academic training.
- When students have not had experiences needed for academic concepts to be understood, simulation or gaming experience of an appropriate kind should precede or be concurrent with academic training in those concepts. *Note:* Sequence for discussion of the cognitive strategy used in a metaskill should be initiated as part of the briefing before the simulation and then followed up during debriefing with the learner reflecting on the correct answer.

Allocation. The instructional capability of high engagement simulation or gaming varies by device. A plan for how to use a specific instructional strategy should reflect these training capabilities.

The following principles apply to the allocation of training among instructional approaches for high engagement simulation and gaming:

- For any task to be trained in a simulation, key cues and responses related to the task should be represented meaningfully.
- If precise visual-motor skills are to be learned, precise performance must be practiced.
- The experience levels of students should be considered when allocating instruction to simulation or gaming with differing training capabilities.
- Instruction should be allocated to individual simulation or gaming devices according to their effectiveness and efficiency in training portions of tasks.
- Instruction can be allocated to a simulation device with characteristics different from given target equipment or situation, provided that the skills being practiced are not specific to that equipment or situation and the cognitive fidelity is maintained.
- When two or more simulations or games are equally useful for instructing a particular set of skills, allocation of instruction to a simulation device should consider relative cost, overall program efficiency, and preferences of instructors and students.

Organizing Practice. An important consideration in the design and conduct of training using high engagement simulation or gaming is organizing tasks to be learned into practice sessions. Generally, cue and response complexity should be reduced in early stages of skill development, managing cognitive load. The degree of cognitive loading on the student needs to be adjusted so that learning is facilitated, not hindered. For example, a new pilot put into the B-52 WST for a full crew mission simulation without first being fully trained in his individual role quickly became overwhelmed. Skills should eventually be practiced in situations representing the full complexity of operational performance.

The following principles apply to organizing tasks for practice in high engagement simulation and gaming sessions:

- Tasks should be separated, or grouped, for practice to reduce cue and response complexity (simplify) early in training, then provide opportunities to put the ''simple'' in context of the more complex ''big picture.''
- Complex tasks should be separated into parts, or subtasks, if it makes the parts easier to learn.
- Tasks should be separated from each other during early learning if separation aids cue discrimination.
- Tasks separated for practice should be divided so as to maintain integrity of the parts and to maintain cognitive fidelity for the learner.
- When tasks normally occur together or in a contiguous sequence, but are practiced separately, some cues and responses related to the omitted tasks should appear at appropriate times during practice.

Duration and Frequency. Practice schedules depend on a number of factors that vary with the tasks being trained, when they occur during the training program, and the relative skill level of the learner. The primary factors for consideration are as follows:

- Effectives of interference;
- Level of previous learning; and
- Amount forgotten during training.

The following principles apply to the duration and frequency of practice in high engagement simulation and gaming:

- Practice sessions should be of a duration and frequency that permit steady progress during learning.
- Schedules for practice should prevent intertask interference as much as possible.
- Practice during task acquisition should not continue to the point that a student becomes unduly bored or fatigued.

- As skill mastery progresses, practice sessions can be both longer and more frequent.
- Practice sessions should be frequent enough to prevent unacceptable deterioration of skills.
- Practice sessions spaced out over time are better than massed practice.
- Build in time for feedback and reflection during and after practice.

NOVICE TO EXPERT

Understanding how to design advanced high engagement strategies for simulation and games requires an understanding of the characteristics of expert performers. Understanding the expert performer begins with an understanding of how the expert thinks. Experts use pattern recognition, an ability to encode entire cue patterns as opposed to a conglomerate of individual cues (Cannon-Bowers, 1997). One of the main differences between an expert and a novice for performing a skill is the greater scope of the expert's generalization system for the skill and the greater number and detail of mastered discriminations within this system. Experts also have the ability to make situation assessment, relying on memory templates and stored mental models. They have critical thinking skills using metacognition and mental simulation, metacognition being the ability to monitor the use of cognitive strategies. Associated concepts of learning that apply to expertise development are transfer, learning hierarchy, and encoding (Ericcson, Charness, Hoffman, & Feltovich, 2006).

Transfer

The continuity of behavior is attributable to transfer as a process of building new learning on old learning. Experience is a complex matrix of discriminations that is generalized selectively to each new experience. Everything from simple motor skills to profound understandings depends on what experiences have gone on before and how they are brought to bear on the present. This is transfer as a process of learning.

The idea underlying almost all of the preceding discussions is that one learns something new by building on what is already known. As knowledge increases, there is an ever-expanding generalization, and discrimination, of a complex knowledge base integration. At the same time, new learning experiences add to and are integrated with the knowledge base and can be built upon in the future.

At any point in the use of high engagement strategies, there should be deliberate provisions for reflection such as drawing forth from learners whatever is in their background that can be effectively and efficiently brought to bear on training issues at hand. Training design for given objectives should anticipate how what is learned in one set of exercises can be incorporated into the knowledge base in such a way that it can be further built upon in later training.

Learning Hierarchy

Another way of depicting learning using high engagement strategies is the learning hierarchy for organization of skills and related elements that shows when transfer would occur during learning, building learning on previous learning. At the bottom of the hierarchy are those skills and elements that should be learned first, to include the cognitive strategies. Their priority is due simply to the fact that, once they have been learned sufficiently, they facilitate learning of skills and elements at the next hierarchical level. In turn, what is learned at the second level facilitates acquisition at the third level, and so on. To take advantage of this process, it is necessary to know what learning facilitates other learning and to conceptualize the hierarchy accordingly.

It is important to recognize that a learning hierarchy is not the same as what is called a ''skill hierarchy.'' The skill hierarchy usually refers to an organization of skills and elements defined by their relation to job requirements. Such an organization would rarely conform to expert schemas for skill performance. The design sequence begins with the big picture showing integrated metaskills and transformed tacit knowledge. This big picture view establishes the ''hat rack'' upon which the follow-on lower-level learning activities can be connected. These lower-level activities are defined by task complexity with management of cognitive load on the learner. Thus, from the beginning, the learner is developing the knowledge schema for transformation to tacit knowledge and the procedures with pattern recognition needed for metaskill performance. A common error is to put all the procedures, facts, and concept instruction early in the hierarchy, leaving integrated metaskills for later. This separation can cause an increase in time to proficiency and may even result in negative training.

Encoding

Development of expert schemas has a beginning with encoding. Encoding is the process of taking the incoming ''message'' and putting it into meaningful ''code'' or units. Generally, there are four encoding steps that comprise what is often referred to as ''pattern matching.'' When a performer or student approaches a task, she:

1. Selects factors (e.g., stimuli) for attention.
2. Transforms these factors into meaningful representations (e.g., cue recognition).

3. Abstracts transformations into patterns, sorting out what goes with what in what order.
4. Spells out the resulting patterns as needed to understand the task.

The novice learner typically falls short on all four of the encoding processes. After a skill is learned, these four processes become more or less automatic. Expert learners are more likely to pick up all relevant stimuli (Step 1), have a greater variety of useful transformations available (Step 2), have better habits of abstracting key features of patterns (Step 3), and be more versatile in obtaining the final pattern match (Step 4). Experts are also more likely to recognize failure to obtain an adequate match.

A major goal of instructional design for high engagement strategies is to provide the knowledge and training of processing skills that make these four steps automatic and comprehensive. Instruction starts with building on the prior experiences of the novice learner and provides a framework on which new learning can be organized into meaningful units. The learning strategies are explicitly designed to help the learner accomplish the four encoding processes, not expecting them to happen automatically. For example, the novice driver education student is taught how to drive on a curved roadway simulation by (Step 1) looking at cues down road rather than following the line in the middle of the road off the front of the vehicle. The student is taught to (Step 2) anticipate the curve and (Step 3) read cues that indicate the degree of turn and the safe speed for operating the vehicle around the curve. The student is taught (Step 4) to match the steering wheel turn with the right amount of pressure on the gas pedal in a pattern for safely negotiating the curve.

Analysis of Expertise

Develop an in-depth understanding of the content and support requirements of an expert's tacit knowledge. This involves at least the following tasks:

1. Describe the tacit knowledge that can be identified and articulated in a way that allows us to manage and train it in a simulated environment.
   Note: Cognitive task analysis provides a method for ferreting out tacit knowledge from the expert performer. Tacit knowledge is generally not readily recognizable by the expert performer because it has been transformed into its current implementation. The skilled surgeon has difficulty articulating the basic principles learned in medical school that have become the foundation for performance, yet cognitive task analysis can reestablish the connection to show what began as discrete elements is now transformed into a more meaningful whole for the surgical application.
2. Define the ''raw material'' of tacit knowledge, even though some tacit knowledge may resist being translated into words or rules.

Note: Although this latter goal may seem illusory, it should be possible to observe and measure enhanced levels of higher-order performance, for example, problem solving, when a simulation provides the raw material needed to support underlying tacit mechanisms. This should be possible, even though we do not understand all of the inner workings of a given tacit process.

3. Define the metaskills of the expert performer and the process of metaskill development. This requires an understanding of the mental processes that the expert performer uses to organize and/or reorganize information when problem solving, decision making, developing a sense of situational awareness, experiencing a sense of knowing, and so forth.
4. Given this information, formulate a knowledge management capability that identifies and structures the raw material needed for development of expertise, the tacit knowledge and metaskills of the expert, such as the ''top gun'' combat pilot.
5. Put it all together in a cognitive ''whole'' to help describe the mechanics and support requirements of a cognitive reality.
6. Build a blueprint that shows how ideas and strategies begin with schema development that find their source in feelings, attitudes, beliefs, and perhaps even inherited traits. This blueprint will require an understanding of:
   a. What knowledge, including clues and rules, has the potential of being transformed into tacit knowledge.
   b. How schema and models transformed in tacit knowledge are processed and combined so as to provide solutions to new and novel problems.

Corollary in Pilot Training

Development of expertise takes into account normal learner progression. Devol (1998) used Maslow's hierarchy of needs to explain the normal progression of a pilot through warfighter training, as shown in Figure 12.5. A pilot's lower level needs must be met before progression to the satisfaction of higher level needs. The early stages of self-preservation are to meet physiological and safety needs. Life support systems that work and safe flight experience contribute to progress. Flying as a two-element wingman and then in four-ship formation are initially self-centered activities.

Next the pilot comes to feel a part of a flight, experiencing success in team interactions and growing in self-esteem to meet social and ego needs. The pilot gains a sense of and a willingness to give and receive. Success is experienced in the ranges of team performance required for the flight's mission

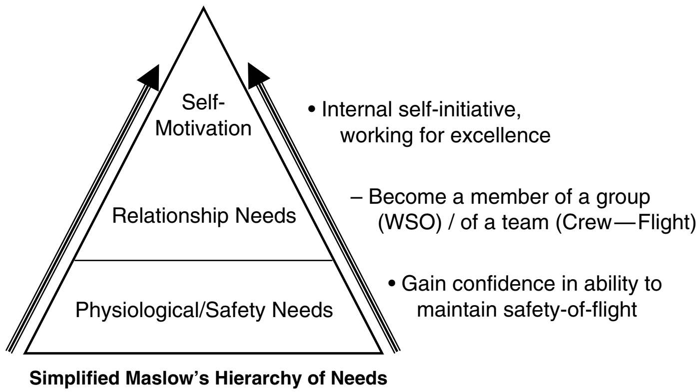


Figure 12.5 Normal Progression of Pilot Development Through Warfighter Training.

accomplishment. The pilot's self-confidence and a feeling of independence grow with these achievements. Status and recognition come as progression goes from two-ship to four-ship lead and then to instructor. Personal reputation leads to selection for weapons school or Red Flag exercises.

As the pilot advances, self-potential is realized, continued self-development is desired, and self-fulfillment needs are met. The pilot emerges as an expert and is called upon to participate in tactics reviews, check ride evaluations, and select leadership opportunities.

INSTRUCTIONAL STRATEGIES

In order for high engagement strategies to be applied in virtual, constructive, and live simulation, serious games, augmented reality, or Second Life, to teach, then the right conditions for learning must be applied. The acquisition of different types of knowledge and skill requires different conditions for learning (Gagné, 1985). If any instructional experience or environment does not include the instructional strategies required for the acquisition of the desired knowledge or skill, then effective, efficient, and apparent learning of the desired outcome will not occur (Merrill, Drake, Lacy, Pratt, & ID2 Research Group, 1996). Merrill provides the following summary:

- There are different kinds of knowledge and skill (Gagné assumption).
- The different kinds of knowledge and skill each require different conditions (strategies) for learning.
- IF an instructional strategy does not include the first principles of instruction, THEN it will not teach.

The conclusion is this: ***If a simulation or game does not teach, it has no value.***

Another way of stating this conclusion is this: *If the simulation or game is not teaching the right cognitive and behavioral skills using the right instructional strategy, then it is only a game*. Simulations or games can lead to a great deal of activity but little learning. A game or simulation can bring about cognitive overloading in ways that are counterproductive to learning. With the wrong simulation or game, the learner's development toward expertise can be inhibited, negative training can ensue, and severe consequences may occur.

SUMMARY

Advances in simulation and gaming opened a new era for instruction, particularly advanced graduate-level training for developing expertise. As high engagement strategies are implemented, learners are scheduled for operations or mission training in simulation much the same way they are scheduled for training in the operational vehicle or organizational environment. The realization of effective high engagement strategies is dependent on the proper instructional system design using the first principles of instruction and the systematic sequencing of instructional media in a building-block approach following principles that enhance learning. Simulation and serious games are selected as instructional media when they are less costly, less dangerous; when they are more convenient, more flexible; when they focus on important aspects of a training problem; and when they are designed to support and facilitate learning. Simulation and serious games can also be used in solving design and research problems.

Instruction that leads to development of expertise is above and beyond basic vehicle and mission training that only achieves operation-ready status. Instruction for expertise can incorporate simulation and gaming for developing metaskills and providing aging and experiencing that bring about the transformation to tacit knowledge. Through this type of advanced instruction, a learner builds expertise as part of a team, capable of facing novel situations and responding appropriately for successful mission accomplishment.

Application of instructional system development (ISD) to simulation and games will guide development of high engagement strategies if the focus is on the development of tacit knowledge and integrated metaskills, and also capitalizes on the instructional capabilities for evolving novice learners into experts. This focus is accomplished by systematically applying principles for enhancing learning. Learners are advanced to more complex, integrated learning activities in simulation and gaming so that a transformation to tacit knowledge occurs and higher-order metaskills are developed. Like any new instructional technology, if the right instructional strategies are not employed to bring about learning, then it has no value.

In order for high engagement strategies for simulation and games to provide instruction, the principles that promote learning must be incorporated. When positive transfer of simulation training corresponds to successful performance in the operational vehicle or organizational environment, then simulation effectiveness will be demonstrated. Integration of new instructional technology requires a total instructional or training systems approach. This total context or ''big picture'' provides insight into where the high engagement simulation or game should be used. When the learner is provided this big picture at the beginning of the instructional experience, then this context facilitates long-term schema development. Here the context of the real-world, operational application is cognitively connected to the patterns of instructional events much like a hat rack connects each different hat to the center pole.

In order to achieve integration of simulation and games into the instructional or training system, the instruction is designed to prepare the learner for each scenario and then provide feedback following the simulation or game experience. The scenarios should be created to manage cognitive load in a simple to complex developmental sequence. Scenario scripts can be designed by subject-matter experts to include all elements necessary to achieve high engagement strategies complementary to the operational world. The scenarios should be set in typical real-world situations to which learners may be assigned. Mission-orientated testing uses these scenario scripts for the validation method to demonstrate the simulation or game instructional capability.

Assessment of learner performance should indicate degree of progression to metaskill mastery and transformation to tacit knowledge. Grade sheets can be based on how much involvement is required by the instructor during task performance by the learner. Predetermined performance capabilities such as mission-essential competencies should incorporate cognitive components.

The functions covered by the total instructional or training system design should trace back to the instructional, support, and operational requirements. The design goal is achievement of instructional outcomes, rather than engineering a high engagement simulation or game. The more typical focus of an engineering design has been on the simulation of the physical reality that supports cognitive processes. In contrast, the cognitive reality design approach articulates the tacit knowledge and metaskills required for operational success and then builds the physical implementation to achieve the instructional capabilities. In cognitive reality, the high-end instructional strategy for the simulation or game leads to performance of the desired instructional outcomes. In order to understand the nature of cognitive reality, simulation or game design must include some form of cognitive analysis. This level of design leading to advanced instruction will require a multi-disciplinary team approach. High engagement strategies give new meaning to instructional capability for carrying out advanced instruction. High engagement strategies have demonstrated the capacity for taking individual and team skills to levels never before thought possible.

References

Augmented Reality. Retrieved March 12, 2008, from http://www.augmented-reality.org/

Bills, C. G. (1987). Integration of weapon system trainer (WST) for large body aircraft into combat crew training. *Proceedings of National Aerospace Electronics Conference (NAECON)*, pp. 1054–1059. Dayton, Ohio.

Bills, C. G. (1993). Information for designers of instructional systems: Design guide for device-based aircrew training. Unpublished manuscript.

Bills, C. G. (1997). Effects of structure and interactivity on internet-based instruction. Doctoral Dissertation. Kent, OH: Kent State University.

Bills, C. G. (1998). Performance measurement system—It's time has come. *Proceedings of the Interservice/Industry Training Simulation and Education Conference (I/ITSEC)*, Orlando, Florida.

Bills, C. G. (2009). Diagram of developmental learning from novice to expert. Unpublished manuscript.

Bills, C. G., & Burkley, J. T. (2002). F-16 MTC, from design to full warfighting capability. *Proceedings of the NATO RTO Studies, Analysis and Simulation Symposium*, Brussels, Belgium.

Bills, C. G., & Devol, D. (2003). How F-16 MTC team achieved operational training capability. *Proceedings of the Interservice/Industry Training Simulation and Education Conference (I/ITSEC)*, Orlando, Florida.

Bills, C. G., & Wood, M. E. (1999, December). New insight for training development of 21st century advanced warfighter training. *Proceedings of the Interservice/Industry Training Simulation and Education Conference (I/ITSEC)*, Orlando, Florida.

Brock, T. R. (2007). Training NASA astronauts for deep space exploration missions: A research study to develop and validate a competency-based training framework. Doctoral Dissertation, Cappella University.

Cannon-Bowers, J. A. (1999, February). Advanced learning principles. Paper presented at the meeting of the Joint Strike Fighter Training Integrated Product Team (TIPT), Orlando, Florida.

Caro, P. W. (1978). Some current problems in simulator design testing and use. Human Resources Research Organization paper presented at Military Operations Research Symposium, Fort Eustis, Virginia.

Carolan, T., MacMillan, J., Entin, E. B., Morley, R. M., Schreiber, B. T., Portrey, A., Denning, T., & Bennett, W., Jr. (2003). Integrated performance measurement and assessment in distributed mission operations environments: Relating measures to competencies. *Proceedings of Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*, Orlando, Florida.

Clark, R. C., & Mayer, R. E. (2008). *e-Learning and the science of instruction: Proven guidelines for consumers and designer of multimedia learning* (2nd ed.). San Francisco: Pfeiffer.

Colegrove, C. M., & Bennett, W., Jr. (2006, December). *Competency-based training: Adapting to warfighter needs* (AFRL-HE-AZ-TR-2006-0014). Williams AFB, AZ: Air Force Human Resources Laboratory.

Devol, D. (1998). Using Maslow's hierarchy of needs to explain pilot warfighter development. Unpublished manuscript.

Ericcson, K. A., Charness, N., Hoffman, R. R., & Feltovich, P. J. (2006). *The Cambridge handbook of expertise and expert performance*. New York: Cambridge University Press.

Fishburne, R. P., Jr., Spears, W. D., & Williams, K. R. (1987). Design specification development for the C-130 model aircrew training system: Final report. Irving, Texas: Seville Training Systems.

Fried, M. P., Sadoughi, B., & Bills, C. G. (2006). Surgical educators incorporating surgical simulation take lessons from flight training. *Proceedings of the Interservice/Industry Training Simulation and Education Conference (I/ITSEC)*, Orlando. Florida.

Gagné, R. M. (1985). *The conditions of learning and theory of instruction* (4th ed.). New York: Holt, Rinehart and Winston.

Gagné, R. M., & Merrill, M. D. (1990). Integrative goals for instructional design. *Educational Technology*, *38*(1), 22–30.

Kaufman, R. A. (1997). Avoiding the ''dumbing down'' of human performance improvement. *Performance Improvement*, *36*(5), 22–24.

Kaufman, R. A., & Keller, J. (1994). Levels of evaluation: Beyond Kirkpatrick. *Human Resource Development Quarterly*, *5*(4), 371–380.

Kirkpatrick, D. L. (1959). Techniques for evaluating training programs. *Journal of the American Society for Training and Development*, *13*, 3–32.

Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research and Development*, *50*(3), 43–59.

Merrill, M. D. (2007, October 12). First principles summary.doc. Retrieved February 27, 2008, from http://cito.byuh.edu/merrill/text/papers/FirstPrinciplesSummary.pdf.

Merrill, M. D. (2008). First principles of instruction. Retrieved February 27, 2008, from http://cito.byuh.edu/merrill/text/papers.htm.

Merrill, M. D., Drake, L. D., Lacy, M. J., Pratt, J. A., & the ID2 Research Group. (1996). Reclaiming instructional design. *Educational Technology*, *36*(5), 5–7.

Morrish, T., Stoner, P., & Gurcak, A. (1997). Virtual environment deployable simulation (VEDS). *Proceedings of the Interservice/Industry Training Simulation and Education Conference (I/ITSEC)*, Orlando, Florida.

Nullmeyer, R. T., & Rockway, M. R. (1991). Integrated aircrew training management information systems: An organizational perspective. *Proceedings of the Interservice/ Industry Training Systems Conference (I/ITSC)*, 416–426, Orlando, Florida.

Nullmeyer, R. T., & Rockway, M. R. (1984). Effectiveness of the C-130 weapon system trainer for tactical aircrew training. *Proceedings of Interservice/Industry Training Equipment Conference*, Washington, D.C.

Polanyi, M., & Prosch, H. (1975). *Meaning*. Chicago: University of Chicago Press.

Roberson, A. J., & Bills, C. G. (2006). Virtual William Tell competition: The F-16 MTC Turkey Shoot. Unpublished manuscript.

Rossett, A. (2004, October). Simulations and games: Revisiting their strategic value. *The eLearning Developers Journal*, pp. 1–6.

Second Life. Retrieved February 26, 2008, from http://secondlife.com/

Serious Games[1]. Retrieved February 26, 2008, from http://seriousgames.ning.com/

Serious Games[2]. Retrieved February 26, 2008, from http://seriousgames.org/index2 .html.

Spears, W. D. (1983). *Processes of skill performance: A foundation for the design and use of training equipment* (NAVTRAEQ-VIPCEN78-C-0113-4). Orlando, Florida: Naval Training Equipment Center.

Spears, W. D. (1985). *Plans for C-130 model aircrew training system (MATS) evaluation and associated research and development*. Williams AFB, AZ: Air Force Human Resources Laboratory.

Tennyson, R. D., & Michaels, M. (1991). *Foundations of educational technology: Past, present and future*. Englewood Cliffs, New Jersey: Educational Technology Publications.

Spears, W. D. (1992). Learning foundations of training design. Unpublished manuscript.

U.S. Department of the Air Force. (1993). *Information for designers of instructional systems: Design guide for device-based aircrew training* (AF HAND 36-2235 Vol. 7). Washington, DC: U.S. Department of the Air Force.

CHAPTER THIRTEEN

Video Game–Based Learning

An Emerging Paradigm for Instruction

Kurt D. Squire

Interactive digital media, or video games, are a powerful new medium. They offer immersive experiences in which players solve problems. Players learn more than just facts—ways of seeing and understanding problems so that they ''become'' different kinds of people. ''Serious games'' coming from business strategy, advergaming, and entertainment gaming embody these features and point to a future paradigm for e-learning. Building on interviews with leading designers of serious games, this article presents case studies of three organizations building serious games, coming from different perspectives but arriving at similar conclusions.

This chapter argues that such games challenge us to rethink the role of information, tools, and aesthetics in a digital age.

Over the past decade, e-learning has been a dominant paradigm for the electronic development, management, and distribution of learning materials. But as many critics have noted, most e-learning is nothing more than online lectures or course notes, and the basic organizing metaphors of traditional classroom learning—knowledge as discrete and abstract facts, learning as ''acquisition'' of content, and therefore instruction as the organization, dissemination, and management of that content—have gone unchanged (cf. Bednar, Cunningham, Duffy, & Perry, 1992; Fodor, 2000; Sfard, 1998). The promise of

Excerpted from Kurt D. Squire, *Performance Improvement Quarterly, 21*(2), 2009. Reprinted with permission of John Wiley & Sons, Inc.

e-learning, to make customized, accessible learning experiences, has given way to more mundane pursuits such as free online content. In the words of Cross and Hamilton (2002):

> ''Corporate e-learning is a powerful paradigm, but it has strayed from its inspired beginnings. Poised to become a driver of business performance, e-learning lost its way as vendors reached for quick economic gains at the expense of long-term strategic position. e-Learning devolved into quick-to-sell IT-only content libraries, bland web course designs, and unfocused, minimally tailored portal solutions. This was a boon to the training department, but not the business as a whole, and the value of hassle-free turnkey campuses and trainer-empowering LMSs [learning management systems] became the low hanging fruit in the marketplace.'' (p. 1)

In short, many e-learning leaders recognize that publishing content online is not synonymous with improving learning or performance. In fact, so-called content (that is, declarative knowledge in the form of information bits or facts) is, and always has been, ''cheap''; even before the Internet, one need only go to the public library for access to the world's information. What has been more difficult is effective design of instruction to deliver the kind of social and material experiences necessary to make sense of that content, to make it meaningful and useful for future action.

As traditional e-learning stands in flux, a new paradigm of digitally mediated learning, commonly called game-based learning, is emerging (Aldrich, 2004; Prensky, 2001; Squire, 2003). Recently, research projects, organizations, centers, grants, books, and studies have emerged exploring new visions for game-based technologies in learning (cf. Games-to-Teach Team, 2003; Media-X, 2003; Sawyer, 2002; Shaffer, Squire, Halverson, & Gee, 2004). Driving this change are several factors, among them (1) recognition that games are a multibillion-dollar industry, rivaling Hollywood in revenues and cultural influence[1] (ESA, 2004); (2) digital games being one of the only (other than pornography) unambiguously profitable uses of the Internet (Kolbert, 2001); (3) digital games being routinely listed as the most ''important'' and influential medium by college students (Games-to-Teach Team, 2003); and (4) games being a powerful socializing force, such that those who play computer and video games have different attitudes than their peers about work, play, and their coworkers (Beck & Wade, 2004).

Underlying this move toward game-based learning environments is more than strategic opportunity or marketing; the shift toward games also entails intellectual recognition among many that they represent experiential learning spaces, spaces where learners have rich, embodied, collaborative, and cooperative interactions during which they think with complex tools and resources in the service of complex problem solving (Gee, 2003; Squire, 2003). Gee argues

(2004) that, as games become more complex, they have begun using intelligent tutors, scaffolding, and affinity groups for learning to help players understand their increasingly sophisticated interfaces and systems. Rapid iterations in a highly competitive market have resulted in highly evolved interfaces and learning systems designed to teach players to play them. In short, many game designers have developed an expertise in (some) fundamental principles of instructional design, in particular the idea of experience design, which Wilson and Myers (2000) and others have argued is fundamental to situated views of cognition.

As a result, game-based training has gone from a relative niche market to a roughly $30 to $75 million market (Erwin, 2004). The games industry is transitioning into big business, and many small developers are facing difficult financial times (Erwin, 2004). As this study shows, a number of developers such as BreakAway Games are taking the interactive design expertise honed in the games industry and applying it to advertising, training, and marketing. The military in particular is hiring game designers for their knowledge of how to create compelling user experiences that can be the basis for changing understandings, behavior, beliefs, and even identities (Swartout, 2004). As these game players, designers, and even entire companies migrate into the training space, traditional e-learning developers may have to rethink how they conceptualize their practice. This case study investigates three questions:

Games [are] a powerful socializing force. Those who play computer and video games have different attitudes than their peers about work, play, and co-workers.

1. What new models of learning and training are emerging?
2. What kinds of institutional changes are accompanying this change?
3. What implications does game-based learning present for instructional designers and performance technologists?

Drawing on a critical review of existing game-based learning literature, a content analysis and review of game-based learning products, and interviews with game-based learning designers, it shows how digital and video games are emerging as a new model for situated learning environments.

I argue that games problematize contemporary work in e-learning, which focuses on supplying, organizing, and repackaging content, offering new models that put primacy on experience.[2] Game-based learning can be understood as a particular kind of designed experience, where players participate in ideological worlds, worlds designed to support a particular kind of reaction, feelings, emotions, and at times thoughts and identities, which game-based learning designers are leveraging for education and training.

METHODOLOGY: COMPARATIVE CASE STUDY

This study uses comparative case study techniques to build a framework for game-based learning. Consistent with Stake's case methodology (1995), it employs a combination of historical research methods, document analysis, interviews with trainers and game developers, and critical study of game artifacts to theorize contemporary serious games as an emerging model of e-learning. The cases were reported elsewhere in greater detail (cf. Squire, 2005). This study analyzes the cases for emergent themes toward understanding serious games as a model for situated learning and their implications for instructional and performance technologists.

PRELIMINARY REVIEW

The initial review of existing work examined the Serious Games archive; the emerging literature in games studies on e-learning, advergaming, and ubiquitous gaming; and the situated learning literature. Several successful programs were identified and contacted for further exploratory study. The researchers conducted eighteen informal interviews with representatives from ten learning organizations, ranging from small independent contractors to Fortune 500 companies (gamelab, Root Learning, Digital Mill, E.I. Lilly, BreakAway Games, YaYa Media, SimuLearn, SimQuest, DESQ, and the U.S. Army).

Data Sources

From these initial interviews, three game-based learning companies (YaYa Media, BreakAway Games, and Root Learning) doing game-based learning work were selected (see Table 13.1). They were selected for their relation to four themes that emerged from preliminary interviews and analysis: (1) games as spaces for experiential learning, (2) games as a context for discussion,

Table 13.1 Overview of Case Contexts

Organization	*Background*	*Size*	*Offices*
BreakAway Games	Entertainment games, military consulting	100	Baltimore
Roots Learning	Business strategy, consulting	75	Toledo, Chicago, London, Zurich
YaYa Media	Business strategy, marketing, and advertising	50	Los Angeles, New York

(3) games as tools to think with, and (4) games as a space for exploring new identities. Each theme has also emerged within the research literature on games and education and therefore was worth examining further[2] (cf. Gredler, 1996; Squire, 2003). These cases are not necessarily representative of the field as a whole; rather, they were selected as purposive samples to probe theoretical issues in game-based learning and generate a productive framework for game-based learning.

For each case, we interviewed company CEOs, developers, and trainers, reviewed games and other materials, and interviewed clients and vendors to triangulate data. On the basis of these data, researchers generated profiles of each company. It is worth noting that none of the featured companies started in instructional design, technology, or e-learning; they come from business strategy, marketing, and the games industry.

Data Analysis

For each case, the author compiled all notes, which numbered approximately thirty pages each. These were condensed into vignettes, which are furnished here for context. Next, a team of three researchers examined the notes and began coding interactions for themes, using a database application.

Multiple passes were made at the data in an attempt to generate assertions that were concise, not overlapping, and as strong as possible while still being ''true'' to the data. The researchers then outlined four conclusions, which are detailed in the Implications section of this article. These conclusions or assertions draw from all three cases and attempt to synthesize the findings so as to build a more general theory of game-based learning. Researchers shared these findings in draft form with participants, in order to understand whether they accurately reflected participants' beliefs. After sharing the results, the researchers made another round of edits to clarify misleading assertions and ensure the accuracy of participants' comments.

Limitations of Study

Underlying this approach is an interpretivest epistemology, a way of knowing that assumes the researcher brings his or her own questions, values, and assumptions to the study and is thereby an integral part of the research study. Whereas most traditional case study research has relied on an objectivist epistemological stance, whereby cases or other sets of qualitative data are used to develop ''grounded'' findings, the interpretivist tradition makes no claims that the assertions are inherently ''in'' the data, but rather that the findings are co-constructed among the interpreter, the phenomena at hand, and the reader (Cresswell, 2003; Glaser & Strauss, 1967; Merriam, 1998; Stake, 1995). From the interpretivist perspective, it is assumed that no findings are (or could) be made outside of the theoretical traditions and language in which they are situated. Further, in contrast to traditional grounded theory approaches, it assumes that in many instances better progress can be made by targeting research to extend particular theoretical notions than by relying exclusively on data to guide the research. Thus the reader might regard these findings and assertions as interpretive—as a design theory of what directions the field may go in (Reigeluth & Frick, 1999).

RESULTS: CASES OF GAME-BASED LEARNING

BreakAway Games

BreakAway Games is one of the many games companies that were spun off of the legendary Baltimore, Maryland–based game developer MicroProse after its breakup. BreakAway's positioning near Washington, D.C., allowed it to develop learning simulations, particularly war games and support tools, in addition to its traditional game lineup. Most of BreakAway's early experience was with 2D war games, including Peloponnesian War (a game still used in the army college to teach about ancient warfare), and BreakAway continued making entertainment games, including Waterloo; Austerlitz; and expansion packs for the Tropico, Cleopatra, and Civilization series. BreakAway soon found that their expertise in creating emotionally compelling media that offers particular kinds of experiences was directly applicable to other endeavors, including marketing and training with the U.S. military and defense contractors.

Currently, BreakAway is developing a number of proprietary systems and technologies such as Entropy Based Warfare (campaign analysis and war game assessments) and Integrated Gaming System (supporting war gaming). In addition, they have a number of trademarked technologies for 3D terrain generation, multi-user support tools, and simulation tools. Ironically, they are also preparing to launch a game for A Force More Powerful, a group dedicated to using nonviolent conflict to achieve democracy and human rights (Figure 13.1). One does not normally think of nonviolent peace activists as funding million-dollar games, but the challenges behind training such activists—for example, that it must enable a globally distributed workforce to espouse a particular ideology for solving problems—means that game-based solutions are especially attractive.[3] Homeland security is an arena with similar constraints, and BreakAway, like many companies, also has a game-based solution for training emergency responders.

Root Learning

Root Learning is a ''strategic learning company'' with roots in business strategy. Instructively, their mission is not to ''deliver content'' or ''train new knowledge, skills, attitudes, and beliefs,'' but rather to ''engage and connect people to create results in a context that respects their humanity, intelligence, and capacity to

Figure 13.1 BreakAway's A Force More Powerful.

grow.'' Crucial to Root's identity is that the company perceives itself both as scientists and artists, educators, and businesspeople.

Root Learning's core products include their learning maps, planning documents generated by holding strategic discussions with company leaders and participants (see Figure 13.2). In brief, a learning map is a document generated through a type of needs analysis. An interdisciplinary team of technologists, artists, and designers observe and interview participants to generate a metaphor describing the training or strategy problem.

These metaphors help stakeholders understand the problem in broad terms, understanding why training may be needed before ever introducing content. Participants do interact with content via the learning maps by reading index cards of information, discussing problems, and playing mini-games where they consolidate or apply information they have encountered.

Critically, information in learning maps is not the goal of the exercise, but secondary to supporting the particular message or worldview. Interaction among participants is critical to this model, and, in fact, the gameboard might be seen as a framework for facilitating discussion.

Although Root Learning's background is in interactive, participatory learning environments, it is only now branding its approach as a games company. Root's current work in e-learning, which includes simulations such as Blockbuster, draws more obviously from digital gaming metaphors, tropes, and interfaces (see Figure 13.3).

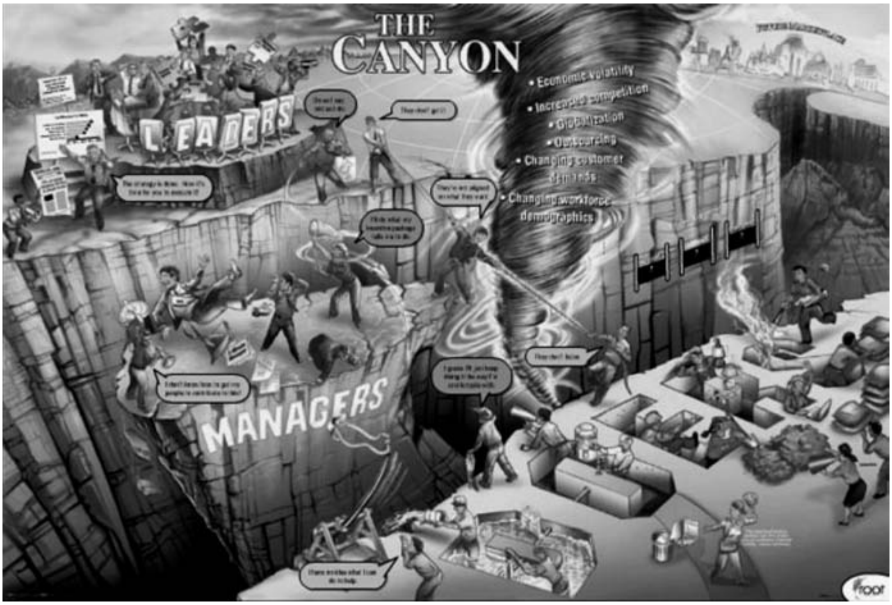


Figure 13.2 Root Learning Map.

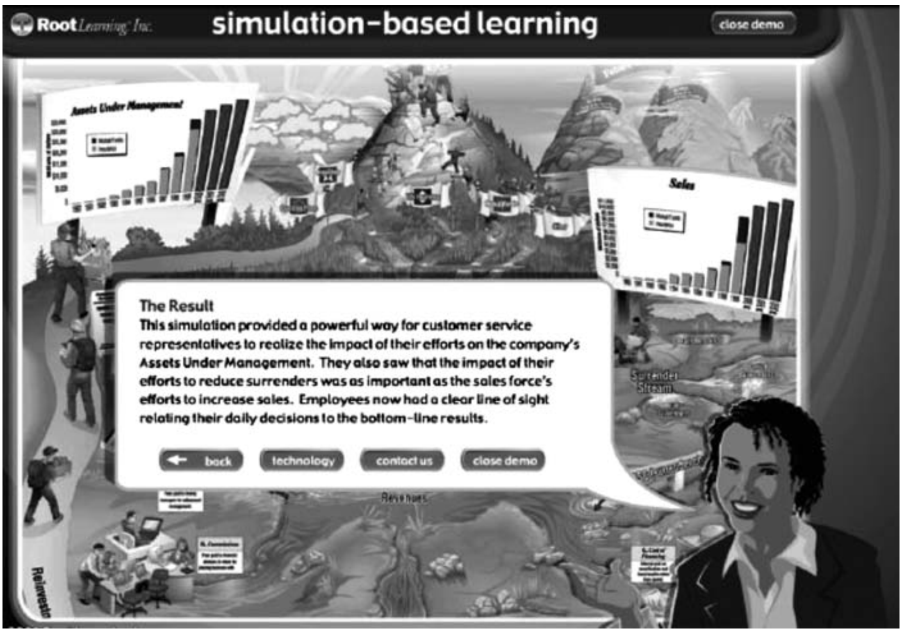


Figure 13.3 Root Learning's Digital Simulation–Game Media.

YaYa Media

YaYa Media's roots are in both the games industry and business strategy. YaYa has carved out a niche in advergaming and is entering training as well (Chang & Pfeffer, 2003).[4] Founded with funding from Michael Milken as a ''leading interactive technology company,'' YaYa's initial business charge was to invent new ways for marketing and advertising a digital economy where consumers' attention is increasingly scarce and technologies such as TiVo threaten the future of traditional broadcast advertising. As such, YaYa is most famous for branding advergaming, a genre of advertising based on gaming principles that simultaneously advertises and gathers marketing data. A nearly game, Chrysler Get Up and Go, typifies the YaYa approach (see Figure 13.4). Users log in to the game, try to match their personality to one another (and a Chrysler vehicle), and win a free vacation to a location best suited to match their personality, all based on a Cosmo magazine-style quiz.

Other games include an accounting game ''Bizzfun,'' and a Jeep driving game for the Chrysler/Jeep sales force (see Figure 13.5).

Originally designed for advertising, these games are now being used for training as well. Critical to understanding YaYa's approach is that their engine not only presents users with content but also collects data on users' choices, preferences, and habits. In such a constrained environment, it is relatively simple to track players' progress and identify patterns (such as which color is most popular with the eighteen-to-twenty-four-year-old age group). Thus, YaYa has found that some of the core questions behind advergaming—how to entice users, present customized content based on players' choices, how to aggregate and respond to this data, and how to encourage customers to build allegiance to

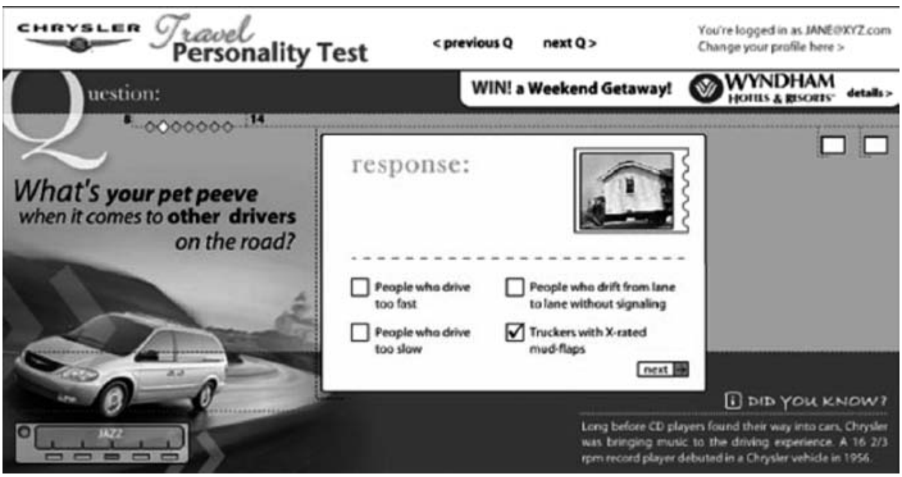


Figure 13.4 Screenshot from YaYa Media's Chrysler Travel Game.

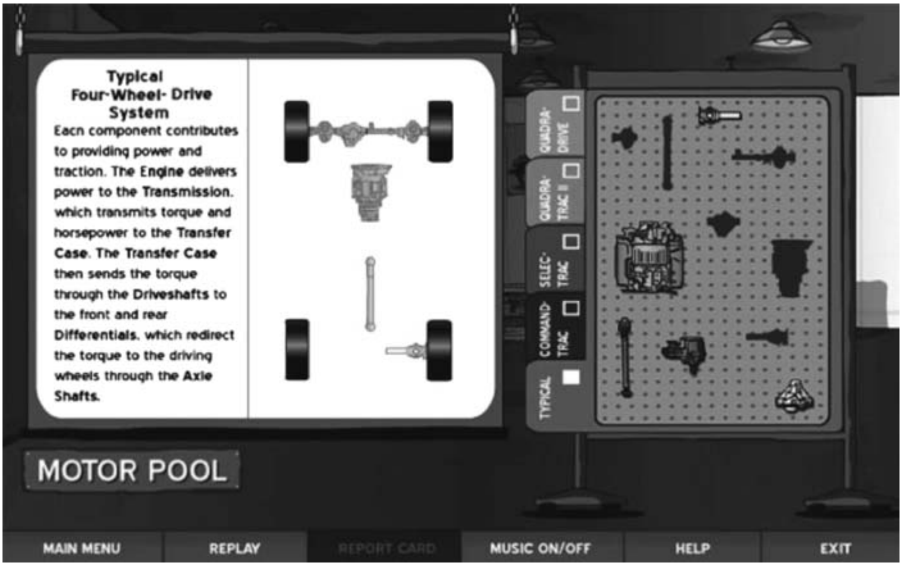


Figure 13.5 Screenshot from YaYa Media's Jeep Driving Training Game.

the brand—are all problems that instructional and performance technologists deal with as well. YaYa's Chrysler game has been used to train sales employees about user preferences, their fashion game is used to change teens' attitudes toward accounting, and their basic game engine has been used in other training scenarios.

FINDINGS: A SITUATED FRAMEWORK FOR UNDERSTANDING GAME-BASED LEARNING

The move toward game-based learning represents more than a shift to a new medium; it is a shift toward a new model of e-learning that focuses lesson content and more on designing experiences to stimulate new ways of thinking, acting, and being in the world. Movements such as the MIT Open Courseware project show that ''good'' online content is cheap; powerful learning experiences are harder to produce. The emergent paradigm of game-based learning is built on a number of principles: (1) create emotionally compelling contexts for learning; (2) situate learners in complex information management and decision-making situations where facts and knowledge are drawn on for the purpose of doing; (3) construct challenges that confront and build on users' preexisting beliefs; (4) construct challenges that lead to productive future understandings; (5) anticipate the users' experiences from moment to moment, providing a range of activities to address learners' needs; (6) invite the learner to participate in constructing the solutions and interpretations; and (7) embrace the ideologically driven nature of education and training.

Instructional development models (such as ADDIE or rapid prototyping) for game-based learning draw from traditional development processes, but also require new techniques. They frequently include (1) managing expectations, (2) offering an early holistic model of the product for clients, (3) iterative design cycles, (4) early user feedback, (5) an increased role for visual designers, (6) business models with blurred lines between marketing and strategy, and (7) distributing instructional design tasks across roles. These findings imply that if instructional and performance technologists adopt ''designing experiences'' as a metaphor for their practice, then they may need to embrace some ''fuzzy'' areas such as aesthetics, which is discussed in the final section of this chapter.

From Content to Context

Participants across all three cases reported that a primary driver of their move to game-based learning was clients' desires for more engaging and immersive experiences. These terms are commonly used in games but rarely discussed in learning. More engaging, immersive e-learning is more than ''fancier window dressing for content''; it is a transformation of assumptions about what it means to think, learn, and teach. Tom Crawford describes Root's interest in gaming:

> ''We're always looking for innovative, fun, engaging pieces, so games are kind of a 'no duh' to move to. We ask, 'How can we get people engaged and get them to learn?' People look at our maps and everyone says, 'It's a gameboard.' So we're giving in to what they're telling us. But the most important piece of our strategy and philosophy of life is really that e-learning has missed the boat. The industry has focused on content, getting out the content, but they leave out the context.''

For Root Learning and their clients, why something matters is much more important than the content itself, which fits with situated and functional views of cognition (cf. Barab, Hay, Barnett, & Squire, 2001; Cognition and Technology Group at Vanderbilt, 1993).

Whereas there is a saying in e-learning that ''content is king,'' a situated view of knowledge would say that it is the context in which learners develop knowledge that is king.

Creating Context

The first thing that games do is create an emotionally compelling context for the player. Many games use cut scenes, short movies designed to situate the player in the game world and context. But there are other, simpler ways. Root's materials build on nostalgia, curiosity, visual appeal, and presumably employees' interest in the bottom lines of their companies. Good games connect with the player emotionally and are an entrée or invitation into the world that is to be learned.

The context creation (much like the problem in problem-based learning) is the bridge from where the player is to where he or she wants to go.

A common misconception about games and simulations is that they are perfect representations of reality. Inherently, they are simplifications of reality (much like any representation, just as a book, picture, or film is also an incomplete representation of reality). Games are spaces in that they are worlds that we participate in the construction of, but they are also built according to particular values, as we saw with Root's learning maps. They call our attention to some aspects of reality while obscuring others. Part of what makes games so powerful as a medium for learning is that they allow us to build worlds that are instantiated according to a particular set of rules.

The Root materials work on several levels to frame the experience according to these rules. First, the ''maps'' draw on board game tropes to immerse the learners in an experience in which they are gathered together around a common task in a setting where informal talk, collaboration, and discussion are encouraged.

But further, they make very powerful use of the core metaphor (going down a road, jumping a chasm) to situate the learner while putting forth an argument for how the particular problem should be viewed. In short, they use graphic artists to create a visual metaphor for engaging with the topic. These metaphors are far from innocuous; they communicate subtly (and not so subtly) what the problem is about. They also set the agenda for the activities to follow.[5]

It's worth noting that Root Learning uses artists—not instructional designers—for task analysis. For Root, it is less important that they create an exhaustive (or even reasonably thorough) statement of the problem. What's more important is that they build a common metaphor for talking about the experience that is understandable to all parties. The assumption is that the core challenge is communicating the proper way of framing the problem, and then particular knowledge, skills, procedures, and beliefs can follow. Thus, games structure experiences around problem solving—problems of the designer's choosing, the player's choosing, or, when games work best, a hybrid of both. A core theoretical and design problem of games is that they are not only designed by the developer but are spaces to be inhabited by players whereby their actions fill out the game world.

INTELLECTUAL AND EMOTIONAL ENGAGEMENT: INVITING PARTICIPATION FROM THE USER

In creating a context for experience, games invite players to inhabit the game space. How different games and game genres work is beyond the scope of this paper and is still being understood by game researchers (see, for example, Games-to-Teach Team, 2003; Gee, 2007; Squire, 2005; Steinkuehler, 2004a, 2004b). But critical to most games is that they both establish challenges and goals for learners to meet (save one's job, rescue the company) and establish seductive identities and capacities for players (high-performing manager, ace delivery driver).

The first thing that games do is create an emotionally compelling context for the player. Many games use cut scenes, short movies designed to situate the player in the game world and context.

In the case of Root Learning, the physical layout visually, metaphorically, and literally invites the learner inside the map to participate in constructing the learning environment (mostly through various related activities, such as matching games). The images are designed to evoke pop culture nostalgia, drawing the user's past identity into the experience. It immediately ties the brand (and learning) not just to an abstract pedagogy but to personally meaningful emotional experiences. Participants are invited to scan different eras and pick out what year the ''Pepsi challenge'' hit, when Michael Jackson's hair caught fire, and so on. As the eye moves down the street (and through time), it encounters shops showing different trends, with shrinking rates of profitability presented in order to elicit concern over participants losing their jobs.

YaYa Media's games operate more like traditional video games, creating alluring roles for players to inhabit. Bizzfun, their accounting game, is designed to show high school students how accounting and communication skills can lead to ''exciting'' careers in business, such as in the fashion industry.

Although Bizzfun may not have been designed to target women specifically, a majority of its users are women. The roles it creates (powerful fashion industry leaders) explicitly ''show'' high school students how the skills they already have will make them successful in any business and speaks (in their own language). Similarly, YaYa's Jeep racing game teaches sales representatives about Jeep options not through PowerPoint but by letting them design and race their own virtual vehicles that include realistic Chrysler parts.

Problem-Driven Activity

In these cases, key factual information—the things that you might find in a PowerPoint—is organized and presented to give players a compelling experience, all of which emphasizes a particular worldview (Pepsi must adjust sales strategies, accounting is fun, Jeep upgrades are good). As opposed to traditional instructional approaches, which typically contain the need for instruction, generalities, examples, definitions, practice, and feedback, game-based approaches are organized around situations, roles, activities, and practices. Although game-based approaches are interactive, co-constructed by users, and experiential, there are still overarching narratives at work. Root Learning executives explain how their approach treats data as subservient to an overarching narrative: ''Most executives feel that everything is important. We ask 'What are the key pieces of data people need to do their jobs differently?' One thing that artists do is filter through and say, 'This is the key piece. This is the lynchpin to the story.'''

Of course, this is not any old story, but the story that Pepsi executives want their employees to believe. Root designers claim that few object to this approach because they always make companies' goals explicit and use the discussion and debriefing times to address the validity of this interpretation explicitly.

BreakAway Games' A Force More Powerful is a clear example of this principle. Players complete a series of missions around nonviolent political action missions, designed to teach players the principles of nonviolent political action. As players hold demonstrations, free hostages, or take over communications stations (such as radio stations), they learn the principles and strategies of nonviolent action. Missions are designed around historical scenarios modeled to include important variables, including ''strategic and political factors, ethnicity, religion, literacy, material well-being, media and communications, resource availability, economic factors, and the role of external assistance.'' Through these missions players develop not just factual and descriptive knowledge of tactics such as ''training, fund-raising, community organizing, leafleting, protests, strikes, mass action, civil disobedience, and noncooperation'' but an appreciation for their functions in situ, for their strategic role within historical situations, and understanding of how and when they're used for action.

As anyone who has read a game FAQ knows, commercial entertainment games are actually overloaded with information—names, terms, procedures, and strategies (or moves) that players must master to be competent. As such, the information of game-based spaces differs from most traditional learning environments (and even learner-centered ones, e.g., Jonassen, 1997) in that players are given loads of data to manage through tools, databases, and online forums. A BreakAway game designer explains how this skill—understanding how people navigate multiple information streams—is essential for the next-generation designer:

> ''I think the value of games in the future will be understanding human psychology and how you interact with information as opposed to traditional instructional design skills necessarily. How do human beings react to multiple sources of info to come to an analysis? That's what we're good at without knowing. We handle massive amounts of data—letting people manage copious amounts of data very well. That's the future. It's about how this data comes across and how you analyze it and come to a conclusion.''

Indeed, even the simplest of game interfaces includes dozens of pieces of information, most of which have been streamlined for efficient use through several generations of testing with thousands of users. Consider these screenshots from Firaxis's Civilization III (Figures 13.6 and 13.7). Through evolution

Figure 13.6 A Relatively Simple Screenshot from Firaxis's Civilization III.

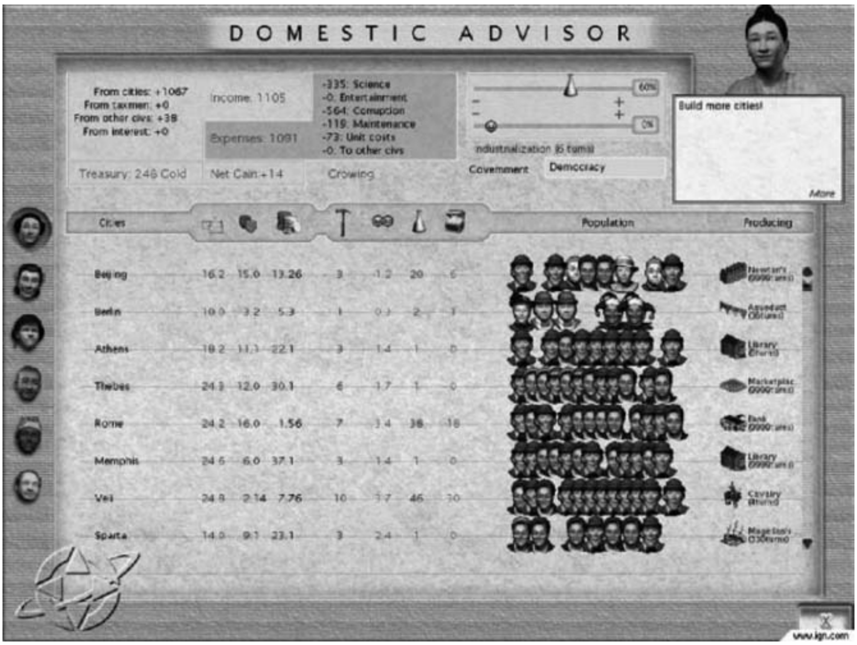


Figure 13.7 A More Complex Screenshot of Advisors from Firaxis's Civilization III.

over thousands of games in a highly competitive environment, successful design interfaces have been taken up and used, whereas bad or confusing interfaces are abandoned. Players enjoy complexity—especially the pleasure of experiencing amplified output that comes from playing with powerful tools. What they do not like is ''uninteresting decisions'' (that is, boring games), games where they are left with too many easy or inconsequential decisions (micromanagement)—decisions where there is no learning to be had.

From a training perspective, driving this move toward games is a shift from caring what the person knows or can store in his or her head toward a concern for what the person can do, given access to a full set of tools, resources, and social networks that is consistent with the situated view of knowledge (cf. Hutchins, 1995; Salomon, 1993). One of the primary benefits of games is that they can immerse players in ''smart contexts'' where they have access to and are given reason to use tools, resources, and social networks. Games such as Lineage 2 are designed to be played by hundreds of thousands if not millions of people, and mastering the game's quests, economy, and political structure requires collaborative problem solving. From this paradigm, training programs seek to give people not just user manuals or explanations of tools but also, and more crucially, experiences that demand complex information, where they use tools to make sense of multiple information streams.

Game design, perhaps more than any other area of design, is on the cutting edge of creating and supporting these digitally mediated distributed communities of expertise.

CHALLENGES THAT CONFRONT AND BUILD ON USERS' PREEXISTING KNOWLEDGE AND BELIEFS

With games, knowledge is not presented to the learner but arises through activity—activity that occurs in relation to preexisting knowledge and beliefs and the projected identities that are established for players (Gee, 2003). In the case of Root Learning, mini-games, which include matching games where players identify the fastest-growing product sector or most profitable retail outlet, elicit learners' preexisting knowledge and beliefs. Further, the mini-games are designed to draw on learners' desire to be informed participants in popular culture or knowledgeable workers, which when combined are powerful contributors to conceptual change (cf. Gardner, 1991). Critically, participants confront these beliefs in a social setting where participants must (1) actually commit to a view publicly and (2) explain their choices, which makes their cognition visible to participants, creates opportunities for reflection, and creates a mechanism for addressing conceptual changes. This design allows learners to share stories, theories, and experiences with their products, further tying the learning experience to their work outside the learning context.

This ''game-like'' approach gives the experience an entertaining feel where it is safe to disagree (much like a family game of Trivial Pursuit) and also challenges players' core assumptions about their practice. A Root designer explains:

> ''Our model is really about challenging assumptions. And we can do it in a way that no PowerPoint presentation can—by letting them challenge their own assumptions. Our basic theory is that most people are intelligent, rational people and when presented with information will come to their own conclusions. They come to the same conclusions that the organizations do, although most organizations are actually afraid of giving them information. Rather than being afraid, we try to put it in their hands and let them talk about it. Liberating information doesn't cause problems. It creates solutions.''

This notion that information should be ''liberated'' is one commonly associated with the gaming generation, a generation familiar with open source software, websites, and communities such as Wikipedia (Squire & Steinkuehler, 2005).The idea is that information in and of itself is cheap; what is valuable is the right conceptual knowledge, or organizing set of assumptions and ideas. Effective game-based learning involves structuring these challenges so that learners develop the kind of understandings that designers would like.

Games are unlike other interactive learning systems in that they contain failure states, conditions where players' choices can lead to negative consequences; game constraints push up against players' behavior, limiting what they can and cannot do. In entertainment games such as Ninja Gaiden, ''boss'' monsters ensure that players have learned all of their characters' moves (such as blocking and defending). In educational games, level design or time constraints can induce participant failure, using ''seductive failure states'' to entice learners into making mistakes that are tied to their misconceptions about a domain (Games-to-Teach Team, 2003). This is a design mechanism that has been commonly exploited in research prototypes developed at academic institutions, but less so within industry, perhaps because it necessitates approaching game design from a more traditional instructional design perspective and requires substantial background research into learners' previous conceptions (cf. Barnett, Squire, Higgenbotham, & Grant, 2004; Klopfer, Squire, & Jenkins, 2004).

In a few cases, game designers have used games' capacity to generate learning through failure as explicit selling points, suggesting that it could be a core affordance of the medium. As a Root designer explains, ''For us, learning to recover is more important than seven bullet points. How do you come about learning to recover? Making success of a failure is a key to learning through games.'' Good games should give you contexts to practice failure (and recovery safely). They are environments in which learners can and do take risks, trying on different learning strategies, learning through an abductive process of inquiry rather than a linear one of question and answer (Squire, 2005). In fact, in their studies of gamers, Beck and Wade (2004) found that this willingness to take risks and learn through failure is a characteristic trait of the gamer generation that contrasts them from their older peers.

Knowing Through Practice

Games are fundamentally about doing. Perhaps the biggest difference between game-based and more traditional approaches to learning is that game designers most often start with the user experience, specifically with what the user does. Legendary game designer Shigeru Miyamoto (creator of Mario, Zelda, and Pikmin) likes to say that he starts with verbs—what it is a player can do in a world. Imagine listing the verbs available to a learner in a classic e-learning scenario. Most likely, they are read and look; if the person is lucky, maybe chat. Many game developers begin with these verbs and then create structured problems, which build player mastery and add nuance to player skills through extensive practice involving repetition and variation. Although some marvel at the fact that games take twenty, thirty, forty, even one hundred hours to complete, in fact what is happening here is that game designers are allowing players to learn new skills and apply them in a variety of situations (cf. Chronicle). Most games structure levels so that these skills are combined and put together in new ways through time. The game Viewtiful Joe, for example, structures levels so that players must combine and use knowledge in a variety of settings, the kind of practice schedule that is useful in generating transferable knowledge and skills (Squire, 2005). Game designers build levels with new challenges, player capacities, and constraints to maximize novelty, which leads to player learning (and staves off boredom). As designer Raph Koster (2004) argues, game designers are locked in an eternal battle with their players, creating newer and newer challenges to stay one step ahead of players' skills.

Academically developed games in research contexts have shown that this structure can work, and these designs are beginning to enter the commercial space as well (Barnett, Squire, Higgenbotham, & Grant, 2004). In the case of A Force More Powerful, BreakAway's designers have created a variety of levels and scenarios so that players can try strategies in different situations and, in so doing, develop a kind of deep expertise that comes through multiple cases. As players encounter the scenarios, they practice routine skills, develop a mastery-level understanding of game basics, and acquire more flexible understanding of game content. This variety of levels both enforces mastery and prevents overgeneralization from a minimum of cases. A Force More Powerful, which also ships with a robust level editor, enables students and designers alike to create levels communicating and extending this knowledge.

The most promising model for games and training could be these kinds of levels, delivered episodically, which serve as refresher courses tailored to a particular employee's needs, much like a personalized tutorial or just-in-time experience. With their Blockbuster game, Root is creating a module that will be the equivalent to twenty hours of training, yet doled out over several months, made available on time and on demand. The idea is that players can begin by mastering basic skills in the game and then try these basics out in limited conditions in an apprentice situation (in the real world). Next, they can return to the simulation for further training (as opposed to doing a lengthy training up front). Each of the 137 modules they have designed includes context, content, practice, and then elements that take them out on to the floor to complete. In this way, the game starts to span across the real and virtual spaces, a particularly promising form of training called ''augmented reality.'' One can imagine sales representatives, or even employees themselves, identifying training needs and selecting the appropriate training.

Whereas most game-based training solutions have been thus far conceptualized as off-site, traditional instruction, their biggest potential may be in such distributed training scenarios.

Modeling the End User

Games' open-endedness poses unique challenges for instructional designers; although games differ in the amount of control users have, compared to traditional instruction, games give learners a tremendous amount of choice and freedom in choosing what to do. Doug Whatley of BreakAway Games describes some of the issues in designing within a game-based pedagogy: Most training is highly linear. You have your objectives up-front. Then you add information so that the learners can spit it back.

Creating a world in which the user is completely free, where the experience is open-ended, is a little different circumstance. We have to know lots more about them and bring it back into the environment.

It may be surprising that in designing an open-ended simulation, designers worry that they need to know more about the end-user. Good game design involves designing experience around what players might be thinking and doing, including carefully graduating complexity for the end-user.

The holy grail for game-based learning designers is to model the end-user on the basis of data gathered in situ, much like an intelligent tutor. YaYa's game engine, which was developed for data gathering for marketing (How can we infer what types of products twenty-five-year-old men in Madison, Wisconsin, prefer through their in-game choices?), suggests where the field is headed.

YaYa's engine can already gather data on users' choices, compare these choices to existing models, and potentially serve up custom content accordingly. Designers of intelligent tutors, for example, have become very good at creating models of users' behaviors (albeit within limited domains) and then programming the tutoring system to respond with customized content fitting learners' actions (Aleven & Koedinger, 2002). To date, no educational game offers this kind of assessment of learners' actions in situ, or this kind of adaptive content. Games developed in research contexts have done extensive user testing to identify learners' existing knowledge and conceptions and then structure levels accordingly. As such, the model of the user exists not in the artificial intelligence but in the design (Jenkins, Squire, & Tan, 2004). A number of commercial entertainment games are exploring how to use real-time data to customize content and adjust difficulty, suggesting paths that educational game designers might explore (Wright & Laurel, 2004). To date, however, most researchers are finding it difficult to take data generated in game-based activity and infer back cognitive understandings.

Embracing Ideology

Running through these findings is a notion that designing games for learning is not just about conveying content, but representing the world according to a particular set of rules aligned to particular viewpoints and ideologies about how the world works. A perhaps overlooked capacity of games is to frame problem situations in particular ways, to include those variables, situations, and issues that instructional designers deem important while leaving out others. Organizations turning to game-based learning share a concern in training workers to make ''better'' decisions, meaning decisions that are more in line with their goals for the company.

In one example, Root was trying to help Pepsi truck drivers understand Pepsi's move to rebrand their business because large retail stores (such as Sam's) were generating most of Pepsi's profits (cf. Harris, 2005). The map of the business terrain shows a ''logical'' progression from the 1950s to the 1980s, following a trend along different business models, including depictions of the changing beverage marketplace and trends in retail distribution.

The physical layout frames the problem landscape in a particular way (including some features and leaving out others). In this case, the map framed Pepsi's problem as maintaining continuous growth and shifting focus toward large retail stores that are more profitable than smaller convenience stores. Part of what makes Root's products powerful is that they are systematically organized to frame problems in particular ways (that is, continued growth in profit and reduction in cost is necessary).

When a company like Pepsi produces a game for delivery drivers aimed at training them to spend more time with big-box retailers and less time with mom-and-pop stores, Pepsi is more than just teaching knowledge or skills (both of which they are doing); Pepsi is also trying to get drivers to adopt its corporate values, where profits, expanding markets, and efficiency are more important than maintaining traditional customer relations or worker job satisfaction. What makes games like this—or the Jeep/Chrysler game or the A Force More Powerful game—distinctive is the way they model problem spaces according to a particular ideology and then invite users to interact with them. As such, games seem well poised in organizations that want employees to think strategically with knowledge on the fly, seeing problems the way that organizations want them to be seen, and acting in ways that are in accordance with organizational values.

However, as participatory systems, games invite learners to enter the problem space, thinking with information and making decisions in real time, which to be effective demands an openness toward information that is uncommon in most organizations. A designer from Root Learning explains: ''Most organizations [like most of education] are built on a military-style setup of command-and-control hierarchy. Information is made available on a need-to-know basis. The idea is that if people [lower-level employees] have the information, they will be dangerous. But there's no way that I [as a leader] can manage information and decision making on a task-to-task basis. It's just impossible. There is too much information and not enough time in the day. So if you want people to do the right thing, they have to have the information to make their own conclusions, and then it will happen.''

Underlying this instructional approach is an ideological shift away from hierarchical organizational models, where every decision must be vetted by upper management, and toward distributed models, where particular cultural values and an ethos permeate, driving employees to make the ''right'' kind of decisions from ''within'' rather than ''without'' (cf. Gee, Hull, & Lankshear, 1996; Levy, 1997). Game-based approaches create learning contexts where information is free, open, and discussed and made ''talkaboutable.'' They hope to create a context in which employees might openly confront and discuss beliefs and willingly take on the corporate ideology or way of seeing problems.

Changing Design Models

The previous section described how game-based approaches to training share unique qualities, some of which differ from traditional instructional design. Designers of game-based learning systems are also reporting unique approaches to instructional design (some of which may be familiar to others in e-learning). To suggest that there is one approach to game design within the commercial games industry would be mistaken; there is no one common method for game design, and there are almost as many approaches as there are game design studios. As e-learning companies begin developing games and hiring game developers, these methods permeate instructional design as well. Participants in serious games are reporting at least seven distinct, crucial themes that characterize how they design games:

1. Managing expectations;
2. Providing a holistic model of the product for clients;
3. Iterative design;
4. Early user feedback;
5. Increased importance of visual designers;
6. New business models; and
7. Distributing instructional design tasks across roles.

The next section explores these phases in more depth.

Managing Expectations

Because there are still relatively few examples of game-based learning systems in existence, expectations between clients and developers can differ greatly. Stakeholders often create their own models in their heads of what the game will be, and they will differ greatly. Deb Tillet, CEO of BreakAway, explains:

> ''The biggest, biggest concern I have with nongaming customers is that they require more education and laying out of expectations. If we are dealing with Microsoft games, they know what the milestones and deliverables are and where we should be each step of the way. We have education sessions with nongame clients about what to expect (and when). The standard military business way of making a big committee, stating the parameters of a project, and then implementing it to spec is not how you do games. So the first thing the client comes in and wants to see is 'What is the final product going to be?' We set a goal and work together. You can't lay out the specifications two years in advance with entertainment technologies.''

People's experiences with games differ greatly, with some fully aware (and expecting) real-time physics and 3D graphics, and others expecting something more like Pac-Man. But clients also need to understand how game development processes differ from traditional instructional design processes, and how this affects milestones, deliverables, and so on.

Supplying Holistic Models of the Learning Experience

Holistic models of the entire user experience can save time and money by quickly and easily illustrating key concepts to the client without wasting valuable time in preproduction. Most game designers find it difficult at the outset to provide fully detailed models of game play. Game developers are notorious prototypers, preferring to develop rudimentary models of game play to figure out what makes good game play and what does not, which creates uncertainties for designers and clients alike. Root Learning has found it useful to create animatics, storyboards that are shot on film to give a sense of a typical user experience and communicate timing, rhythm, and pacing.

Root reports that clients will often express satisfaction with initial design documents yet change their minds once they see a full animatic, which offers a sense for the entirety of a project. Developing early models of the experience also helps designers plan for what game developers call ‘‘feature creep,’’ the continuous addition of features in a project, which adds time, cost, and complexity and usually ‘‘breaks’’ a game design. It is one method for managing these expectations. These animatics function to increase early communication (a common goal of collaborative visual rapid prototyping; cf. Boling & Bichelmeyer, 1997) and create a framework for both clients and game developers to think about the instructional experience.

Iterative Design

One crucial difference between software developers trained in the games industry and the tenets of traditional instructional design processes is that game designers tend to prefer to jump in and begin coding game prototypes rather than conduct needs analysis, create design documents, or write out specifications. For most game designers, the first and most important step is finding a working core game dynamic, an interesting set of interactions that can be polished and expanded into the full game.

Having even simple objects on screen to interact with can give the clients and design team a sense of what is engaging (and not engaging), what is working, and what the experience will be like. Thus, for companies like BreakAway Games, which consist mostly of game designers, rapid iterative prototyping is the norm and fundamental to their instructional design processes. It is not unusual for a company to have a working initial prototype within a few weeks, but then spend the next months polishing and finishing the project.

Early User Feedback

With the many risks associated with game-based learning approaches (high production costs, uncertain outcomes, novel instructional approaches), design teams frequently incorporate user feedback early and often in development. Teams need to know as early as possible whether they are coordinating game play mechanics, art direction, instructional goals, and learners because changing even small features late in production can cause ripple effects, resulting in lost time and money. In entertainment games, developers such as Maxis use rapid user testing, which they call ‘‘Kleenex testing,’’ because tests are quick and dirty and you never encounter the same user twice, with literally thousands of testers to test and refine design concepts in each phase of development (cf. Jenkins, Squire, & Tan, 2004; Wright & Laurel, 2004).

Root Learning even advocates getting users in on design meetings with subject-matter experts as a way of clarifying when something is confusing or contradicts their own experience. A designer explains:

> ''Have them [users] in the design meetings. Knowing what they don't know or already know is a key way of getting there. With controversial topics, we bring them in immediately. We try to find focus groups that are the most contentious and ornery and then test with them. We try to get the most honest, direct feedback to win them over, and everyone else becomes easy. Rather than shy away, we try to bring them in early, whenever we are allowed, to make them part of sessions for open, honest feedback and to make sure the module really works for that group.''

The trick here, as with Maxis's Kleenex testing, is that users cannot be employed too often or they can become part of the design team. If they become too familiar with the product, they will lose their perspective as users. Working with teachers, we found that once a member ''truly'' becomes a part of the design team, he or she loses end-user perspectivity (cf. Squire, Makinster, Barnett, Barab, & Barab, 2003).

New Business Models

With every advance in computing power and storage capacity, the costs of game development rise, with top-end training/simulation games such as Full Spectrum Warrior costing several million U.S. dollars. Putting together the capital to fund a commercial-scale game is difficult, but with an estimated $75 million in ''serious games'' products in development, it is happening.

For many companies, the vehicle for this innovation is creating partnerships and projects that span marketing and training. YaYa Media's Jeep game is being used to advertise new vehicle models as well as train sales representatives about new features and raise their enthusiasm about the product. A number of other serious games blur these lines, including America's Army (U.S. Army recruitment and training), Homes of Our Own (industrial education and public relations for home builders), and A Force More Powerful (nonviolent training and political activism; Macedonia, 2002; Prensky, 2001; Rejeski, 2002). Many new titles are currently in development, and it seems possible that future collaborations between marketing and training will flourish as each sector attempts to respond to changes in the modern media marketplace.

The emergence of a serious games industry is also the result of consolidation and change within the games industry, driving more commercial game developers into looking at games for training, marketing, or other nontraditional purposes as new markets (Sawyer, 2002). Training games allow them to invest in new core technologies, own intellectual property, or gain retail rights to training games (such as owning the entertainment retail rights to a training game). BreakAway Games occasionally offers in-kind services up-front on training games in order to obtain commercial market rights or future contracts. Traditional game-publisher relationships usually leave developers with little power, forcing them to starve between projects, offering incentive for entertainment game developers to be entrepreneurial about locating new markets.

Interdisciplinary Design Teams

As games grow in complexity, so do the teams required to make them. A contemporary entertainment game might employ 120 people, including dozens of programmers and twice as many graphic artists and animators. Because games are a highly visual medium, traditional developers working in game-based learning report visual designers playing a more central role in game projects than in traditional instructional design. Root, for example, employs three staff members on every project, which fairly closely mirrors the breakdown in games companies: graphic artists, program managers, and programmers.

For Root, the most critical step is connecting artists and clients early in the process so that artists can understand users' needs and develop a core metaphor for the project. The project manager meets the client and obtains the basic information (objectives, goals, and institutional constraints), creating an outline for the project. Working closely with clients and a select group of users, the artists create the initial specifications, storyboard, layout, and animatics.

For Root, the main goal of this process is to understand the culture of the organization in order to understand what products will work, what the cultural values of the organization are, and what unspoken messages they might be trying to convey. Crawford says:

> ''The trick to their success is bringing together diverse people in order to talk about the client, making it something of an anthropological study of diverse ideas from people who normally wouldn't get together. MBAs and artists just don't get together; it's not logical. When you do you get a unique product working with two different thought patterns and learning styles; you get something that will appeal to all learners.''

Designers of games for learning are finding, like entertainment game designers, that a productive tension among programmers, artists, and storytellers is critical to a successful learning product. Traditional instructional designers are noticeably absent from this equation.

Distributing Instructional Design Functions

Notably, most processes for developing game-based instruction distribute traditional instructional design competencies across multiple roles and do not employ many (if any) traditional instructional designers. All of the groups I interviewed found instructional designers somewhat redundant to the skills offered by graphic artists, programmers, scriptwriters, or user-interface designers. Some companies did employ producers with instructional design experience, but more frequently they hired them on the basis of other expertise, such as interactive media design. Game design is a craft demanding knowledge and skills in psychology, interface design, art production, user testing, and software design. Executives in these companies felt that game design itself was an excellent preparation for instructional design of any sort. As one executive commented:

> ''Traditional instructional designers are stuck in old paradigms, which are all about objectives, content, and pen-and-paper assessments. We need people who can think holistically, imagine user scenarios, and understand the culture of organizations. There is no one way to do things.''

The strong sense I had after many interviews with these game-oriented companies was that an ideal instructional design curriculum, from their perspective, would include courses in narrative, usability studies, cognitive science, software production, and basic art. It was especially surprising that traditional graphic designers were preferred for their ability to interact with clients, iterate ideas, and understand different cultures. Traditional instructional designers were criticized for being ''too married to text'' and unable to work with visual media effectively.

Perhaps most important, the culture of traditional instructional design programs was seen to be at odds with the culture of contemporary media, particularly game cultures. Not only are text-based representations privileged over graphic or interactive representations, but the values underlying traditional instructional design practices—controlled information, predictability, linearity, hierarchies, and centralized control—are at odds with the values underlying the new media landscape: open access to information, flexibility, nonlinearity, user autonomy, customization, and permeable boundaries. Pushed by game cultures but also broadly indigenous to the Internet, these values are seen to be at the center of the contemporary new fast capitalist economy (Friedman, 2005; Gee, Hull, & Lankshear, 1996; Squire & Steinkuehler, 2005). As researchers have noted, the Nintendo generation has brought the cultural values of their media with them (just as the television generation did), and those companies that learn to speak their language and harness their creativity will be at a competitive advantage over those that do not (Beck & Wade, 2004).

IMPLICATIONS: GOING DIGITAL

Games' flashy graphics and powerful simulation capacities are both enticing and intimidating to instructional designers. Some see this technology as enticing learners; others may see it as pandering. But underlying many companies' shift toward gaming technologies is a recognition that today's media landscape has dramatically shifted from that of a decade ago, with more demands for learners' attention, more information at their disposal, and the business environment changing more rapidly than ever before (Jenkins, in press; Rushkoff, 1999). A number of managers are noting that today's workers thirty-five and under operate with a different motivating logic than their older counterparts, and video games have been shown to be a powerful predictor of workers' changing attitudes toward work, leisure, and life goals (Beck & Wade, 2004). This section focuses on one of the many implications for instructional designers, the need to embrace aesthetics and compelling learning experiences as more than just superfluous, which is integral to effective training.

Getting Beyond Textualism

One of the deep disconnects between contemporary learning theory and instructional design, as it is generally conceived, is what historian Paul Saettler (1990) refers to as textualism. Textualism is the belief that knowledge is ''true'' when represented through textual definitions. Textbooks, workbooks, and lectures work relatively well for generating this kind of knowledge—written explanations, definitions, and so on. Unfortunately, such descriptions, when not buffered by embodied experiences, often lay inert. This problem of ''shifting signifiers'' is represented by the student who can recite any formula from a physics textbook but cannot use them to explain basic phenomena in his or her environment (Gardner, 1991; Perkins, 1992; Whitehead, 1929).

Traditional instructional design practices, with careful formulation of objectives, functional specifications, and measurable indicators of performance, are deeply wedded to text. These companies working in digital game–based learning had less need for these traditional instructional systems design competencies and more need for skills such as communicating corporate culture through visual media, devising creative solutions to novel problems, and the ability to rapidly change directions mid-project. Traditional instructional design competencies were generally distributed across teams made up of business strategists, marketers, artists, interface designers, and programmers. One cannot help but wonder whether part of the problem is the cultures of instructional design programs themselves. In many graduate programs, it is difficult to imagine students turning in a needs analysis in the form of a picture, let alone a digital model. Although this type of practice may be foreign to instructional designers, it is common practice in other disciplines and fields.

Creating Compelling Experiences: More Than Just Eye Candy

Games have the capacity to give learners situated, embodied understandings of complex phenomena. What the boundaries and features of these understandings developed in digital environments are, we are only now investigating. But as we live in increasingly digitally mediated environments, most companies and work environments prefer employees who understand the properties of digital media, just as earlier generations preferred those adept at written text.

Implicit to this view is a focus not just on games per se but also and equally on visual media, culture, and literacy. Educators' concerns about ''eye candy'' show a deep misunderstanding, if not distrust, of popular culture and the arts.

Most e-learning [is] designed along the lines of the old paradigm of instruction—resulting in something akin to a trivia contest—as opposed to instantiating the kind of experimentation, problem solving, and collaboration that characterizes the new gaming age.

Eye candy functions in games in important semiotic ways, cueing emotions, conveying meanings, and tipping off users to new semiotic possibilities. Understanding these mechanisms is critical to game-based pedagogy; for example, one can see how Root Learning designers use such cues to create a sense of emotional connection and immediacy for Pepsi or Blockbuster workers who could otherwise quit their jobs for other similar service-sector jobs. Good artwork can serve to ramp up emotional intensity, perhaps making the player feel pressured, nervous, angry, sad, or compassionate. Game designers use these kinds of tools to make games such as Harvest Moon (a farming simulation for kids), Katamari Damacy (a game about rolling a ball), or World of Warcraft (a massively multiplayer game featuring many mundane activities) compelling.

There is a well-known saying among makers of serious games: ''If you want to take all of the fun out of it, get a bunch of educators involved'' (cf. Aldrich, 2004, 2005; Gordon & Zemke, 2000; Prensky, 2001). Questions such as ''Can I strip away the graphics and save money?'' or ''Could we succeed by making something more serious?'' are common at game-related conferences. Not every piece of instruction needs to be a fully functioning 3D environment, but instructional designers might embrace some of these conventions, enhancing the opportunity to create compelling experiences for learners.

Will e-Learning Go Digital?

Even if e-learning designers do not immediately jump to build game-based learning modules, they might look to digital games for design inspiration. Gaming communities are the cutting edge of consumer-grade simulation, artificial intelligence, and community design. One route for e-learning designers is not necessarily to design games per se but to at least understand how and why they work and then use this understanding as a means for designing other forms of educative media. One example of this, also described by Jon Goodwin of E. I. Lilly, was to allow users choices in customizing characters, enabling them to think about different variables at work in a situation (such as business or accounting; cf. Games-to-Teach, 2003).

Perhaps most important, examining games might help e-learning designers understand the mechanisms by which digital cultures work. To date, most e-learning has been designed along the lines of the old paradigm of instruction—resulting in something akin to a trivia contest—as opposed to instantiating the kind of experimentation, problem solving, and collaboration that characterizes the new gaming age. If digital cultures do embody a different set of values and ideas about learning—a set that next-generation workers are already bringing to the workplace and to training—then games could be the perfect research and development laboratory for instructional designers. Games are but one way that e-learning has an opportunity to truly go digital, to embrace the values and ideas that are indigenous to the digital age and embodied by gaming. This shift seems to be occurring in the military and certain sectors of training. It remains to be seen how e-learning will react.

ACKNOWLEDGMENTS

The author would like to thank Zhan Li for his work in the early portions of this research, particularly in helping conceptualize the study and encouraging me to focus on gamer cultures. I would like to thank Constance Steinkuehler for her careful editing of the paper. Extra special thanks to Jill Burger for her patience and willingness to work with me in crafting this report, especially for helping frame it in a way that will (hopefully) be useful.

Thanks to James Paul Gee for his intellectual contributions to and support of the ideas here, and to Judy Brown for connecting me to the Masie community.

Finally, I thank the folks at Root, BreakAway, and YaYa Media, as well Jon Goodwin from E. I. Lilly and joystick101.org.

NOTES

1. Entertainment Software Association. (2004). Top 10 industry facts. Retrieved August 1, 2004, from http://www.theesa.com/pressroom.html. In fact, the relationship between Hollywood and games is much more complex.
2. Pedagogical models such as ''learning through experience'' are common in educational research, although they have been underspecified and under-theorized (for example, Gredler, 1996; Squire, 2002).
3. See http://nationaldefense.ndia.org/issues/2005/Feb/UF-Strategists_Learn.htm for more information.
4. Keith Ferrazzi. Stanford Business School case OB-44, written by Jeffrey Pfeffer and Victoria Chang, November 15, 2003.
5. Commercial video games do the same thing with cover stories, scenarios, and cut scenes. They situate the player into a particular role. This serves several ends; it explains why the game isn't simulating everything in the world (few see the opening of Doom and want to kiss the Martians), making it unnecessary to program in all of those potential interactions.

References

Aldrich, C. (2004). *Simulations and the future of learning*. San Francisco: Pfeiffer. Volume 21, Number 2/2008 DOI: 10.1002/fig33.

Aldrich, C. (2005). *Simulations and the future of learning: An innovative (and perhaps revolutionary) approach to e-learning*. San Francisco: Pfeiffer.

Aleven, V. A. W. M. M., & Koedinger, K. R. (2000). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science*, *26*(2).

Barab, S. A., Hay, K. E., Barnett, M. G., & Squire, K. (2001). Constructing virtual worlds: Tracing the historical development of learner practices/Understandings. *Cognition and Instruction*, *19*(1), 47–94.

Barnett, M., Squire, K., Higgenbotham, T., & Grant, J. (2004). Electromagnetism supercharged! In Y. Kafai, W. Sandoval, N. Enyedy, A. Dixon, & F. Herrera (Eds.), *Proceedings of the 2004 International Conference of the Learning Sciences* (pp. 513–520). Mahwah, NJ: Lawrence Erlbaum Associates.

Beck, J. C., & Wade, M. (2004). *Got game: How the gamer generation is reshaping business forever*. Boston: Harvard Business School Press.

Bednar, A. K., Cunningham, D., Duffy, T. M., & Perry, J. D. (1992). Theory into practice: How do we link? In T. M. Duffy & D. H. Jonassen (Eds.), *Constructivism and the technology of instruction: A conversation* (pp. 88–100). Mahwah, NJ: Lawrence Erlbaum Associates.

Boling, B., & Bichelmeyer, B. (1997). Filling the gap: Rapid prototyping as visualization in the ISD process. Presentation made to the Association of Education Communications and Technology. Last retrieved January 21, 2004, from http://www.indiana.edu/_iirg/ARTICLES/prototyping/gap.html.

Chang, V., & Pfeffer, J. (2003). *Keith Ferrazzi*. Cambridge, MA: Harvard Business Publishing. Retrieved May 1, 2008, from http://harvardbusinessonline.hbsp.Harvard.edu/b01/en/common/item_detail.jhtml; jsessionid = ESNZEWHWW11YWAKRGWDSELQBKEOYIISW?id = ob44&referral = 2342.

Cognition and Technology Group at Vanderbilt. (1993). Anchored instruction and situated cognition revisited. *Educational Technology*, *33*(3), 52–70.

Cresswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage.

Cross, J., & Hamilton, I. (2002). The DNA of e-learning. In J. Cross & I. Hamilton (Eds.), *Beyond e-learning*. New York: Internet Time Group.

Erwin, S. (2004, December). Videogame market: A huge source of untapped technology. National Defense. Last retrieved August 31, 2008, from http://www.nationaldefensemagazine.org/issues/2004/Dec/VideogameMarket.htm.

Fodor, J. (2000). *The mind doesn't work that way*. Cambridge, MA: MIT Press.

Friedman, T. L. (2005) *The world is flat*. New York: Farrar, Strauss, and Giroux.

Games-to-Teach Team. (2003). Design principles of next-generation digital gaming for education. *Educational Technology*, *43*(5), 17–33.

Gardner, H. (1991). A disciplined mind. New York: Simon & Schuster.

Gee, J. P. (2003). *What video games have to teach us about learning*. New York: Palgrave.

Gee, J. P. (2004). *Situated language and learning: A critique of traditional schooling*. New York: Routledge.

Gee, J. P. (2007). *Good video games, good learning*. New York: Peter Lang.

Gee, J. P., Hull, G., & Lankshear, C. (1996). *The new work order: Behind the language of the new capitalism*. Boulder, CO: Westview.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory*. Hawthorne, NY: Aldine.

Gordon, J., & Zemke, R. (2000). The attack on ISD. *Training*, *37*(4), 42–53.

Gredler, M. E. (1996). Educational games and simulations: A technology in search of a research paradigm. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 521–539). New York: Macmillan.

Harris, P. (2005). Case study: Simulated learning adds fizz to PepsiCo's sales efforts. *Performance Improvement Quarterly*. Retrieved January 20, 2005, from http://www.learningcircuits.org/2005/jan2005/harris.htm 34 DOI: 10.1002/piq.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.

Jenkins, H. (in press). *Convergence cultures*. New York: NYU Press.

Jenkins, H., Squire, K., & Tan, P. (2004). You can't bring that game to school! Design of Supercharged. In B. Laurel (Ed.), *Design research* (pp. 244–252). Cambridge, MA: MIT Press.

Jonassen, D. H. (1997). Instructional design model for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology Research and Development*, *45*(1), 65–95.

Klopfer, E., Squire, K., & Jenkins, H. (2004). Environmental detectives: PDAs as a window into a virtual simulated world. In M. Kerres, M. Kalz, J. Stratmann, & C. de Witt (Eds.), *Didaktik der notebook-universitat* (pp. 259–274). Munich: Waxmann Verlag.

Kolbert, E. (2001, May 28). Pimps and dragons. *New Yorker*. Retrieved June 1, 2001, from http://www.newyorker.com/printable/?fact/010528fa_FACT.

Koster, R. (2004). *A theory of fun for game design*. Indianapolis, IN: Paraglyph Press.

Levy, P. (1997). *Collective intelligence: Mankind's emerging world in cyberspace*. New York: Plenum.

Macedonia, M. (2002). Games soldiers play. *IEEE Spectrum Online*.

Media-X Conference. (2003). Gaming to learn. Last retrieved August 31, 2005, from http://mediax.stanford.edu/news/sep05_03.html.

Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass.

Perkins, D. N. (1992). *Smart schools: Better thinking and learning for every child*. New York: Free Press.

Prensky, M. (2001). *Digital-game-based learning*. New York: McGraw-Hill.

Reigeluth, C. M., & Frick, T. W. (1999). Formative research: A methodology for improving design theories. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory, Vol. II* (pp. 633–652). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rejeski, D. (2002). Gaming our way to a better future. Adrenaline Vault. Last retrieved August 31, 2005, from http://www.avault.com/developer/getarticle.asp?name5 drejeski1.

Rushkoff, D. (1999). *Playing the future*. New York: New River Trade.

Saettler, P. (1990). *The evolution of American educational technology*. Englewood, CO: Libraries Unlimited.

Salomon, G. (Ed.). (1993). *Distributed cognitions: Psychological and educational considerations*. New York: Cambridge University Press.

Sawyer, B. (2002, September 30). *The next ages of game development*. Adrenaline Vault. Last retrieved August 31, 2005, from http://www.avault.com/developer/getarticle .asp?name5bsawyer1

Sfard, A. (1998). On two metaphors for learning and on the dangers of choosing just one. *Educational Researcher*, *27*(2), 4–13.

Shaffer, D. W., Squire, K., Halverson, R., & Gee, J. P. (2004). *Video games and the future of learning*. Retrieved December 12, 2004, from http://www.academiccolab.org/ resources/gappspaper1.pdf.

Squire, K. (2002). Cultural framing of computer/video games. *GameStudies*, 2(1). Retrieved from http://www.gamestudies.org.

Squire, K. (2003). Video games in education. *International Journal of Intelligent Simulations and Gaming*, *2*(1).

Squire, K. (2005). Educating the fighter. *On the Horizon*, *13*(2), 75–88.

Squire, K. D., Makinster, J., Barnett, M., Barab, A. L., & Barab, S. A. (2003). Designed curriculum and local culture: Acknowledging the primacy of classroom culture. *Science Education*, *87*, 1–22.

Squire, K. D., & Steinkuehler, C. A. (2005). Meet the gamers: Games as sites for new information literacies. *Library Journal*. Retrieved from http://www.libraryjournal .com/article/CA516033.html.

Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage. Volume 21, Number 2/2008 DOI: 10.1002/fig35.

Steinkuehler, C. A. (2004a). A discourse analysis of MMOG talk. In *Proceedings of the Other Players Conference*, Copenhagen Denmark. Retrieved from http://www.itu.dk/ op/proceedings.htm.

Steinkuehler, C. A. (2004b). Learning in massively multiplayer online games. In Y. B. Kafai, W. A. Sandoval, N. Enyedy, A. S. Nixon, & F. Herrera (Eds.), *Proceedings of the*

Sixth International Conference of the Learning Sciences (pp. 521–528). Mahwah, NJ: Lawrence Erlbaum Associates.

Swartout, W. R. (2004). Building better systems for learning and training: Bringing the entertainment industry and simulation technology together. ICEC 20.

Whitehead, N. (1929). *The aims of education and other essays*. New York: Macmillan.

Wilson, B. G., & Myers, K. M. (2000). Situated cognition in theoretical and practical context. In D. Jonassen & S. Land (Eds.), *Theoretical foundations of learning environments* (pp. 59–88). Mahwah, NJ: Lawrence Erlbaum Associates.

Wright, W., & Laurel, B. (2004). Sim smarts. In B. Laurel (Ed.), *Design research* (pp. 253–259). Cambridge, MA: MIT Press.

CHAPTER FOURTEEN

Training Complex Psychomotor Performance Skills*

A Part-Task Approach

Peter J. Fadde

This chapter focuses on training complex psychomotor performance skills, advocating a part-task approach that involves de-coupling the conjoined cognitive and motor domains for targeted training. Psychomotor performance skills typically include two types of component skills: *production* of motor actions and *recognition* of environmental conditions that trigger actions. Production and recognition skills are often intertwined in a seamless cycle of adaptive action that appears effortless when observed in an expert performer—whether that is a surgeon performing an arthroscopic ligament repair, a head sawyer segmenting a log to maximize the lumber footage, or a linebacker in American football knifing into the backfield to make a tackle-for-loss.

Despite the intertwined nature of the production and recognition components of psychomotor performance, there are benefits to keeping them artificially separated for the sake of targeted *part-task* training. Actually, it is quite typical of psychomotor training approaches to isolate and target production skills for part-task training, often using behavioral principles of chaining small, sequential steps or shaping a skill sequence from simple to complex. Newer theories of training psychomotor performance in sports favor *decision training* over

* *Note:* I would like to thank Edward Fadde for providing inspiration and expertise for the semi-truck driver training scenario that is portrayed in this chapter.

behavioral motor training (Vickers, 2007). Decision training entails incorporating recognition skills earlier in the acquisition and practice of psychomotor production skills, for instance, having a quarterback in American football practice reading defenses while practicing footwork drills. More traditionally, integrated training of recognition and production components of psychomotor skills occurs during whole-task practice. However, whole-task practice can be expensive and instructionally inefficient. Full team football scrimmages, for example, produce much less coaching of individual players than small-group drill periods. Instructional inefficiency, along with increased risk of injury in competitive play, is why college football coaches typically minimize the number and length of full-contact team scrimmages (J. Tiller, personal communication, May 24, 2003).

While whole-task practice, including high-fidelity simulation, is assumed to facilitate transfer of learning to performance, high instructional costs suggest that it should be used judiciously (Alessi & Trollip, 2001). In many cases, it can be instructionally efficient to keep the production and recognition components of psychomotor skills separate for the sake of targeted training activities that are optimized for either the psycho or the motor part and are therefore more efficient. This part-task approach to training psychomotor performance skills is based on the simple but profound notion that recognition and production components can be decoupled for targeted training and then re-coupled for transfer to performance.

The part-task production/recognition approach to training psychomotor performance skills is based on sports science research showing that experts' performance advantage over skilled but less expert performers often lies in the area of recognition skills rather than production skills and, further, that recognition skills can be targeted for part-task training that then leads to improved performance of the overall skill (Williams & Ward, 2003). The part-task training approach has far-reaching implications for training psychomotor performance skills beyond sports, especially those that are typically associated with simulator-based training such as aviation, surgery, and use-of-force in law enforcement and the military (Fadde, 2007).

CHAPTER PLAN

Before exploring the part-task production/recognition training approach, I summarize eight principles for training psychomotor performance skills. These principles are drawn primarily from the sports area and include both traditional and newer approaches. I then list guidelines for designing psychomotor training in a distinctly non-sports domain. The guidelines are based on the *2002 National Guidelines for Educating Emergency Medical Service (EMS) Instructors* (NHTSA, n.d.). The EMS training guidelines relate to a particular type of psychomotor training that is common in corporate, educational, and military contexts—that is, procedural training of adult professionals in a group workshop environment.

I then return to the part-task training of the production and recognition components of complex psychomotor performance skills. The theory and methods that support the part-task approach come from expertise research in the field of sports science, which I will review. I then outline an extended scenario that draws on the provided principles and guidelines of psychomotor training as well as the emerging sports science research on recognition training. The scenario involves designing a part-task production/recognition training program that is intended to improve the truck backing skill of over-the-road truck drivers. The production skills training component is designed to be completed at a closed-course training facility, while the recognition training component is designed to be delivered over the Internet.

PRINCIPLES OF TRAINING PSYCHOMOTOR PERFORMANCE SKILLS

Derived largely from established theories of motor learning (Schmidt & Wrisberg, 2004) and newer theories of sports coaching (Vickers, 2007), these eight principles of training psychomotor performance skills relate to practice scheduling, provision of instruction, the learner's focus of attention, feedback provided to learners, the role of feedback in simulation, and technology-based feedback. The principles recommend using:

1. Blocked practice for faster initial learning; spaced practice for better retention and transfer; decision practice for highly motivated learners;
2. Explicit instruction for faster initial learning; implicit instruction for better retention and transfer;
3. Internal focus of attention for initial learning; external focus of attention for more skilled performers;
4. Knowledge-of-performance feedback early in skill development; knowledge-of-results feedback later; fade feedback as skills develop;
5. Artificial simulation feedback early in learning; natural simulation feedback later in learning;
6. Constant, augmented feedback for initial learning; delayed augmented feedback (such as video) with more advanced learners;
7. Questioning by trainer to help advancing learners develop self-coaching; and
8. Part-task drills to train recognition skills separate from motor skills.

Practice Scheduling

The key decisions in designing practice schedules include: blocked or variable, concentrated or spaced, component or whole-skill, and chaining or shaping. Initial training of motor skills has a long tradition of relying on blocked practice of component sub-skills. For instance, each stroke in tennis has component motor skills such as footwork, grip, backswing, and striking motion. These might be taught in sequence (forward chaining): grip followed by footwork followed by backswing followed by striking motion. Alternatively, component sub-skills might be practiced through reverse chaining whereby the learner is ''given'' proper footwork, grip, and backhand and then practices executing the ball strike—giving the learner the satisfaction of a well-executed stroke, and then retracing to practice the individual components that led to it.

Component sub-skills might also be taught in an easy first sequence. This approach might also start with the striking motion because that is what is most natural (easiest) for beginning players. The training sequence might go next to footwork or backswing—whatever is the next-easiest or most natural for the player. Such a sequence might focus on grip last because it is likely to be the most abstract component for a novice player. In any case, *chaining* involves individually mastering component sub-skills through blocked practice trials.

Another alternative for initial learning and practice of motor skills is *shaping* of a learned motor skill sequence in which the learner practices a particular stroke (serve, volley, forehand, backhand) as a complete sequence but begins with a simplified performance context such as a coach tossing the ball so that it bounces to the location at which the learner is oriented to strike the ball. Practice of the full stroke would then progress through increasingly difficult ball-striking contexts.

Once a number of strokes are learned completely, if not yet refined, then the issue of blocked versus variable practice arises. A tennis player might practice serve, return of serve, volley, backhand, forehand, overhead, and drop shots in a single practice session, each in a distinct *block* of practice trials. Alternatively, a practice session might be arranged so that the player hits a few serves followed by a few volleys followed by a few overheads, and so on in a *variable* practice sequence. Or a coach could hit a ball toward the practicing player and then call out which stroke the player is to execute while the ball is still approaching to create *random* practice.

Newer theories of coaching advocate using contextualized *decision practice* that moves away from part-task drills and toward whole-task practice. Also called *hard first* practice because it incorporates complex situations right from the earliest practice sessions, this approach might have a developing tennis player practice shots in game-type sequences, for example, serve, approach, volley, and overhead shots in succession. Research shows that whole-task practice leads to slower immediate learning but better retention and transfer of learning to performance (Vickers, 2007).

Vickers conducted a study involving batting practice with college baseball players in which a *behavioral training* group practiced hitting the same type of pitch (fastball, curveball, or changeup) in blocks of fifteen pitches. A *decision training* group hit the same number of pitches, but with the types of pitches mixed randomly—much as in game conditions. The behavioral group showed greater improvement in early practice sessions, while the decision training group actually suffered a decrement in performance compared to baseline during early practice sessions. Only after eight practice sessions did the decision training group catch up to the behavioral training group. Ultimately, though, the decision training group performed better than the behavioral training group on retention and transfer tests that involved hitting mixed pitch types. Decision training, then, appears to have long-term benefits but comes with a steep learning curve that requires highly motivated coaches or trainers as well as learners.

In addition to questions of scheduling drills within a practice session, there are also scheduling options between practice sessions. Blocked or concentrated practice sessions are contrasted with *spaced* practice sessions. A developing tennis player, for example, might practice for six hours in a single session, or the player might practice for two hours in each of three separate sessions.

In general, spaced practice is considered superior to concentrated practice, variable practice superior to blocked practice, and whole-task practice superior to component practice. However, each different practice schedule may be more or less appropriate for particular learners and learning goals. For example, blocked practice is known to be less than optimal for retention and transfer—which are usually the goals of training. However, blocked practice may be called for in situations in which learners are resistant to training or lack confidence and therefore need quick and observable results to remain motivated. While general principles of optimal practice can be supported, the informed teacher, trainer, or instructional designer doesn't reject any of the options for practice scheduling outright but rather picks and chooses among practice schedules based on the present learners, goals, and context.

Principle 1: Blocked Practice for Faster Initial Learning; Spaced and Variable Practice for Better Retention and Transfer; Decision Practice for Highly Motivated Learners

Instruction. In many ways, the contrast between traditional behavioral practice and more recently articulated decision practice represents the continuing debate in learning science between direct instruction and situated or constructivist learning. Both approaches have strengths and weaknesses that recommend them in particular contexts and with particular learners. A similar contrast can be made between explicit instruction and implicit instruction. A study that involved teaching intermediate tennis players to recognize pre-serve cues shown by an opponent server found that a group given explicit instruction in what cues to look for had better initial success. A comparison group that was given implicit instruction on *where* to look but not what to look for had less initial success but ultimately better retention of the skill (Smeeton, Hodges, & Williams, 2005).

Principle 2: Explicit Instruction for Faster Initial Learning, Implicit Instruction for Better Retention and Transfer

Learners' Focus of Attention. In another study of batting practice by college baseball players, one group was directed to focus on the internal production of motor movement and another group was directed to focus on external knowledge of results (Castaneda & Gray, 2007). The researchers found that more skilled batters performed best with an external focus of attention and were hampered by an internal focus on execution of skills that were already mastered to a point of largely unconscious control. Alternatively, less skilled batters performed better when focusing attention internally on execution of skills.

Principle 3: Internal Focus of Attention for Initial Learning, External Focus of Attention for Skilled Performers

Feedback. Feedback can be described as internal (also termed intrinsic or inherent) feedback—which is natural feedback from our senses—or external (also termed extrinsic or augmented) feedback—such as that provided by viewing videotape or a coach's comments. Internal, kinesthetic feedback in psychomotor skills is often an issue with patients recovering from injury or illness (Schmidt & Lee, 2005). Internal feedback is also an essential aspect of developing expertise in golf, diving, gymnastics, and other *closed* sports and non-sports skills in which the goal of performance is to execute skill sequences as precisely as possible with relatively little adjustment for the actions of an opponent or changes in the environment. Such closed skills are less the focus of this chapter than *open* skills that still involve performers executing skill sequences but also dynamically adapting skill execution depending on actions of an opponent or changing environmental conditions.

Extrinsic/augmented feedback provided by a coach during or after performance is one of the key strategies involved in the design of psychomotor training. While it has long been the tradition of sports coaching to provide abundant levels of corrective feedback to performers at all levels of skill, research now favors *bandwidth feedback* that involves reducing and delaying feedback as a learner's skill level increases (Vickers, 2007). Interestingly, bandwidth feedback is consistent with the behavioral principle of fading reinforcement as a behavior is strengthened. Schmidt and Lee (2005) state that ''When augmented feedback is provided frequently, immediately, or otherwise in such a way that various processing activities are not undertaken, then there will likely be a decrement in learning'' (p. 398).

Extrinsic/augmented feedback can be in the form of *knowledge of results* or *knowledge of performance*. The Castaneda and Gray (2007) baseball batting study suggests that externally focused knowledge-of-results is a more appropriate mode of feedback for skilled learners, and internally focused knowledge-of-performance feedback is more appropriate for less skilled performers.

Principle 4: Knowledge-of-Performance Feedback Early in Skill Development; Knowledge-of-Results Later; Fade Feedback as Skills Develop

Simulation Feedback. Training of psychomotor performance skills often includes simulation of performance situations during whole-task practice. Feedback in a simulation activity is typed as *artificial* or *natural* (Alessi & Trollip, 2001). Artificial feedback involves the instructor or the instructional system (for example, computer-based simulation) correcting the learner during the simulation when he or she makes an incorrect decision or takes an inappropriate action. Artificial feedback is preferable early in learning to avoid reinforcing undesirable behavior. In a natural feedback condition, which is appropriate for applying or assessing skill learning, feedback is delayed and the learner will not become aware of an incorrect decision or action until the simulated patient dies or the airplane runs out of fuel mid-flight.

Principle 5: Artificial Simulation Feedback Early in Learning; Natural Simulation Feedback in Later Practice

Technology-Based Augmented Feedback. Augmented feedback is artificial, such as the score given for a successful skill execution or the verbal comments of a coach. Augmented feedback can be immediate or delayed. Generally, newer learners benefit more from immediate augmented (also called extrinsic) feedback provided during practice while experienced learners, who are often highly aware of their own performance from the inherent feedback they receive from their own bodies, can find immediate augmented feedback to be distracting or confusing.

Technology provides a key type of delayed augmented feedback. Video, and before that film, have been extraordinarily valuable learning tools for performers of complex psychomotor skills in a range of domains, from sports to surgery. Affordable video analysis tools are now available to teachers, trainers, and coaches that allow one performance by a learner to be compared side-by-side with or overlaid on another performance by the same learner or by a model performer (Dartfish, 2008). Video analysis tools allow portions of a videotaped performance to be coded and compiled so that, for example, a wrestler and his coach can study all of his takedowns—and his opponents' takedowns of him. For logistical as well as instructional reasons, technology-based augmented feedback is almost always delayed—which is considered to be less beneficial to learners just developing new psychomotor skills but more beneficial to advanced learners, who are able to recall their earlier performance and process a coach's retrospective feedback or their own observations.

Principle 6: Constant Augmented Feedback for Initial Learning; Delayed Augmented Feedback (Such as Video) with More Advanced Learners

Questioning and Self-Regulation. As performers advance to levels of expertise or near expertise in a performance domain, the role of formal and systematic training becomes less clear. While top athletes continue to practice the sub-skills of their craft on a daily basis, including receiving direction and motivation from a professional coach, most professions and skilled crafts do not have a culture of practice that includes direction and regular feedback from a coach. Performers become largely responsible for their own progression as performers. Within this progression, however, a trainer or mentor may have an opportunity to help the performer progress by using the activity of *questioning*. That is, the trainer or mentor asks questions of a performer that lead the performer to reflect critically on his or her performance. Questioning can be a step toward the performer developing the type of self-regulation and self-coaching that typify expert performers in a wide range of domains. The goal of questioning is that advancing learners progress along a path of decreasing dependence on a coach or trainer and increasing self-awareness and self-control (Vickers, 2007).

Principle 7: Questioning to Help Advancing Learners Develop Self-Coaching

Part-Task Versus Whole-Task Practice. One of the key design considerations in the training of psychomotor performance skills is whether to take a *part-task* or a *whole-task* approach to practice. The emphasis of modern instructional design and learning theory, as well as coaching theory, is in the direction of involving contextual, whole-task practice earlier and more often during instruction (Merrill, 2002; Vickers, 2007). The problem with whole-task practice such as sports scrimmages, war gaming, and simulator training is that it can be expensive. Whole-task activities, which are almost by definition contextual and experiential, tend to be instructionally inefficient in comparison to part-task methods such as drill-and-practice (Alessi & Trollip, 2001).

While experiential, whole-task learning has clear benefits for transfer of learning to performance, there are also benefits to conducting part-task, drill type training—not only of the motor production component (as is typically done) but also of the *recognition* component. The part-task recognition training approach is based on sports expertise research that has revealed recognition skills as differentiating expert performers in many reactive sports skills. Sport expertise research has further shown that such recognition skills are eminently trainable using techniques derived from expert-novice studies, most notably the method of video-simulation (Ward, Williams, & Hancock, 2006).

Principle 8: Part-Task Drills to Train Recognition Skills Separate from Motor Skills

Later in the chapter, I summarize the sports expertise research that establishes a foundation for the part-task approach to training of recognition skills using video-simulation along with emerging research that extends the approach beyond sports. Following the research review, I unfold a hypothetical example that describes the design of a program to train veteran truck drivers in the complex psychomotor performance of backing a fifty-three-foot trailer into a loading dock while avoiding often unseen and sometimes moving obstacles. The design of the truck-backing training program incorporates traditional part-task training of the motor production aspects of the skill as well as innovative part-task training activities that target the recognition aspects of this complex psychomotor performance skill.

GUIDELINES FOR TRAINING PSYCHOMOTOR PERFORMANCE SKILLS

Before moving to the review of sports expertise research and the design of the truck-backing training program, I offer the following set of guidelines for designing a particular type of psychomotor training that is common in corporate, military, and higher-education settings. That is, training adult, pre-service or in-service professionals in highly procedural psychomotor skills within a group workshop setting. These guidelines, which are based on the *2002 National Guidelines for Educating Emergency Medical Services (EMS) Instructors* (NHTSA, n.d.), relate to a non-sports psychomotor performance context that also involves rapid decisions and actions. The *EMS* guidelines have been adapted to include established and emerging instructional design principles and include recommendations related to five aspects of designing a psychomotor training program: levels of psychomotor performance, demonstrating psychomotor skills, practicing psychomotor skills, feedback during practice, and group training.

Levels of Psychomotor Performance

1. Imitation
   a. Trainee repeats what the instructor does: ''See one, do one.''
   b. Avoid modeling incorrect behaviors because trainees will do as you do.
   c. Some skills are learned entirely by observation; no need for formal instruction.
2. Manipulation
   a. Provide guidelines as a foundation for learning new procedures (skill sheets).
   b. Use forward or backward chaining of component sub-skills to build a sequence.
   c. Use blocked practice for rapid learning of newly acquired sub-skills.
   d. Interrupt and correct incorrect behavior in beginners.
3. Precision
   a. Trainees practice sufficiently to produce skill without mistakes.
   b. Trainees can perform the skill in a limited setting only. Shape behavior through increasingly challenging settings.
   c. Allow advanced trainees to identify and correct their own mistakes. Involves trainees visualizing themselves performing the skills.
   d. Let trainees develop their own style within acceptable behaviors.
4. Articulation
   a. Trainees explain why the skill is done a certain way.
   b. Trainees describe what adjustments can be made to skill sequence and when.
   c. Trainees recognize and self-correct errors.
5. Naturalization
   a. Trainees perform basic skills accurately, quickly, and with low cognitive effort.
   b. Trainees multi-task effectively by minimizing cognitive load on routine motor tasks, thus freeing mental resources for more complex cognitive tasks.
   c. Trainees perform skills confidently and competently in a variety of scenarios.

Demonstrating Psychomotor Performance Skills

1. Whole-part-whole demonstration of skills.

   a. Instructor introduces the skill by demonstrating the entire skill from beginning to end while briefly naming each section.
   b. Instructor demonstrates the skill sequence again, explaining each step in detail. Trainees may interject questions during the step-by-step demonstration.
   c. Instructor models the entire skill sequence by performing it in real time without interruption or commentary.
2. Whole-part-whole provides trainees with multiple observation opportunities.
3. Whole-part-whole appeals to both analytic learners who prefer step-by-step description and global learners who prefer an overview.

Practice of Psychomotor Performance Skills

1. After demonstration of skill sequence, trainees practice sub-skills using checklist.
2. Trainees memorize the steps of the skill until they can verbalize the sequence.
3. Trainees perform the sequence stating each step as they perform it (no checklist).
4. Trainees perform the sequence while answering questions about their performance in order to increase meta-cognitive awareness.
5. Trainees should be allowed to progress at their own pace. The need for direct supervision should lessen as trainees' skills increase.
6. When ready, trainees perform the skills in context of scenarios or simulations.
7. Trainees should be allowed ample time to practice before being tested.

Feedback During Psychomotor Skill Practice

1. Interrupt and correct incorrect or inappropriate behavior in beginners.
2. Practice sessions should end on a correct performance or demonstration of the skill.
3. Under limited supervision, allow advanced trainees to identify mistakes and make corrective adjustments in themselves and others (delayed feedback).
4. Provide trainees with positive feedback to reinforce correct behaviors.
5. Allow adults to develop their own style after mastery has been achieved. Focus on acceptable behaviors instead of rote performance.

Situated Skill Learning in Group Training Sessions

1. Assign students in each group to roles (depending on size of group) during scenarios and simulations, including:
   a. Evaluator. Uses a skill sheet, videotape, or audiotape to create a record. Multiple students can evaluate and compare observations.
   b. Information provider. Uses script to ''run'' scenario.
   c. Team leader. The primary decision-maker of the group.
   d. Partner or assistant. Gathers information to inform decisions.
   e. Patient, customer, or other central person in scenario. Portrays symptoms or behaviors according to role in scenario.
   f. Bystanders. Can depict helpful or distracting roles.
2. Instructor should not interrupt the scenario, except for safety concerns, but rather make notes for debriefing session to follow scenario.
3. Instructor provides group performance evaluation in debrief session.
   a. Use positive-negative-positive format when possible.
   b. Provide constructive criticism and areas for improvement.
   c. End with reinforcement of critical aspects of skill performance.
   d. Participants comment from the perspective of their roles.
4. Rotate roles for next scenario or simulation.

This list of guidelines is adapted from the *2002 National Guidelines for Educating EMS Instructors* (NHTSA, n.d.). Burke (1989), Kolb (1984), Millis and Costello (1998), and Watson (1980) were cited as sources in the original U.S. government document. These guidelines are specific to training a particular type of learners (adult, professional) in particular types of psychomotor performance skills (procedural, adaptive) in a particular training context (workshop). When similar skills, for example cardiopulmonary resuscitation (CPR), are taught to different learners in different contexts and with different goals, then different guidelines might apply.

The sports-based principles for training psychomotor performance skills and the EMS-based guidelines for workshop-style psychomotor training serve as a summary of current theory and practice in psychomotor training. The following sections explore an emerging part-task approach to training complex psychomotor performance skills that addresses motor production skills and recognition skills separately with targeted, optimized, and therefore instructionally efficient training activities. While still based in sports science, the foundational research has been conducted not in the area of kinesiology that has generated motor skill training principles but rather in an area of sports psychology that pursues *sports expertise* research.

SPORTS EXPERTISE RESEARCH

While sports scientists in the kinesiology area have traditionally focused on the production of motor skills, a group of sports expertise researchers in the sport psychology area have focused on decision making in *open* sports such as tennis, basketball, soccer, and hockey, rather than on skill production in *closed* sports such as golf and gymnastics. As shown in Table 14.1, closed sports are primarily concerned with the consistent execution of motor actions while open sports involve dynamically adapting actions to changing conditions, especially the actions of an opponent. This chapter extends this focus on open rather than closed skills to other domains of psychomotor performance.

The focus of sports expertise researchers on the recognition component of psychomotor skills goes back to the early 1980s and is based in general theories of expertise and expert performance that are rooted in classic chess research. Studies of expert and novice (less expert) chess players revealed that the experts enjoyed a software advantage in the form of chess-specific schema rather than a hardware advantage such as prodigious memory (Simon & Chase, 1973). In classic experiments, it was shown that expert chess players were not substantially better than less expert players at arranging chess pieces on a blank board to replicate the arbitrary arrangement of pieces on a board that they viewed for only a short time. However, when the stimulus chessboard had a meaningful arrangement of pieces from an actual game, then the experts were much better at replicating the arrangement. The researchers inferred that the expert chess players were able to encode information into chunks that could be more easily remembered and then decoded, thereby circumventing the limitations of working memory.

The classic chess experiments generated a distinct approach to the study of expertise and expert performance that has been modeled and researched in performance domains ranging from aviation to physics problem solving and including sports performance (Ericsson, Charness, Feltovich, & Hoffman, 2006).

Table 14.1 Open Versus Closed Psychomotor Performance Skills

	Open Skills	*Closed Skills*
Sports	Tennis, basketball, soccer	Golf, gymnastics, bowling
Other domains	Vehicle operation, surgery	Product assembly
Performance goal	React to opponent or environment by adjusting motor sequence	Reproduce motor sequence accurately

Sports scientists working within the expert-novice paradigm of expertise research have shown that the seat of expertise in reactive sports skills such as blocking shots on goal in hockey and soccer or batting a ball in baseball and cricket lies in the experts' ability to ''read'' an opponent's actions and anticipate the outcome more than in the production of motor actions. That is not to say that motor actions are not important but rather that production skills don't *differentiate* expert from less expert performers. Whether baseball batters or vascular surgeons, performers need to have mastered requisite production skills to be ''in the game.''

From a training perspective, the key question is whether the recognition skills that differentiate expert performers can be systematically improved through training activities. That question has been addressed by sports expertise researchers who have developed and implemented training programs that essentially repurpose the tasks used to measure expert recognition skills into training tasks to target and improve those same recognition skills. Most of these recognition training studies have targeted the ballistic and reactive skills of returning serve in tennis (Farrow, Chivers, Hardingham, & Sasche, 1998; Haskins, 1965; Scott, Scott, & Howe, 1998; Singer, Cauraugh, Chen, Steinberg, Frehlich, & Wang, 1994) and batting in baseball (Burroughs, 1984; Fadde, 2006). All of these training programs used the *video-simulation* method developed for expertise research studies in which participants view and react to a video or film display of an opponent's action (serve or pitch).

In most of the reported recognition training studies, video-simulation training of recognition skills was associated with improved performance of the full skill in either live performance-based tasks (near transfer) or in laboratory-based simulations (Williams & Ward, 2003). At least one study produced far transfer of recognition training to full-context performance. In that study, college baseball players who received ten fifteen-minute pitch recognition training sessions performed better than a control group of players from the same team in game batting performance as measured by official National Collegiate Athletic Association (NCAA) batting statistics (Fadde, 2006).

The baseball pitch recognition training program (Fadde, 2006) illustrates the close link that can be made between experimental research and skill training, not only in application of theory and findings but also adoption of research methods such as video-simulation. Video simulation involves research participants or trainees viewing a visual depiction of an opponent's action, in this case delivering a pitch, and then identifying the type of pitch or predicting the location of the pitch in the strike zone. The visual display is edited to black (temporal occlusion) at various points in the pitcher's delivery and resulting ball flight. For research purposes, the ability of more expert and less expert batters was compared at various occlusion points. An expert-novice study by Paull and Glencross (1997) found that novices performed as well as experts when more than one-third of ball flight was shown. On the other end, the performance of both experts and novices was reduced to chance at occlusion points before the moment-of-release of the pitch. The window of expert advantage, then, is between release of the pitch and one-third of ball flight. Fadde (2006) repurposed the research design as a training design by arranging stimulus video of pitches in a sequence of progressive difficulty that started by showing trainees video clips with about one-third of ball flight shown and progressing, with mastery, through clips showing less ball flight and ultimately to clips occluded at the point-of-release of the pitch.

Few psychomotor performance skills, in sports or other domains, are as ballistic as returning a 120-mile-per-hour serve or hitting a ninety-mile-per-hour pitch. However, if the recognition-action link, which appears to be inextricably linked in these skills, can be de-coupled for targeted training and then re-coupled to improve performance, then that argues convincingly for applying the approach in other, less ballistic, psychomotor performance domains. Indeed, the case has been made for applying part-task recognition training approaches in areas of performance well beyond those typically associated with psychomotor skills, such as classroom management and radiology (Fadde, 2007).

There are two key implications of part-task training of recognition skills as a component of psychomotor performance. The first is that, since recognition skills have not been systematically addressed in training in the same way that production skills have been, there is an opportunity to improve on the proven methods of psychomotor training. The other implication is that part-task training of recognition skills, separate from production skills, can be delivered much less expensively than full-task training that often involves high-fidelity simulations. In the hypothetical example of training tractor-trailer drivers to back their rigs that is elaborated later in this chapter, that means that the recognition component of this complex psychomotor performance can be trained over the Internet. Of course, drivers still need to train production skills by backing up real trucks in controlled conditions. Ultimately, the production and recognition components are re-coupled to facilitate and assess transfer of learning to real-world performance.

TRAINING THE RECOGNITION COMPONENT OF PSYCHOMOTOR SKILLS

Figure 14.1 depicts a college softball player engaging in part-task, motor skill practice. The coach leads the batter through a sequence of batting drills that emphasize different parts of her swing, a typical approach to the development of

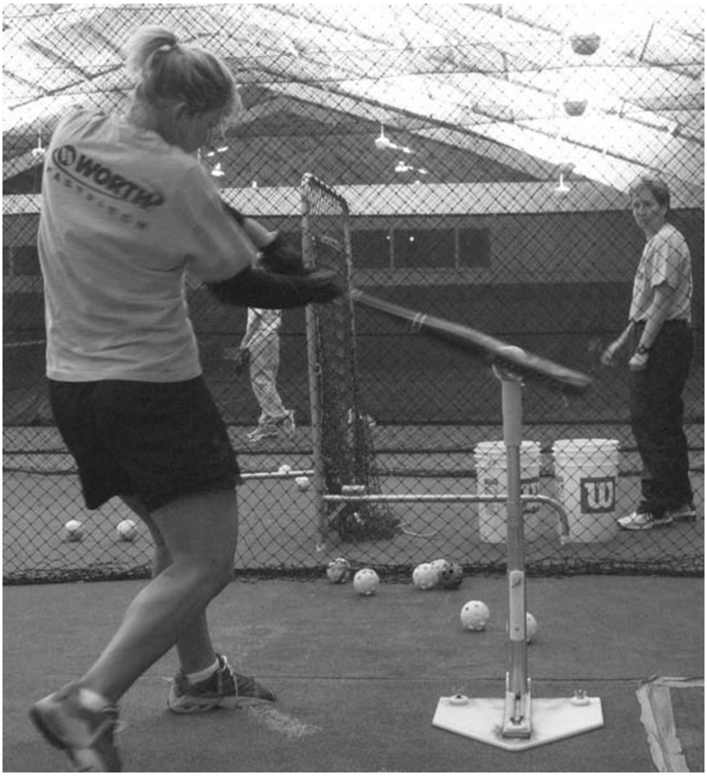


Figure 14.1 Softball Batting Practice.

complex psychomotor skills in sports. This part-task coaching approach uses the behavioral principle of *chaining* to break the batter's swing down into component segments, targeting each segment with specific drills, and then recombining the component skills to form a cohesive full-skill. The coach then uses the behavioral principle of *shaping* to add context and degrees of difficulty in moving batting practice training tasks closer to game performance.

Figure 14.2 shows a softball player engaged in a computer drill that focuses on the recognition component of the complex psychomotor skill of softball batting. Working entirely in the cognitive domain, the computer program quizzes the player on the type or location of a pitch thrown by the pitcher shown on the video screen. The design of the computer-based pitch recognition training program (*Interactive Video Training of Perceptual Decision Making*, 2007) is based on expert-novice research in sports science in which a sizable body of research has isolated early recognition and anticipation skills as the ''seat of expertise'' in many sports skills that involve rapid decision making and actions.

Figure 14.2 Pitch Recognition Training (Video Simulation).

Not only has sports expertise research provided a theoretical foundation and empirical findings to support training skills, but it has also provided model tasks that were designed to measure skills but that are readily repurposed for training skills. This is an equally important contribution as the skills that are being considered here have not traditionally been part of systematic coaching or training designs. As can be seen in Figure 14.2, *video-simulation* training involves learners viewing a video display that depicts the point of view (POV) of a live participant. The participant then engages in drills that require recognizing the type of pitch being thrown or predicting the location of the pitch in the hitting zone. Improving the pitch recognition component of batting, separate from physical batting actions, can lead to improved performance by batters who already possess requisite physical and technical batting skills. There is reason to believe that a video-simulation approach targeting recognition components can also lead to improved performance in a wide range of non-sports psychomotor skills.

BACK TO THE FUTURE: DESIGNING A PART-TASK PSYCHOMOTOR PERFORMANCE SKILLS TRAINING PROGRAM

The principles and guidelines offered earlier in the chapter, along with the recognition training approach described above, are integrated through a hypothetical scenario addressing the design and development of a training program, called *Back to the Future*, that is intended to improve the backing skills of veteran over-the-road truck drivers. *Back to the Future* demonstrates the benefits in terms of instructional effectiveness and efficiency that the part-task production/recognition approach brings to training complex psychomotor performance skills.

In this scenario, instructional design consultants from Human Performance and Learning Corporation (HPLC) design a training program for the mid-size national transportation firm PJF Fleet. The scenario reflects a typical but challenging training context in which the trainees are already advanced performers—although lacking in a newly emphasized skill set. As the instructional design consultants face an array of design decisions, they apply instructional design theory, research, and principles such as those provided in this chapter and also elsewhere in this volume. As the real world intrudes upon pure design, the consultants must also consider the client's priorities, deadlines, and resources.

Scenario: *Back to the Future* Truck Driver Training Program

PJF Fleet, a national trucking firm, has recently expanded its business into offering dedicated account service (DAS). DAS accounts essentially use PJF Fleet trucks and drivers as their own contracted fleet. PJF Fleet's training problem is that many potential DAS clients require deliveries in urban areas, a type of driving that is unfamiliar to most PJF Fleet drivers, since the company had been strictly an over-the-road carrier. Even some of the firm's ''million-milers'' (drivers who have logged over one million highway miles without accident or incident) have very limited experience driving a tractor-trailer in urban environments. The greatest area of concern to the firm, to clients, and to drivers themselves involves backing forty-eight-foot and fifty-three-foot trailers into loading docks located in congested urban areas. The firm has contracted Human Performance and Learning Consultants (HPLC) to create a training program to, in the words of the company's human resources director, ''teach our drivers how to drive backwards.''

Domain of Learning. The first thing that HPLC does is to determine what domains of learning are involved in the training project, since different domains (affective, cognitive, psychomotor) call for different training strategies. Obviously, the target skill involves complex physical movements and is therefore in the psychomotor domain. Backing tractor-trailers in urban areas also involves problem solving and planning, so there is a cognitive aspect. The cognitive aspect includes both *declarative knowledge* of company policies and applicable traffic laws and *procedural knowledge* of proper backing techniques. With any in-service training program that is required, there is also an affective aspect that can impact the motivation of the learners. However, PJF Fleet has assured HPLC that the training will not be required, but rather that new DAS accounts will be among the highest paying for drivers and will allow drivers to drive in a geographically limited region rather than cross-country. They expect that drivers will vie for the new accounts. PJF Fleet does want to assure DAS clients that their drivers are certifiably skilled at the type of driving—or in this case, backing—that DAS accounts for. PJF Fleet will require that drivers take and pass the backing program (the first of a probable series of urban driving training programs), but only if the drivers want to be considered for DAS accounts.

HPLC first looks at the declarative knowledge aspect of the training. However, an interview with the PJF Fleet driver/instructor who had been assigned to act as subject-matter expert (SME) suggests that the information to be learned—including policies, laws, and ''official'' techniques related to backing procedures—is not overly detailed or challenging to learn. The challenge, the SME insists, is lack of experience. With lack of experience comes lack of confidence. And, as the SME points out, maneuvering a fifty-three-foot trailer into a docking position while blocking multiple lanes of traffic cannot be done tentatively.

When interviewed about the kind of knowledge that a driver needs to have in order to successfully perform backing maneuvers, as opposed to pass a paper test, the SME immediately lists an array of tips that drivers use—most of which have not appeared in the technical literature and some of which contradict official policy and techniques. For example, the PJF Fleet policy requires that drivers conduct a GOAL (Get Out And Look) at least once during *every* backing maneuver and at any point in the maneuver when the driver is unsure of the location of the trailer. ''But if you did that,'' notes the SME, ''you'd never get the job done. Cars start honking and trying to get around you, and pretty soon a cop is there saying 'Driver, you've got to move this rig NOW.' And they don't take any 'yeah, buts.' So then you're calling the dispatcher and saying you can't make the delivery. That's actually what the policy says to do. But that's going to make a p.o.'d customer and make the driver look bad.''

Declarative knowledge, then, turns out to be less of a learning issue than is *procedural knowledge*—that is, knowledge that influences action. HPLC needs to figure out how to extract the real knowledge that drivers have. They also need to determine what the client's *real* goals are for the training program: do they want printed policies reinforced more than they want success in urban environments?

Training Goals. HPLC has considerable experience with designing certification training programs and immediately recognizes the need to clarify the underlying goals of the training program with PJF Fleet. Simply put, HPLC needs to know whether the training is actually intended to improve performance or if it is ''check off'' training intended to certify that employees had been given required information. PJF Fleet's vice president for safety and operations assures the consultants that this training program is indeed intended to improve performance. When presented with the SME's initial list of truck backing tips, the V.P.—who had been a driver for twenty years—laughs and says, ''Yeah, that's the real stuff.'' HPLC is now emboldened to conduct a *cognitive task analysis* (see Chapter Seven of this volume) in order to determine what experienced drivers actually *think and do* in order to plan and execute successful backing maneuvers in difficult situations.

Cognitive Task Analysis. Before conducting cognitive task analysis (CTA), HPLC consultants check industry and academic literatures looking for research or recommendations concerning backing tractor-trailers. Although they find numerous prescriptive guidelines, they do not locate information on how expert drivers actually perform maneuvers or on the performance problems that less skilled drivers have with such maneuvers. HPLC therefore conducts its own version of an expert-novice research study for the purposes of discovering what expert drivers do and think while performing, especially what they do and think that is different from less expert drivers.

HPLC recruits three representative PJF Fleet drivers who have considerable experience with over-the-road driving but little experience with local (that is, urban) driving and therefore with difficult backing situations. HPLC also recruits three independent drivers who routinely make local deliveries using ''sleeper cab'' tractors, which have a compartment behind the driver and passenger seats that serves as a sleeping room for over-the-road drivers. Most local service deliveries are made using ''day cab'' tractors, which have a back window behind the driver rather than a sleeping compartment. Day cabs are much easier to back because drivers can turn and get a ''visual'' (direct rather than mirror view) out of the back window. Backing using a sleeper cab allows the driver to get a turn-and-look visual out of the driver-side window but requires relying on the passenger-side mirror—which can result in a ''blind'' back. Simply, sleeper cab tractors are not designed for precise backing. However, the value-added proposition of PJF Fleet's new DAS service is that the same trucks that have transported the client's goods over-the-road will make the local delivery, without an interim stage of redistributing goods for local delivery.

Three PJF Fleet drivers and the three ''expert'' drivers participate in a series of representative backing maneuvers at PJF Fleet's closed-course training facility. HPLC measures the drivers' performance in terms of the speed and accuracy with which the maneuvers are executed. HPLC also videotapes each trial and then conducts *retrospective think-aloud protocol* in which the drivers talk through their cognitive and psychomotor processes while reviewing the video (''I'm trying to feel my perimeters. Feels like about two feet of clearance on the blind side. Should be able to cut the wheels . . . but getting nervous. OK, now I need to get out and check''). HPLC pays particular attention to when drivers stopped for a GOAL, sometimes rewinding the tape and asking the driver where he thought the trailer was in relation to obstacles set up on the closed-course.

HPLC's form of cognitive task analysis reveals critical differences between expert and representative (less expert) drivers in two areas: (1) maneuvering the forty-eight-foot and fifty-three-foot trailers precisely and confidently and (2) accurately estimating the proximity of the perimeters of their tractor and trailer to obstacles.

Instructional Goals. Based on the CTA, HPLC decides to address two key elements of backing in the training program, which they title: *Trust Your Mirrors* and *Steer Your Rig*. Within the psychomotor performance of backing, *Steer* addresses production skills and *Mirrors* addresses recognition skills. Identifying production and recognition as separate components has important implications for how the training modules can and should be delivered. HPLC presents and gains approval from the client for a program that by now has picked up the name *Back to the Future* (the client suggested the name as evoking the new opportunities represented by DAS accounts).

After the CTA, HPLC and the SME are able to create a set of representative backing tasks, including sighted-straight-line back, blind-side straight-line back, sighted jack-knife back, and blind-side jack-knife back. They also generate preliminary criteria for the speed as well as accuracy that would represent *mastery* of the backing tasks. Having determined the instructional objectives, the next step is to ascertain the current level of skill of the target learners and thereby determine the *performance gap*—not of individual learners (that would be determined in a pre-test stage of training) but rather of the group as a guide to creating training content.

Learner Profile. HPLC now reviews the work experience of a representative sampling of PJF Fleet drivers who have indicated an interest in DAS and finds that they vary considerably in the amount of urban driving experience that they have logged. Ironically, some of the newer drivers coming from other jobs or recent truck driver training programs have more experience and some of the veteran drivers are the least experienced. Earlier the SME had noted that many of the veteran drivers have a great deal of pride and are likely to be resistant to any kind of ''training'' program. Indeed, drivers (and practicing professionals of all types) often express disdain for training programs.

In part to demonstrate the value of training to drivers, HPLC decides to develop a performance-based *pre-test*. If drivers pass an in-truck performance test, then they will get full credit for the *Steering* component of the training program. The test-out procedure will also serve as a pre-test for the learners who do not pass the test and should convince trainees of their need for the training program as well as provide trainers with a profile of each individual trainee's performance gap.

Training Resources and Constraints. PJF Fleet will be fully compensating drivers for missed driving time in addition to paying travel, housing, and per diem costs. A primary constraint for HPLC, then, is to minimize on-site training. Resources available at the PJF Fleet training facility include one permanently installed simulator along with a portable simulator that can be scheduled into the training facility. PJF Fleet's internal training department assures HPLC that the simulators can be programmed to present backing scenarios and estimates that they could generate scenarios, including graphics representing urban obstacles, for a cost of approximately $2,000 each. The closed driving course can also be set up with simulated backing environments that could include overhead obstacles such as power lines in addition to the usual parked cars, trees, and fire hydrants. The PJF Fleet training department estimates that they could arrange a variety of closed-course backing environments for a cost of $1,000 each, which includes hiring local high school students to set up cones and obstacles on the closed-course during training sessions.

Although the availability of the simulators is tempting, HPLC opts for in-cab, closed-course training for the *Steer Your Rig* motor component of the *Back to the Future* program. There are several reasons. The first is that learners are generally more satisfied with ''live'' training using authentic equipment than they are with simulators. In addition, the SME suggested that PJF Fleet drivers associated the simulators with ''punishment'' training that was required after any moving vehicle incident.

A subtler factor is that, because the recognition component of the training program will be addressing contextual problem-solving aspects of performance, it would be acceptable to focus the motor training part of the program on developing motor skills in an essentially context-free environment (for a full discussion of highly engaging learning environments, see Chapter Twelve in this volume). One of the benefits to the part-task psychomotor training approach separating the production and recognition components of performance is that each mode can then be optimized—usually resulting in both subtask learning environments being less expensive *and* more effective than whole-skill simulation—whether live or simulator based.

HPLC now has enough knowledge of the instructional goals, the learner profile, and the client's resources and constraints to create an initial design for the *Back to the Future* training program. They focus first on the more conventional motor skill component of the training program.

Stages of Instruction

Adapting Alessi and Trollip's (2001) stages of instruction, HPLC addresses the following stages of instruction:

Assessment (pre-test)

Instruction

Practice/Application

Assessment (certification test)

These stages apply to single instructional offerings, such as a workshop, in addition to complete programs of instruction. The key consideration for HPLC in designing the *Back to the Future* training program is to determine which stages of instruction require bringing learners to the central training facility and which stages of instruction can be completed without drivers coming in from the road. HPLC has much more flexibility in designating the delivery mode for the various stages of instruction because they conceive of production skills and recognition skills as separate components of the performance and of the training program. The stages of instruction that HPLC addresses are:

Assessment (Pre-Test). The pre-test assessment will serve two instructional purposes; one is that it will reveal the level of existing skill that drivers have, which PJF Fleet trainers can use to adapt the instruction materials and activities to the learning needs of the group and of individual drivers. Some drivers may need pre-training remediation and others may pass out of the *Steer* portion of the training program. Drivers who pass out of the *Steer* module will still need to complete other training activities involving declarative knowledge of legal and policy information and the *Mirrors* recognition training module, and will have to pass the non-driving parts of the certification test.

The other purpose of the skills pre-test is to demonstrate to the learners that the desired level of backing skill is not perfunctory. The pre-test should be demanding so that any drivers who pass out of the module have clearly demonstrated superior backing skills. The demanding pre-test provides a target level of mastery that learners know they will work toward achieving and should convince them that they need the training (affective objective). The pre-test will be conducted on the closed-course range and involve a variety of backing tasks: straight line, jack knife, sighted and blind side, and with stationary obstacles and moving obstacles—with each maneuver scored in terms of speed as well as accuracy. The pre-test will be given before training to drivers who wish to test out, and then will be given at the outset of training to all remaining trainees.

Instruction. Instruction in psychomotor skills typically comes in the form of demonstration and modeling. The *National Guidelines for Educating EMS Instructors* (NHTSA, n.d.) suggests taking a *whole-part-whole* approach to demonstrating psychomotor procedures. That entails the instructor demonstrating the whole process from beginning to end, naming each component step. The instructor then goes through the process again, step-by-step, offering expert ''tips'' and answering questions from trainees. The instructor then demonstrates the whole process in real time without taking questions or making observations.

One of the *guidelines* is to have a credible model demonstrate the skills, whether that demonstration is live or in a mediated form. For the *Back to the Future* training program, HPLC decides to produce a series of videos demonstrating the techniques and procedures to be used in the various truck backing situations. By producing videos rather than having live demonstration, HPLC can control the accuracy and consistency of the demonstrations, which can have legal as well as instructional value.

HPLC also decides to use the SME to provide narration to accompany video of an expert driver executing the required maneuvers. The SME will ride in the passenger seat of the truck cab, describing each step that the driver takes in the first stage of the whole-part-whole demonstration of each maneuver and then asking the driver questions in the middle part of the demonstration. HPLC also creates a checklist of tips and procedures for each of the common backing situations. These can be kept by the trainees to use as job aids during *Back to the Future* practice and testing activities and later for continued use in the field.

HPLC plans to have trainees watch the demonstration videos on a portable DVD player in their truck cabs while sitting on the closed-course, and then transition directly into guided practice activities. HPLC provides a portable DVD player to each trainee with the instructional truck backing videos compiled on a disc. The trainees will be allowed to keep the portable DVD player (bulk cost: $70 each) if they successfully complete the on-site portion of the *Back to the Future* program. A PJF Fleet executive has suggested including a DVD of the Michael Fox comedy movie classic and, although there is no research or theory basis for doing so, HPLC embraces PJF Fleet's ''branding'' of the training program as a way to address the affective objective of having drivers accept and value the training program.

Consistent with recommendations emerging from research in cognitive as well as psychomotor learning, HPLC directs the SME to deliver externally focused implicit instruction (results oriented) rather than internally focused instruction (technique oriented). In other words, describe what the tractor or the trailer needed to be doing at a particular point in a backing procedure rather than providing step-by-step instruction in how to execute the maneuver.

Practice. Guided practice is the most important element of mastering psychomotor skills. As noted in the *Guidelines* discussion, practice can be optimized for rapid initial learning or for retention and transfer. For the *Back to the Future* training program, the emphasis is clearly on transfer of learning to performance. Further, HPLC is dealing with motivated and ''captive'' learners in that drivers are required to pass (or pass out of) the training in order to be considered for the DAS accounts for which they have applied.

Part-Task Versus Whole-Task Practice. Referring back to the earlier section discussing *Guidelines* for psychomotor performance training, the optimal design of practice would seem to be variable, spaced, whole-skill and smart practice of the sort described by Vickers (2007) as decision training. After conducting the cognitive task analysis (CTA), though, it had become clear that in this situation there is little need to train PJF Fleet drivers in the component motor sub-skills involved in maneuvering a tractor and trailer. As the SME put it, ''They know how to work their rig.'' So the behavioral training strategy of chaining component sub-skills is not relevant. *Shaping*, however, is a basic motor learning strategy to be considered along with the alternative *hard first* approach—which, as noted earlier, might better be termed *hard from the start*. In this case, HPLC identifies four distinct types of backing maneuvers to be trained: sighted straight line, sighted jack knife, blind-side straight line, and blind-side jack knife (the ultimate challenge of tractor-trailer backing). Additionally, testing and/or practice tasks have been envisioned in which moving obstacles and limited time frames are added in each of the four basic maneuvers in order to increase difficulty and realism.

Shaping in this case refers to the individual backing maneuvers. For example, shaping of the blind-side straight-line back would involve creating a simplified version of the maneuver on the closed course, perhaps with orange cones marking the target. As a driver/trainee masters the simplified version of the blind-side straight-line back, he or she progresses to more challenging and contextual versions (for instance, with a wooden construction representing a dock) of the same maneuver. Hard first practice suggests bypassing a simplified version of the task and jumping right into a more representative version of the task. A logistic advantage of the hard first approach is that the course would not have to be constantly re-set to depict progressively more difficult versions of each task. HPLC designed a shaping approach in which repeated trials on each basic maneuver would remain essentially the same but would have progressive difficulty, a key element of drill-type practice, introduced through adding time limits and physical obstacles in later trials.

Practice Scheduling. HPLC had first considered an optimal design of backing practice that would be *variable* and *spaced*. That is, trainees would make a blind-side straight-line back followed by a sighted jack-knife back and so on, in no particular order. In addition, relatively short practice sessions would be spread out over several training sessions. However, in one of those instances of logistical real-world considerations overriding optimal instructional design decisions, HPLC was told that (1) they could not have workers running around re-setting the closed course for every trial by every driver and (2) they had to minimize the number of days that drivers were off the road.

Ultimately, the four backing tasks, and the time-pressured condition for each one, are designed to be conducted in blocks. A natural progression is built into the practice activities by starting with the sighted straight-line task that was easy for almost any experienced driver, then moving to the blind-side straight-line back, the sighted jack-knife back, and finally the notorious blind-side jack-knife back. HPLC agrees to the client's goal of training up to ten drivers per training session.

Feedback and Guidance During Practice. Traditional models of sports coaching and other types of motor instruction typically involve the teacher, trainer, or coach providing concurrent, intrinsically focused (technique) feedback during practice trials. Generally, less skilled learners benefit from such *knowledge-of-performance* feedback, while more skilled learners benefit more from delayed *knowledge-of-results* feedback. With the PJF Fleet drivers being highly skilled in the performance domain of truck driving, if not yet in the targeted psychomotor skills of truck backing, recognizing that accomplished performers often have individual styles rather than insisting on rote display of techniques should help to cultivate a positive training environment.

The SME asked HPLC to consider a favorite technique for providing feedback during practice in which he stands on the step right outside the driver's window while the truck is in motion, literally getting in the driver's ear with running instructions, tips, feedback, and encouragement. The SME had used the technique previously in teaching novice drivers how to execute blind-side backs and reported that the technique was both effective and popular with drivers. However, HPLC's emerging training design includes multiple trainers working with up to ten drivers on the closed course at one time. Although the ''in your ear'' method fits the mode of constant augmented feedback that can be effective for initial skill learning, the method does not scale up well.

HPLC also has concerns that concurrent, augmented feedback—while appropriate for novice learners—might override the intrinsic feedback that drivers need to develop to be their ''own coaches'' in the field. Instead, HPLC settles on delayed augmented feedback in which every trainee's practice maneuver will be videotaped, with the video being recorded by a computer with video analysis software. The trainees who are on the course at the same time will be split into two groups. One will be executing maneuvers with tractor-trailers on the course and the other group will be reviewing the videos of their practice trials with the video-trainer. The groups then change positions and trainers for the next trial.

Test-Instruction-Practice Integration. The training tasks designed to develop the motor component of the truck-backing skills are essentially an extension of the pre-test tasks: sighted straight line, blind-side straight line, sighted jack knife, and blind-side jack knife with a time-pressured version of each added. In conjunction with PJF Fleet's internal training department, HPLC designs the pre-testing stage to transition into the instructional stage and then into the structured practice stage—all to be accomplished in a three-hour time block. They will ''own'' the closed course for the duration of the *Back to the Future* project. Ten drivers will be trained in a morning session and ten more in an afternoon session (minimizing housing costs and drivers' time off the road) until every driver who has applied for a DAS account has been trained. The training will involve five trucks and five drivers working with two trainers, an on-course trainer, and a video-review trainer. PJF Fleet will further hire temporary help to rearrange the cones and barriers that define each different backing task. While the cones and barriers are being reset, the drivers will watch the seven-minute demonstration video that goes with the next backing task. The drivers will perform the four backing tasks in blocks of thirty minutes each, with ten minutes of set-up time before each block of practice.

Watching the demonstration videos between blocks of backing practice is intended to provide a degree of *spaced practice*, which is considered to facilitate transfer of training to performance. The workers slightly rearrange the cones between trials to add a degree of *variability* within the single-task, blocked practice. The time-pressure condition is added in the final ten minutes of each block, which further adds a degree of *progressive difficulty*.

Ultimately, the blocked, single-session practice schedule is not optimal for transfer of learning to performance. HPLC would have liked to space practice over several sessions, perhaps covering a training period of several weeks, with the intent of *over training* the skills to the point of automatic and effortless execution of basic skills so that drivers' cognitive capacity would be freed for contextual problem solving in the field. The number and variability of practice trials would not seem to be sufficient to reach such a level of learning. Instead, HPLC is relying on both repetition and spacing of practice to come from the recognition component of the truck backing training program—which trainees will engage in before, during, and after the closed-course motor training sessions. In the recognition training component, trainees will be able to get far greater repetition of trials and variability of situations—although practicing without the motor skill production component.

Assessment of Performance. HPLC would have liked to arrange a post-training retention and transfer test that returned trainees to the closed course after they had completed both the *Steer Your Rig* and *Trust Your Mirrors* components of the *Back to the Future* training program. A final assessment of individual trainees' performance improvement as well as summative assessment of the training program could be accomplished using a set of representative backing maneuvers similar to those used in the closed-course training but with more contextual realism—for example, ''dressing'' the closed course with real or mocked-up cars, trees, power lines, and pedestrians. However, the client's reluctance to remove drivers from the road makes a closed-course assessment unlikely.

Instead, HPLC decides that trainees reaching criterion performance within both the closed-course production skills training and the online recognition skills training meet certification requirements, to the satisfaction of PJF Fleet and DAS customers, that drivers are ready to confidently and competently handle the variety of backing maneuvers that the customers' urban locations offer.

Mirror-Trusting Practice. The cognitive task analysis of truck-backing behavior that HPLC conducted had revealed that the two aspects of backing that drivers struggled with were the actual manipulations involved in properly orienting forty-eight-foot and fifty-three-foot trailers with a loading dock and with being able to use the mirrors on the tractor cab to avoid obstacles. So the *Steer Your Rig* closed-course training component is designed to improve the psychomotor production skills associated with manipulating the trailer in various types of backs. The *Trust Your Mirrors* component focuses entirely on the recognition aspect of backing.

Design Issues for *Back to the Future*

At this point, HPLC has to make decisions about a number of training issues based on instructional design theory and on their own research into training these advanced learners. These design issues include:

- ''Off-book'' behavior by experts;
- Recognition training: psychomotor (without the motor);
- Video-simulation on the Internet;
- Practice as implicit instruction;
- Simulation and fidelity;
- Producing video-simulation training; and
- Design of practice.

''Off-Book'' Behavior of Experts. HPLC's analysis of the expert-novice study (representative backing tasks on the closed course) that they had conducted in the course of performing a cognitive task analysis (CTA) revealed two behavioral differences between the more experienced (''expert'') drivers and the less experienced (''novice'') drivers. One was that the expert drivers consistently took a longer time to start a backing maneuver, even in time-pressured trials. While both expert and novice drivers almost always conducted a walk-around before beginning a backing maneuver, the experts sometimes appeared to ''take a moment'' in the cab. Retrospective think-aloud protocols and follow-up interviews revealed that experts took the extra time to visualize the physical setting and to mentally rehearse the approach that they planned to take in order to get the trailer properly oriented to the dock.

This focus on *problem representation* is consistently found in expert performers in a range of domains, both psychomotor and cognitive. For example, in a classic expert-novice experiment, physics graduate students (experts) routinely took longer than undergraduate physics students (novices) to start solving physics problems—which the experts then completed more quickly than the novices (Chi, 2006). The mental rehearsal stage is also consistent with Gary Klein's model of *recognition-primed decision-making* (RPD). Klein and associates have studied fire marshals, neonatal emergency room nurses, and military field commanders and found that, when confronted with a performance situation, experts spontaneously generate a course of action and then mentally simulate the action being taken. If the outcome of the mental simulation is satisfactory, the course of action is undertaken. If not, then the expert engages a more effortful process to generate an alternative course of action (Klein, 1998).

The other observable behavior difference among the novice and expert backers was that, once they started a backing maneuver, the expert drivers got out of the cab (GOAL) many fewer times than the novice drivers did. ''See, he doesn't trust his mirrors,'' said the SME while watching a video of a novice driver. The official PJF Fleet policy is that every backing maneuver should include at least one GOAL and that drivers should use a GOAL whenever they are uncertain of the proximity of their tractor or trailer to obstacles. The reality, however, is that drivers can lose control of a pressured situation with impatient ''four wheelers'' honking their horns or attempting to move around a tractor-trailer that is blocking traffic. Expert backing performance represented drivers using a GOAL when necessary, but also minimizing their use of this time-consuming and confidence-draining tactic.

As noted earlier, HPLC recognized a significant mismatch between the official policy and the observed behavior of experts—which is often revealed in a CTA. Another mismatch between policy and practice was that drivers were directed by policy to use a spotter to assist in executing backing maneuvers—or not make the delivery. However, drivers interviewed by HPLC consistently noted that spotters cannot always be recruited on site, and even when they are, spotters are not always consistent or reliable. The drivers' reality is that spotters are a luxury and that drivers need to be prepared to execute backing maneuvers without using spotters. HPLC checked with high-level PJF Fleet officials (primarily the V.P. of safety and operations) before committing to developing a training program based on expert behavior rather than official policy.

Recognition Training: Psychomotor (Without the Motor). HPLC had two goals for the *Trust Your Mirrors* recognition training component of *Back to the Future*. One was that the training would build the types of recognition skills demonstrated by experts. The second was that recognition training should be delivered, if possible, over the Internet so that drivers would need to leave the road and come to the close-course training facility only for the motor production aspect of training.

From the CTA it was clear that two types of recognition training were called for. The first would be oriented toward trainees learning and practicing *situation awareness*—that is, sizing up the delivery scenario in terms of deciding how to set up the backing maneuver in order to minimize the duration and extent of blocking traffic or otherwise being in an exposed position with the tractor-trailer. This type of strategic and deliberate ''before the action'' recognition skill is fully in the cognitive domain and was well within HPLC's experience to produce. HPLC would design two situation-recognition tasks. The first would require the learner to *identify* the type of setting that was depicted in graphical or photo-realistic views of a variety of docking situations. A second task would require learners to choose from alternative courses of action and *predict* the outcome of the chosen course of action.

This kind of training module was well within HPLC's experience and capability to produce and to delivery via the Internet. Situation-awareness training represents the type of cognitive problem-solving skill that is more fully considered in other chapters in this volume (see especially Chapter Ten, Instructional Strategies for Directive Learning Environments, and Chapter Twelve, High Engagement Strategies for Simulation and Gaming) and won't be described in detail here.

The second type of recognition skill, however, was outside of HPLC's previous instructional design experience. These are the type of recognition skills that have been studied and trained by sports expertise researchers, as described in an earlier section of this chapter. The appeal to HPLC of the recognition-only training approach was that it not only provided an approach to systematically training an essential aspect of expertise in the target skills of truck backing but that it could potentially be delivered over the Internet, thereby minimizing trainees' time at the close-course training center.

Video Simulation on the Internet. The traditional instructional design approach to training recognition skills is to combine recognition and production skills in realistic, whole-skill psychomotor practice activities. In domains such as aviation and surgery, whole-task training typically involves *simulation*, either ''live'' or using a simulator. It can be a very expensive approach, but one that is often justified by safety and cost issues with real-world training (Alessi & Trollip, 2001). For the *Back to the Future* project, however, HPLC decided to apply the theory and methods developed in sports expertise research to train recognition separate from rather than combined with psychomotor skill production.

The focus of the *Trust Your Mirrors* module was to have trainees practice making judgments about the ''perimeters'' of their tractor and trailer *during* backing maneuvers. The expert drivers participating in the CTA had repeatedly emphasized the importance of drivers being aware of not only the back end of the trailer but also the top of the trailer and the steps, fuel tank, and tires of the tractor. Drivers could get so focused on maneuvering the trailer into docking position and avoiding obstacles that it would be easy to overlook the other perimeters on the tractor as well as the trailer.

The instructional objective of the *Trust Your Mirrors* module, therefore, was

> ''Given photo-realistic (video) images depicting the driver-side and passenger-side mirror views of an in-progress tractor-trailer backing maneuver, the learner will detect any violations of the safe proximity zone of seen and unseen obstacles in relation to all of the tractor-trailer's perimeters.''

Practice as Implicit Instruction. Within the HPLC design, trainees could potentially engage in the online recognition training before, during, or after engaging in the *Steer Your Rig* closed-course motor production training. The videos produced for use in the closed-course training module would be posted on PJF Fleet's training webpage—along with a link to the *Trust Your Mirrors* module—so that the videos could be reviewed by trainees if and when they felt they needed to have backing maneuvers demonstrated. However, the instructional goal of the *Mirrors* module would be achieved almost entirely through practice rather than instruction. Consistent with principles summarized earlier in the chapter, the advanced learners involved in the truck-backing training program would benefit most from implicit instruction in the form of externally focused *knowledge of results* feedback during practice trials and *augmented feedback* in the form of scores displayed during and after the online drills.

Simulation and Fidelity. Note that the instructional objective, and therefore the training design, does **not** include the trainee manipulating the movements of a virtual truck. This is an essential difference between simulators and video-simulation. As described by simulation researchers at the University of Central Florida's Institute for Simulation and Training, truck simulators are categorized into four levels of fidelity (Allen & Tarr, 2005; Tarr, 2006). Level Four is represented by full-size in-cab simulators with three-dimensional, computer-generated visual display, functional controls, and realistic movements of the truck cab. Such high-fidelity simulators cost between $500,000 and $2,000,000 and are typically, and appropriately, used for research rather than routine training purposes.

Level Three in-cab simulators, such as that shown in Figure 14.3, feature some cab motion and computer-generated through-the-windshield visual displays that change in response to learners' manipulation of a realistic steering wheel, gear shifter, and brakes. Level Three truck simulators typically cost $100,000 to $250,000 and are used for training as well as research. Level Two is represented by non-motion, partial-cab simulators that typically cost between $45,000 and $80,000 and are commonly used for training purposes.

Level Two simulators typically have realistic steering wheels, gear shifters, brakes, and instrument panels. Most modern truck simulators of this fidelity level still feature computer-generated graphic displays (although much less immersive) that change in response to trainee input via steering, shifting, and braking devices. In older multi-seat simulators, the display was video or film, and therefore not responsive to individual learners' input. As many as eight or ten trainees viewed the same projected display and still manipulated steering wheel and brake—but without the visual display changing in response. Level One is represented by desktop truck simulators that display animated graphics on a computer screen and may include non-realistic steering wheel and brake

Figure 14.3 Level Three Truck Simulator.

Photo courtesy of MPRI, a division of L3 Services, Inc.

Figure 14.4 Level One Truck Simulator.
Photo courtesy of J. J. Keller and Associates, Inc., Neenah, WI

components—usually repurposed video game devices (see Figure 14.4). Level One simulators can be purchased for between $2,500 and $15,000 (Tarr, 2006).

The interesting and complex issue of fidelity in simulators is considered in detail in Chapter Twelve, *High Engagement Strategies in Simulation and Gaming* (in this volume). Here the key point is that the video-simulation approach is outside of the continuum of simulator fidelity (Tarr, 2006). In terms of responsiveness to learner input, video-simulation is very low fidelity. In fact, the internal training staff at HPLC was mystified as to why an apparently low-fidelity simulation was being used rather than the Level Two truck simulators that PJF Fleet owned and operated. HPLC explained that the volume of trainees could not be moved through the simulators in the target time frame. In addition, the *Back to the Future* training program already included very high-fidelity simulation in the form of the closed-course *Steer Your Rig* training. The online *Trust Your Mirrors* module would serve to both enhance and focus the live in-cab training.

Although the video-simulation approach might be cast as low-fidelity simulation, it should be noted that the video display is actually higher fidelity than the display in even Level Four truck simulators. If a visual display in a simulator changes based on user input, then the display needs to be computer generated. The realism of the computer graphic program's interpretation of visual objects and movement, then, becomes a limiting factor in the realism of the display.

Video display of a driver's view through the windshield or of side mirrors is actually more realistic, if not responsive.

Clearly, the learning objective in recognition training requires a realistic visual display, so this is a case where a low physical fidelity, high cognitive fidelity simulator is more appropriate than a high physical fidelity but low cognitive fidelity simulator (Foshay, 2006). Indeed, cognitive load theory would argue that high physical fidelity in this case is not only not unnecessary but may actually produce *extraneous cognitive load* that interferes with the target learning (van Gog, Ericsson, Rikers, & Paas, 2005).

Producing Video-Simulation Training. HPLC realized early in the project that the videos and the software programming for the video-simulation activity could have use across the trucking industry and the truck driver training industry. HPLC negotiated with PJF Fleet to have the client pay 25 percent of the cost of developing instructional materials and software with HPLC picking up the rest. In return, HPLC would own the video footage and software programming and PJF Fleet would have a standing license to use it.

HPLC's design for the video-simulation activities was to videotape actual trucks backing into actual locations and covering all four types of backing (sighted straight line, blind straight line, sighted jack knife, blind jack knife). A video production company was contracted to shoot and edit the videos for the video simulation. Four video cameras were used to shoot each truck-backing maneuver. One camera was fixed to view the driver's view of the driver-side mirror and another camera was fixed to view the driver's view of the passenger-side mirror. A third camera was placed behind the truck and simulated a GOAL (get out and look) by the driver. A fourth camera was positioned on a boom about twenty-five feet in the air and placed around fifty feet in front of the truck as the truck was in position to start a backing maneuver. This "bird's eye" camera is typically offered as a computer animation in truck simulators to help build drivers' association between what they can see in their mirrors or a GOAL and the actual position and spacing of objects.

At each videotaping location, each of the types of backing maneuvers was videotaped repeated times with slight variances between backing repetitions. Some backs were executed perfectly, while others depicted a variety of backing miscues. The SME assured that every type of miscue was videotaped and kept a log sheet coding each back. The set of backs was recorded with both a standard-length (forty-eight-foot) trailer and a long (fifty-three-foot) trailer. Each approach was then edited in an identical format with all four camera angles depicting the same action and coded for type of backing maneuver, type of trailer, and type of miscue (if any). Computer programmers at HPLC then licensed an interactive sports training software program and adapted it to fit the truck backing content.

HPLC designed a series of training drills that required learners to pause the video when they detected a backing miscue and then to check the GOAL camera view. A point scoring scheme, which represents augmented rather than intrinsic feedback, was devised that gave trainees a beginning score and then subtracted points for each GOAL taken, but subtracted more points for violating the designated proximity zone of objects on any of the trucks' perimeters. The scoring scheme rewarded learners for extending beyond their comfort zones in ''trusting their mirrors'' but without missing any miscues. As with any drill methodology, the goal was for learners to increase their speed while maintaining their accuracy. The drill characteristics of *repetition*, *feedback*, and *progressive difficulty* were enacted in the video-simulation program (Alessi & Trollip, 2001).

About two hundred video items were produced (ten instances of four types of backing maneuvers at five locations). The item pool for each drill could be designated by type of backing maneuver or by location. Because the items were formatted in the same way, the items could be randomized for presentation, meaning that learners could engage the video-simulation program repeated times. While some items would be repeated, the order of presentation would be different each time. Progressive difficulty could be set by the learner through selecting more difficult backing maneuvers and/or locations. More difficult maneuvers and more difficult locations provided more scoring points.

The key instructional design element of the video simulation was that learners did not manipulate the truck or the mirrors in any way. The learner interacted with the program by *detecting* miscues based on mirror views. In usability testing of the video-simulation program with non-PJF Fleet drivers, participants rated it as challenging but not stressful, and participants consistently underestimated the time that they spent on the *Trust Your Mirrors* drills, suggesting a high level of engagement. A number of usability participants expressed a desire to try the program again in order to beat their own scores or to beat the best score of another participant.

Design of Practice. Consistent with long-established guidelines that suggest providing mature and motivated learners with a high degree of learner control (Alessi & Trollip, 2001), *Trust Your Mirrors* drills were designed so that learners could select and create drills themselves and to decide for themselves when to advance to different or more difficult drills. HPLC produced an introduction video in which the SME explained the benefits of *blocked, spaced, variable*, and *random* practice. He also advised trainees to turn the obstacle warning system on during initial learning and early practice but turn it off for later practice and self-testing. The trainee would make the choice, although the *best scores* display on the PJF Fleet training web page would only record scores achieved with the warning system off.

Video-Simulation Test. HPLC's design for assessment in the *Back to the Future* training program required PJF Fleet trainees to register and log in on PJF Fleet's training website to engage in video-simulation drills. Trainees could engage pre-selected drills or could mix and match to make new drills. When the trainee felt ready, he or she selected *test* mode in which the system did not provide feedback and the final score was saved in a database. PJF Fleet offered drivers reward gifts from the company's gift catalog (used to motivate a range of desired driver behaviors) based on the number of *Back to the Future* drills completed and for the high score on selected drills. When a driver met criteria for performance on the *test* drills and had previously met criterion performance on the closed-course backing drills, then the driver was certified as qualified for DAS assignment.

Consideration was given to devising a *performance-based test* of the complex psychomotor skills of backing a tractor-trailer, perhaps by ''dressing'' the closed course to simulate prototypical urban loading dock scenarios. Without a whole-task, high-fidelity simulation test, it was impossible to be certain that the motor *production* component and the *recognition* component that had been de-coupled for training purposes would be successfully re-coupled with a measurable improvement in performance. However, while truck backing performance is important and mistakes are costly, it isn't in the same category as surgery or aviation, and the cost/benefit consideration didn't justify the cost of creating and implementing a high-fidelity *transfer* test.

As in most training evaluations, satisfaction of learners and client would be measured as well as learners' achievement of the defined objectives of the training program. Measures of performance, and therefore measures of transfer of learning to performance, remain elusive (except in sports) and are usually beyond the interest and ability of corporations or institutions to pursue. While transfer from training to performance might not be measured as often or as thoroughly as we would like, some value can come in the form of *pseudo-transfer* from one simulated environment to another (Lee, Chamberlain, & Hodges, 2001). If and when PJF Fleet's internal training department programs backing scenarios into their in-house simulators (which would be appropriate for training current and future drivers who are not able to attend the *Steer Your Rig* closed-course training sessions), then correlating scores on the simulator with scores on the closed-course tasks and the video-simulation drills could become very interesting. It would be expected (but would be worth investigating) that recognition skill training alone or production skill training alone would not lead to as much improvement in whole-skill performance (in the truck simulator) as would the separate but complementary training of the production *and* recognition skills that make up this complex psychomotor performance.

CONCLUSION

After summarizing current theory and practice of psychomotor training by providing lists of principles and guidelines, I have used the hypothetical *Back to the Future* scenario to illustrate the benefits and demonstrate the process of designing part-task training that addresses motor production skills and recognition skills separately but equally. This paradigm-shifting approach is based on sports expertise studies that date back at least to Haskins' (1965) film-based training of tennis serve recognition, and researchers in sports science continue to conduct video-simulation studies in a widening array of sports (Ward, Williams, & Hancock, 2006).

The sports scientists and cognitive psychologists who are conducting recognition training programs are beginning to investigate instructional design questions that are of interest to teachers, trainers, and instructional designers—and serve as a model for instructional design research and practice. For example, studies have investigated the use of explicit versus implicit instruction (Smeeton, Hodges, & Williams, 2005) and internal versus external focus of attention (Castaneda & Gray, 2007). Beyond sports, a group of researchers is adopting recognition research and training methods developed in sports science to domains such as use-of-force decision making by military and law enforcement personnel (Tashman, Harris, Ramrattan, Ward, Eccles, Ericsson, Williams, Roderick, & Lang, 2006) and critical care nursing (Ward, 2008). The leading researchers in the area have also published in a special issue of *Military Psychology* dedicated to connecting sports science findings and methods to military training (Eccles, 2008; Ward, Farrow, Harris, Williams, Eccles, & Ericsson, 2008; Williams, Ericsson, Ward, & Eccles, 2008).

Sports provide a natural context to draw from in designing training of psychomotor skills. It also provides a rich test bed for research and training in psychomotor learning and performance, in part because athletes and coaches have a ''culture of practice'' (MacMahon, Helsen, Starkes, & Weston, 2007) that other professions don't have and partly because performance is so much more clearly observable and measurable in sports. However, the implications of this recognition-training line of research make it worth investigating as a training approach in a wide range of domains (Fadde, 2007). With the continued improvement of video transmission over the Internet and the growth of web-based training in general, the potential for systematically training essential recognition aspects of expert psychomotor performance using video-simulation delivered online is enticing.

References

Alessi, S., & Trollip, S. (2001). *Multimedia for learning: Methods and development*. Boston: Allyn and Bacon.

Allen, T., & Tarr, R. (2005). Driving simulators for commercial truck drivers: Humans in the loop. *Proceedings of the Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*. Retrieved December 17, 2006, from http://ppc.uiowa.edu/driving-assessment/2005/final/papers/49_Talleah Allenformat.pdf.

Burke, J. E. (1989). *Competency-based education and training*. New York: The Falmer Press.

Burroughs, W. (1984). Visual simulation training of baseball batters. *International Journal of Sport Psychology*, *15*, 117–126.

Castaneda, B., & Gray, R. (2007). Effects of focus of attention on baseball batting performance in players of differing skill levels. *Journal of Sport and Exercise Psychology*, *29*, 60–77.

Chi, M. T. H. (2006). Laboratory methods for assessing experts' and novices' knowledge. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 167–184). New York: Cambridge University Press.

Dartfish Video Solutions (2008). Retrieved April 2, 2008, from http://www.dartfish.com/en/index.htm.

Eccles, D. W. (2008). The expert's circumvention of natural human resource limitations: An example from sport. *Military Psychology*, *20*, S103–S121.

Ericsson, K. A., Charness, N., Feltovich, P. J., & Hoffman, R. R. (2006). *The Cambridge handbook of expertise and expert performance*. New York: Cambridge University Press.

Fadde, P. J. (2006). Interactive video training of perceptual decision-making in the sport of baseball. *Technology, Instruction, Cognition and Learning, 4*(3), 265–285.

Fadde, P. J. (2009). Instructional design for advanced learners: Training expert recognition skills. *Educational Technology Research and Development*. DOI 10.1007/s11423-007-9046-5.

Farrow, D., Chivers, P., Hardingham C., & Sasche, S. (1998). The effect of video-based perceptual training on the tennis return of serve. *International Journal of Sport Psychology*, *29*, 231–242.

Foshay, W. R. (2006). Building the wrong simulation: Matching instructional intent in teaching problem solving to simulation architecture. *Technology, Instruction, Cognition and Learning*, *3*, 63–72.

Haskins, M. J. (1965). Development of a response-recognition training film in tennis. *Perceptual and Motor Skills*, *21*, 207–211.

Interactive Video Training of Perceptual Decision-Making. Software program and training program design. United States Patent Office application #20070005540, assigned to Southern Illinois University, January 4, 2007.

Lee, T. D., Chamberlain, C. J., & Hodges, N. J. (2001). Practice. In R. N. Singer, A. H. Hauseublas, & C. M. Janelle (Eds.), *Handbook of sport psychology* (pp. 11–143). Hoboken, NJ: John Wiley & Sons.

Klein, G. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.

Kolb, D. A. (1984). *Experimental learning*. New York: Simon & Schuster.

MacMahon, C., Helsen, W. F., Starkes, J. L., & Weston, M. (2007). Decision-making skills and deliberate practice in elite association football referees. *Journal of Sports Sciences*, *25*(1), 65–78.

Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research and Development*, *50*(3), 43–59.

Millis, B., & Costello, P. (1985). *Cooperative learning for higher education faculty*. Phoenix, AZ: Oryx Press.

National Highway Traffic Safety Administration (n.d.). *2002 National Guidelines for Educating EMS Instructors.* Retrieved March 30, 2008, from http://www.nhtsa.gov/people/injury/ems/Instructor/TableofContents.htm.

Paull, G., & Glencross, D. (1997). Expert perception and decision making in baseball. *International Journal of Sport Psychology*, *28*, 35–56.

Scott, D., Scott, L., & Howe, B. (1998). Training anticipation for intermediate tennis players. *Behavior Modification*, *22*(3), 243–261.

Schmidt, R. A., & Lee, T. D. (2005). *Motor control and learning: A behavioral emphasis* (4th ed.). Champaign, IL: Human Kinetics.

Schmidt, R. A., & Wrisberg, C. A. (2004). *Motor learning and performance: A problem-based approach*. Champaign, IL: Human Kinetics.

Simon, H. A., & Chase, W. (1973). Skill in chess. *American Scientist*, *61*, 394–403.

Singer, R. N., Cauraugh, J. H., Chen, D., Steinberg, G. M., Frehlich, S. G., & Wang, L. (1994). Training mental quickness in beginning/intermediate tennis players. *The Sport Psychologist*, *8*, 305–318.

Smeeton, N. J., Hodges, N. J., & Williams, A. M. (2005). The relative effectiveness of various instructional approaches in developing anticipation skill. *Journal of Experimental Psychology–Applied*, *11*(2), 98–110.

Tashman, L. S., Harris, K. R., Ramrattan, J., Ward, P., Eccles, D. W., Ericsson, K. A., Williams, A. M., Roderick, D., & Lang, L. H. (2006). Expert performance in law enforcement: Are skilled performers more effectively constraining the situation to resolve representative dynamic tasks than novices? *Proceedings from the 50th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 1213–1217). Santa Monica, CA: Human Factors and Ergonomics Society.

Tarr, R. (2006). *A white paper: Truck driver simulation*. Neenah, WI: J. J. Keller & Associates. Retrieved April 30, 2008, from http://www.jjkeller.com/information centers/safesim/home.htm.

van Gog, T., Ericsson, K. A., Rikers, R. M. J. P., & Paas, F. (2005). Instructional design for advanced learners: Establishing connections between the theoretical frameworks of cognitive load and deliberate practice. *Educational Technology Research and Development*, *53*(3), 73–81.

Vickers, J. N. (2007). *Perception, cognition, and decision training: The quiet eye in action*. Champaign, IL: Human Kinetics.

Ward, P. (2008). Expertise and expert performance in multiple domains. Paper presented at the annual meeting of the American Educational Research Association, New York City.

Ward, P., Farrow, D., Harris, K. R., Williams, A. M., Eccles, D. W., & Ericsson, K. A. (2008). Training perceptual-cognitive skills: Can sport psychology research inform military decision training? *Military Psychology*, *20*, S71–S102.

Ward, P., Williams, A. M., & Hancock, P. A. (2006). Simulation for performance and training. In K. A. Ericsson, N. Chariness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 243–262). New York: Cambridge University Press.

Watson, A. (1980). *Learning psychomotor skills in TAFE: Educational psychology for TAFE teachers*. Canberra, Australia: Advanced Education Council. (ERIC Document Reproduction Service No. ED228406).

Williams, A. M., & Ward, P. (2003). Perceptual expertise: Development in sport. In J. L. Starkes & K. A. Ericsson (Eds.), *Expert performance in sports: Advances in research in sport expertise* (pp. 219–247). Champaign, IL: Human Kinetics.

Williams, A. M., Ericsson, K. A., Ward, P., & Eccles, D. W. (2008). Research on expertise in sport: Implications for the military. *Military Psychology*, *20*, S123–S145.

PART FOUR

EVALUATION

A common lament among observers of ADDIE-based instructional design is that evaluation of effectiveness is honored more in the breach than in practice. If you are a perceptive reader of this handbook, however, you will realize that *at every stage, you cannot make sound design decisions unless you test and validate as you work.* It is easy for an instructional designer to build the wrong instruction: all you have to do is proceed on the basis of an incorrect design assumption. If you (and your client) care about effectiveness of training, and want to minimize the costs of doing it right, then evaluating your work is no longer optional.

This section is brief, because the third volume in this series goes into great detail about measurement and evaluation in the workplace. We have chosen to include here only two chapters that are most relevant directly to ID and that should be part of the practice of every skilled ID.

Chapter Fifteen: An Overview of Level Two Certification Processes. The authors indicate that it is most common in training to directly assess learning outcomes, yet this is often done badly. Stated differently, a poor assessment is actually dangerous, because it gives you, your learners, and your client misleading information. Assessment of learning outcomes is important because it gives your client sound information about training effectiveness and because it gives you sound information to refine and advance your design thinking.

Chapter Sixteen: The Role of Evaluation in Instructional Design. This chapter outlines the methodology instructional designers should use to evaluate their work, and it shows how important decision making uses this information. In the design framework of this book, evaluation is central to the design process, not an afterthought.

CHAPTER FIFTEEN

An Overview of Level Two Certification Processes

Sharon A. Shrock
William C. Coscarelli

Today's business and technological environment has increased the need for assessment of human competence. Any competitive advantage in the global economy requires that the most competent workers be identified and retained. Furthermore, training and development, HRD, and performance technology agencies are increasingly required to justify their existence with evidence of effectiveness. These pressures have heightened the demand for better assessment and the distribution of assessment data to line managers to achieve organizational goals. These demands increasingly present us with difficult issues. For example, if you haven't tested, how can you show that those graduates you certify as ''masters'' are indeed masters and can be trusted to perform competently while handling dangerous or expensive equipment or materials? What would you tell an EEO officer who presented you with a grievance from an employee who was denied a salary increase based on a test you developed? These and other important questions need to be answered for business, ethical, and legal reasons. And they can be answered through doable and cost-effective test systems.

As certification and competency testing are increasingly used in business and industry, correct testing practices make possible the data for rational decision making.

WHY READ THIS CHAPTER?

Corporate training, driven by competition and keen awareness of the ''bottom line,'' has a certain intensity about it. Errors in instructional design or employees' failure to master skills or content can cause significant negative consequences. It is not surprising, then, that corporate trainers are strong proponents of the systematic design of criterion-referenced instructional systems. What is surprising is the general lack of emphasis on a parallel process for the assessment of instructional outcomes—in other words, criterion-referenced testing.

All designers of instruction acknowledge the need for appropriate testing strategies, and non-instructional interventions also frequently require the assessment of human competence whether in the interest of needs assessment, or the formation of effective work teams, or the evaluation of the intervention.

Most training professionals have taken at least one intensive course in the design of instruction, but most have never had similar training in the development of criterion-referenced tests (CRTs). These tests compare persons against a standard of competence, unlike norm-referenced tests (NRTs) that compare persons against other test-takers. It is no longer acceptable simply to write test items without regard to a defensible process. Specific knowledge of the strengths and limitations of both criterion-referenced and norm-referenced testing is required to address the information needs of the world today.

A BRIEF SUMMARY OF TEST THEORY

Suppose you had to take an important test. In fact, this test was so important that you had studied intensively for five weeks. Suppose then that when you went to take the test, the temperature in the room was forty-five degrees. After twenty minutes, all you could think of was getting out of the room, never mind taking the test. On the other hand, suppose you had to take a test for which you never studied. By chance a friend dropped by the morning of the test and showed you the answer key. In either situation, the score you receive on the test probably doesn't accurately reflect what you actually know. In the first instance, you may have known more than the test score showed, but the environment was so uncomfortable that you couldn't attend to the test. In the second instance, you probably knew less than the test score showed due now to another type of ''environmental'' influence. Accurate measurement of what you truly knew could also have been deflected by a detectable pattern in placement of correct answers on a multiple-choice test (your score would go up) or the onset of a nasty sinus infection (your score would go down).

In all of these instances, the score you received on the test (your observed score) was a combination of what you really knew (your true score) and those factors that modified your true score (error). The relationship of these score components is the basis for all test theory and is usually expressed by a simple equation:

$$X_o = X_t + X_e$$

Here X_o is the observed score, X_t the true score, and X_e the error component. It is very important to remember that in test theory, ''error'' doesn't mean a wrong answer. It means score distortion resulting from any factor that causes a mismatch between a test-taker's actual level of knowledge (the true score) and the test score the person receives (the observed score). Error can make a score higher (as we saw when your friend dropped by) or lower (when it got too cold to concentrate). Error also results from flaws in the test itself (poorly constructed, misleading items or items that inappropriately cue test-takers regarding the correct responses).

The primary purpose of a systematic approach to test development and administration is to reduce the error component so that the observed score and the true score are as nearly identical as possible. All the procedures we will discuss and recommend in this chapter will be tied to a simple assumption: the primary purpose of test development is the reduction of error. We think of the results of test development like this:

$$X_o = X_t + X_e$$

where error has been reduced to the lowest possible level. Realistically, there will always be some error in a test score, but careful attention to the principles of test development and administration will help reduce the error component.

RELIABILITY AND VALIDITY: A PRIMER

Reliability and validity are the two most important characteristics of a test. Later on, we will explore these topics and provide you with specific statistical techniques for determining these qualities in your tests. For now, we want to provide an overview so that you will see how these ideas serve as standards for our attempts to reduce error in testing.

Reliability

Reliability is the consistency of test scores. There is no such thing as validity without reliability, so we want to begin with this idea. There are three kinds of reliability that are typically considered in CRT construction:

- Equivalence reliability
- Test-retest reliability
- Inter-rater reliability

Equivalence Reliability. Essential for parallel tests, equivalence reliability is consistency of test scores between or among different forms of the same test. Equivalence reliability is a measure of the extent to which the same test-takers get approximately the same scores on Form B of the test as they did on Form A. Forms composed of different test questions that measure the same competencies and yield approximately the same scores are said to be ''parallel.'' If each of your test-takers has the same score on Form B as he or she had on Form A, then you have perfect reliability. If there is little relationship between the test scores on the two forms, then you have a reliability estimate near zero. There are several reasons why parallel forms of a test might be desirable, for example, pretest/posttest comparisons or as a test security measure.

Test-Retest Reliability. As the name implies, test-retest reliability is the consistency of test scores over time. In other words, did a given group of test-takers receive approximately the same scores on the second administration of the test as they did on the first (assuming no practice or instruction occurred between the two administrations and the administrations were relatively close together)? If your test-takers have the same scores the second time they take the test as they had the first, then you have perfect reliability. Again, if there is little relationship between the test scores, then you have a reliability estimate near zero.

Inter-Rater Reliability. This type of reliability is the measure of consistency among judges' ratings of a performance. If you have determined that a performance test is required, then you need to be sure that your judges (raters) are consistent in their assessments. In Olympic competition, we expect that the judges' scores should not deviate significantly from one another. The degree to which they agree is the measure of inter-rater reliability. This agreement will also vary between perfect (1.0) and zero.

Validity

Validity has to do with whether or not a test measures what it is supposed to measure. A test can be consistent (reliable) but measure the wrong thing. For example, assume that we have designed a course to teach employees how to install a new telephone switchboard. We could devise an end-of-course test that asks learners to list all the steps for installing the new equipment. We might find that the learners can consistently list these steps, but that they can't install the switchboard, which was the intended goal of the course. Hence, our test is reliable, but not a valid measure for the installation task.

As mentioned above, test reliability is a necessary but not sufficient condition for test validity. Establishing reliability assures consistency; establishing validity assures that the test consistently measures what it is supposed to measure.

And while there are several measures of reliability, it is more important as you begin the CRTD process that you have a basic understanding of four types of validity:

- Face validity
- Content validity
- Concurrent validity
- Predictive validity

Of these four, only the latter three are typically assessed formally.

Face Validity. The concept of face validity is best understood from the perspective of the test-taker. A test has face validity if it *appears* to test-takers to measure what it is supposed to measure. For the purposes of defining face validity, the test-takers are not assumed to be content experts. The legitimate purpose of face validity is to win acceptance of the test among test-takers and stakeholders. This acceptance is not an unimportant consideration, especially among tests with significant and highly visible consequences for the test-taker. Test-takers who do not do well on tests that lack face validity may be more litigation prone than if the test appeared more valid.

In reality, criterion-referenced tests developed in accordance with the guidelines suggested in this chapter are not likely to lack face validity. If the objectives for the test are taken from the job or task analysis, and if the test items are then written to maximize their fidelity with the objectives, the test will almost surely have strong face validity. Norm-referenced tests that use test items selected primarily for their ability to separate test-takers rather than items grounded in competency statements are much more likely to have face validity problems.

It is important to note that, while face validity is a desirable test quality, verifying its presence is not adequate to establish the test's true ability to measure what it is intended to measure. The other three types of validity are more substantive for that purpose.

Content Validity. A test possesses content validity when a group of recognized content experts or subject-matter experts has verified that the test measures what it is supposed to measure. Note the distinction between face validity and content validity; content validity is formally determined and reflects the judgments of experts in the content or competencies assessed by the test, whereas face validity is an impression of the test held among non-experts. *Content validity is the cornerstone of the CRTD process and is probably the most important form of validity in a legal defense.* Content validity is not determined through statistical procedures but through logical analysis of the job requirements and the direct mapping of those skills to a test.

Concurrent Validity. Concurrent validity refers to the ability of a test to correctly classify masters and non-masters. This is, of course, what you *hope* every criterion-referenced test will do; however, face validation and even content validation do not actually demonstrate the test's ability to classify correctly. Concurrent validation is the technical process that allows you to evaluate the test's ability to distinguish between masters and non-masters of the assessed competencies. The most common way to accomplish this is to ask subject-matter experts to identify known masters and non-masters. The test is then administered to each group, and a statistic is calculated to determine the extent to which the test can distinguish between these performers of known competence. Concurrent validity procedures are often difficult to apply in the corporate world, though we have seen them used relatively easily in the right circumstances.

Predictive Validity. Whereas concurrent validity means that a test can correctly classify test-takers of currently known competence, predictive validity means that a test can accurately predict future competence of test-takers. Predictive validity is important for many personnel selection devices that are used to choose persons for specific job responsibilities. Tests used to help persons select careers also require high predictive validity. In both of these cases, the test is taken first, while the demonstration of competence—job performance or successful career achievement—comes later; hence the term predictive validity. Test scores and measures of future performance are correlated to establish predictive validity.

CRITERION-REFERENCED TEST INTERPRETATION

In contrast to NRTs, a criterion-referenced test (CRT) defines the performance of each test-taker without regard to the performance of others. Unlike the NRT, whereby success is defined in terms of being ahead of someone else, the CRT interpretation defines success as being able to perform a specific task or set of competencies. There is no limit to the number of people who can succeed on a criterion-referenced test, unlike the NRT.

Criterion-referenced tests should be used whenever you are concerned with assessing a person's ability to demonstrate a specific skill. The medical boards licensing exam is an example of a criterion-referenced test. If you are being operated on, you should want to know that your surgeon is competent to perform the operation, not just that he or she is better than 90 percent of those who graduated. The reason is that merely knowing more than the others in the class does not guarantee that your surgeon can perform the operation; maybe nobody in the class mastered the operation. The danger of NRTs in corporate training situations is that without reference to specific competencies, what test-takers can actually do is unverifiable.

INSTRUCTIONAL DESIGN AND TEST DEVELOPMENT

The systematic design of instruction is a process that is no more complex (nor simple) than the systematic design of tests. The practice of testing is easily within the grasp of professionals who are conversant with instructional design processes. The CRTD process is not about complex statistical calculations, but rather about analytical thinking. Good tests can be developed in a nearly seamless manner along with the development of instruction. Good, objectives-based testing takes very little additional time and resources when you understand how it is done. In many instances, training professionals fear that test creation may double the development process time and costs; it need not. Our experience has been that good, basic test development can be integrated into the ID process and may add no more than 10 percent to the timelines and resource requirements. Of course, other factors such as legal concerns or computer-based testing strategies will add to the cost and complexity of a testing project, though the return attributed to such factors may easily be justified in some circumstances.

For all the reasons we discussed earlier, many organizations are now turning to testing as a means of identifying and verifying competence. In the past, organizations may have been willing to assume that their tests were reliable and valid when, in reality, they were not. For professional and legal reasons, organizations are now rethinking these assumptions. Where once companies could accept token attempts at assessment, they now look for professional practice in testing. Simply writing tests that look like the tests everyone had in school is not a substitute for designing and developing tests that match real jobs. In this chapter we will provide an overview of the model we have synthesized and have seen used in developing tests that are linked to job content and performance, as well as to a curriculum plan.

A CRT DEVELOPMENT MODEL

There is a systematic, although rarely described, process for designing certification tests. We have diagramed the major steps that are often used in the process (Shrock & Coscarelli, 2007, p. 45), and in the rest of this chapter we will provide an overview of each step. Figure 15.1 shows these steps in suggested chronological order.

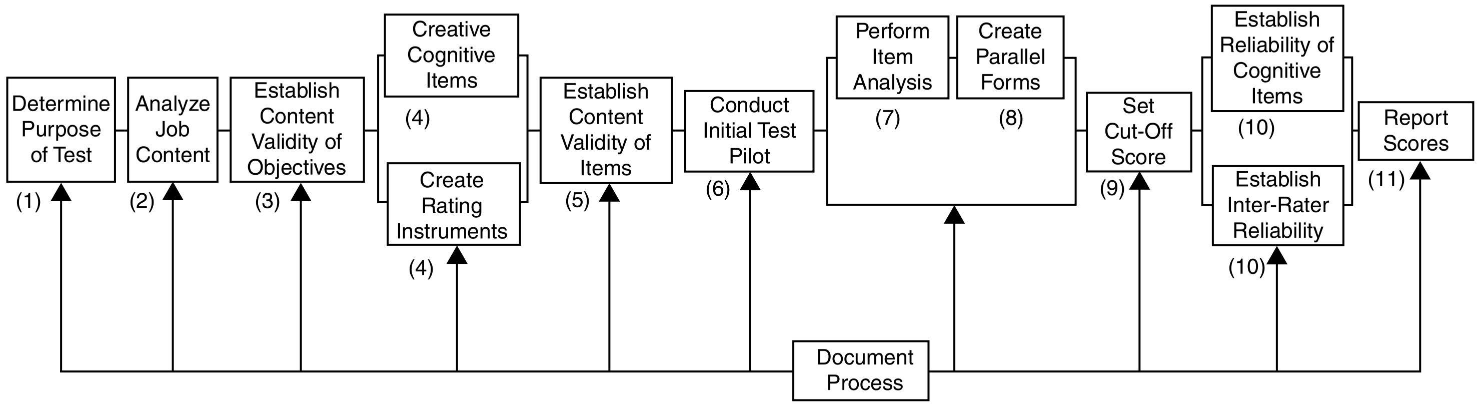


Figure 15.1 A Model for Developing Criterion-Referenced Tests.

From Shrock and Coscarelli, 2007, p. 45

As you review the figure, you will see that the steps that are common to both cognitive and performance tests are in the middle row of the model. The steps on the top address issues primarily associated with cognitive testing. The steps immediately below are used primarily for creating performance tests. And the role of documenting the test development process is shown to apply to all steps throughout the model. The parenthetical numbers at each step refer to the corresponding documentation questions posed in Exhibit 15.1.

Exhibit 15.1 Test Development Documentation Guidelines

1. **Purpose of the Test**

   What is the need for the test?

   How was the need determined?

   Who sponsored the effort? Why were they the sponsor(s)?

2. **Analyze Job Content**

   What job/duty/task is being analyzed?

   If a duty or a task, of what job(s) is it a part?

   What is the hierarchical relationship of job/duty/task elements?

   Include a copy of the analysis.

   Who performed the job analysis?

3. **Establish Content Validity of Objectives**

   For what job/duty/task were the objectives derived?

   In what course(s) are they covered?

   What are the names, titles, and credentials of the subject-matter experts who va1idated the objectives?

   When was the validation performed?

   Include a copy of the objectives and the job analysis.

4. **Create Cognitive Items/Rating Instruments**

   For what job and objectives were the items/rating instruments created?

   For what course is the test intended?

   Who wrote the items/rating instruments?

   What are the titles and credentials of those who wrote the items/rating instruments?

   When were the items/rating instruments written?

   Include a copy of the items/rating instruments.

(*Continued*)

Exhibit 15.1 (*Continued*)

5. **Establish Content Validity of Items and Instruments**

   For what job and objectives were the items/rating instruments created?

   For what course is the test intended?

   Who validated the items/rating instruments?

   What are the titles and credentials of those who validated the items/rating instruments?

   When were the items/rating instruments validated?

   Include a copy of the form on which subject-matter experts indicated the match between the items/rating instrument and the objectives or the job elements.

6. **Conduct Initial Pilot Test**

   When was the test pilot conducted?

   Where was the pilot conducted?

   Who conducted the pilot?

   Who were the sample test-takers?

   How were the sample test-takers chosen?

   What were the pertinent characteristics of the sample test-takers?

   What was the process used to conduct the pilot?

   What changes were made to the test as a result of the pilot?

7. **Perform Item Analysis**

   Who were the test-takers whose test data were used in the item analysis?

   What were the relevant characteristics of this sample of test-takers?

   How were the sample test-takers chosen?

   When were the item analysis data collected? By whom? Where?

   What program was used to analyze the test data?

   What were the results of the item analysis?

   What changes were made to the test as a result of the item analysis?

   Include a copy of the item analysis printout.

8. **Create Parallel Forms and Items Banks**

   For what job/objectives were the parallel forms or itembanks created?

   How many parallel forms were created? When were they created?

How was a decision made about the number of parallel forms to be created? Who made this decision?

What software was used to create the item banks? Who chose the software and why was it chosen?

What are the sizes and structure of the test item banks?

How are items selected from the banks?

What process was used to create the parallel form?

What is the equivalence reliability of the parallel forms? How was it established?

9. **Set Cut-off Scores**

How is the test scored?

What is the cut-off score for the test?

When was the cut-off score determined?

What procedure(s) was (were) used to set the cut-off score?

Include all data pertinent to the procedures(s) used, including a description of test-takers and how they were chosen; judges and their credentials; item difficulty estimates; stakeholders and their suggested cut-off scores.

Who decided what the cut-off score would be? What are their titles and credentials for making this decision?

10. **Establish Reliability of Cognitive/Performance Tests**

What procedures were used to establish the reliability of the test/raters?

When was the data collected upon which the reliability was calculated?

What is the reliability of the test/raters?

Include nil pertinent data used in the calculation of the reliability coefficient.

11. **Report Scores**

To whom are scores reported?

In what form are the scores distributed?

Are the scores accompanied by course means? medians? other descriptive statistics?

Are reported scores composites of subscores? If so, what are they and are they differentially weighted? If so, how? How were the weights determined?

What guidance regarding the use of the scores is provided with them?

Source: Adapted from Shrock and Coscarelli, 2007, pp. 65–67

PLAN DOCUMENTATION

Patricia Eyres (1997, pp. 6–19) lists ten of the most common legal problems for trainers, consultants, and speakers. Of these ten, number nine, ''you find your documentation is ineffective or nonexistent'' (p. 17), is particularly important to the test developer. As she observes:

> ''In any legal dispute . . . the decision maker must make several determinations: what happened, who was injured or damaged, how it occurred, who is responsible, and what penalties or damages should be assessed. As in most conflicts, memories fail, recollections differ, and stories definitely change. To reach a reasonable conclusion on the issues, the agency representative or civil jurors must ultimately determine whose recollections are most credible. The existence of effective, accurate, and consistent documentation helps immeasurably. Conversely, the absence of documentation can destroy even the best technical defense.
>
> ''Documentation is a written record of an event, discussion, or observation by one or more individuals. Any written information, whether formally or informally generated, can be considered documentary evidence if it is pertinent to a legal action or regulatory proceeding. Why is documentation so important? Simply because a written record of events is the best evidence of what actually occurred . . . a common thread throughout the reported cases in all aspects of training liabilities reveals that the absence of documentation hinders an effective defense, *even when the facts would otherwise support your position.*'' (Eyres, 1997, pp. 17–18)

To that, we would add these words from *The Uniform Guidelines on Employee Selection Procedures* (1979), ''Validation studies begun on the eve of litigation have seldom been found to be adequate'' (p. 12002).

TEST DEVELOPMENT DOCUMENTATION

If you follow the model described in this chapter, you will be on the path toward developing a valid test; if you document the process by answering the questions we pose for each part of the model, you will be on the path toward developing a defensible test. Exhibit 15.1 comprises a list of basic questions you should consider addressing as you document the development of your test. Archived, contemporaneous memos are often adequate for documentation purposes. (*Note:* We are not attorneys and do not claim these questions to be legal requirements. Once again, consult with your organization's legal staff if you have any concerns.)

ANALYZE JOB CONTENT

An analysis of the content to be tested, or more significantly, the content of a job to be assessed by a test, is an absolutely critical foundation to the testing process. In the American Educational Research Association's (AERA) *Standards for Educational and Psychological Testing* (1999), Standard 14.10 provides the fundamental underpinning to test development in the corporate setting.

> ''When evidence of validity based on test content is presented, the rationale for defining and describing a specific job content domain in a particular way (e.g., in terms of tasks to be performed or knowledge, skills, abilities or other personal characteristics) should be stated clearly.'' (p. 160)

If you are creating a test that is not linked specifically to a job, but to a content domain for purposes of certification, the need for careful analysis of skills is the same. Standard 14.14 and its associated comment summarize this concern.

> ''The content domain to be covered by a credentialing test should be defined clearly and justified in terms of the importance of the content for credential-worthy performance in an occupation or profession. A rationale should be provided to support a claim that the knowledge or skills being assessed are required for credential-worthy performance in an occupation and are consistent with the purpose for which the licensing or certification program was instituted. *Comment*: Some form of job or practice analysis provides the primary basis for defining the content domain.'' (p. 161)

The *Uniform Guidelines on Employee Selection* (1978), which have become the primary standard for adjudication of testing issues in the courts, state in part:

> ''There should be a job analysis which includes an analysis of the important work behavior(s) required for successful performance and their relative importance and, if the behavior results in work product(s), an analysis of the work product(s). Any job analysis should focus on the work behavior(s) and the tasks associated with them. If work behavior(s) are not observable, the job analysis should identify those aspects of the behavior(s) that can be observed and the observed work products. The work behavior(s) selected for measurement should be critical work behavior(s) and/or important work behavior(s) constituting most of the job.'' (p. 38302)

The courts' response to inadequate job analysis is typified by the opinion in *Kirkland v. Department of Correctional Services* (cited in Thompson & Thompson, 1982):

> ''The cornerstone in the construction of a content valid examination is the job analysis. Without such an analysis to single out the critical knowledge, skills and abilities required by the job, their importance relative to each other, and the level of proficiency demanded as to each attribute, a test constructor is aiming in the dark and can only hope to achieve job relatedness by blind luck.'' (p. 867)

As you can see, there are compelling professional and legal reasons for attending to a careful task analysis of the job or content to be assessed.

ESTABLISH CONTENT VALIDITY OF OBJECTIVES

Establishing the content validity of your objectives appears to be a relatively straightforward process: you've identified the job-task skills; now it should be a simple matter to create objectives that match the skills. So, for example, if the task requires the performer to weld two metal bars at a 90-degree angle, the objective might be: ''Given an arc welder, rods, and two iron bars, weld the bars at a 90-degree angle with no gaps in the weld.'' There should be a close match between the objective and the task (and later, the test item). However, in the corporate world, such seemingly simple prescriptions have a way of becoming harder to fulfill as the demands of technology, time, costs, law, health, and safety issues enter into the planning process. The ''real world'' sometimes forces compromises from this ideal. If for any reason you decide *not* to match the job to the objective and subsequently to the test item, you need to be aware of the consequences of teaching and testing in a way that does not reflect the job—and the consequences are serious! If you find yourself working with the company's traditionally trained psychometricians (who are most likely to be found in the HR department), you need to be aware that their training and tendency is to create tests that are designed to sort test-takers (an NRT assumption) as opposed to tests that measure test-takers against a standard of performance (a CRT assumption). The traditional psychometrician will be interested in sampling items from domains of content that lead to separating test-takers *regardless* of the items' links to the job task analysis. Any time you are asked to develop a test that is not linked directly to the test-takers' job via items that are directly reflected in the corresponding job task analysis—but a test that will have significant consequences such as safety, ''hire-fire,'' or ''promote''—is a time to talk to the legal department.

The Role of Objectives in Item Writing

It is difficult to overstate the usefulness of good instructional objectives in the creation of sound tests (Mager, 1962; Popham, 1978). Most instructional designers are aware of how important objectives are to the creation of instruction; many are less familiar with the role of objectives in testing.

Instructional objectives serve three fundamental purposes for criterion-referenced test developers:

- Objectives ensure that a test covers those learner outcomes important for the purposes that the test must serve.
- Objectives increase the accuracy with which cognitive processes in particular can be assessed.
- The size of the domain covered by the objectives and the homogeneity of the objectives the test is designed to assess are important factors in determining how many items need to be included on the test.

CREATE COGNITIVE ITEMS

The answer to the question ''What are cognitive items?'' used to be easily understood by test designers. ''They are the paper-and-pencil items,'' most would reply. The traditional distinction was that you assessed intellectual skills with ''paper-and-pencil'' items, and you used performance scales to observe a behavior or to judge a product. As computer-based technology evolved, the distinction between the two domains blurred in many areas.

The first wave of change was simply to move the paper-and-pencil items to the computer screen as a delivery strategy. However, as computer processing power evolved, it became possible to create items that could simulate the job in a manner that ''thinking about'' the task and ''doing it'' were indistinguishable; for example, the simulated configuration for installing a local area network and the reality of doing it were essentially the same. In some fields the distinction between ''thinking'' and ''doing'' is not easy to reconcile; for example, listing the steps to install an electrical amplifier is not the same as installing it twenty-five feet above ground. However, in an increasing number of information-age jobs, you may find it useful to consider how the computer can be used not just as a delivery vehicle but rather as a nearly perfect simulation of real-world performance.

Classic Types of Test Items

The classic notion about a test item is that it poses a question and either provides possible answers or distractors among which the test-taker chooses the best answer (called closed-ended items) or allows for a free-form response (open-ended items). Guidelines for using closed-ended and open-ended items are:

- *Closed-ended questions* are often used when you are looking for a response that would have a predictable correct answer. Common item types are multiple-choice, true/false, and matching. Well-written multiple-choice items minimize successful guessing. Closed-ended items are easily scored by a computer and will provide the best chance of establishing an acceptable level of reliability for a test.
- *Open-ended questions* are used when you are assessing writing ability or soliciting a response that is not easily classified as right or wrong. Common types of open-ended items are essay, short answer, fill-in, and numeric. Open-ended questions are difficult for humans to score reliably, but even more difficult to score using computers. Establishing test reliability and validity for tests composed of open-ended items typically requires greater effort on the part of the test designer.

It is important to note that neither the closed-ended nor the open-ended question format is inherently more valid than the other. Test writers are advised to choose a format powerful enough to capture the conditions and verb in the corresponding objective or competency statement, keeping in mind that the scoring of open-ended items presents reliability problems.

Newer Computer-Based Item Types

Computer-based testing has had the effect of increasing the number of test item types and response modes now available to test designers. Most of them are closed-ended in nature. There are some advantages to these new item formats, although it has been our experience that many of them are used primarily for memory-level assessment. Some of these newer item formats actually facilitate the assessment of higher cognitive levels; the potential of this technology for testing will not be realized until test developers become more aware of and skilled at writing items above memory level. Four of the newer item formats are

- *Animated.* Using software technology such as Macromedia Flash, an item can be created that introduces motion as part of the posed question.
- *Drag and Drop.* A list of options is presented to test-takers, and they use their computer mouse to select an option and move it to an appropriate place on the screen.
- *Hotspot.* The test-taker is directed to answer a question by clicking on a place on the screen or touching the screen, depending on the type of screen.
- *Pull Down.* A selection of options is presented by a mouse click on a chosen field. Caution should be used in introducing novel item types or response modes because confusion regarding how to respond to an item would certainly introduce error into test scores. Be sure test-takers are thoroughly comfortable with an item format before employing it in a test.

The Six Most Common Item Types

There are six types of test items most commonly used in cognitive tests. These item types are

- True/False
- Matching
- Multiple-Choice
- Fill-In
- Short Answer
- Essay

Many measurement books provide guidelines for writing items of different types (Burton, Sudweeks, Merrill, & Wood, 1991; Case & Swanson, 2001; Shrock & Coscarelli, 2007). Test writers should be well versed in these general guidelines because following them can greatly reduce the error associated with items that cue correct answers, confuse test-takers, facilitate correct guessing, and/or introduce unreliability into the scoring of the test. Because advice about the mechanics of writing test questions is so readily available, this chapter concentrates on supporting a comprehensive test creation process and those elements of test development frequently overlooked.

The Key to Writing Items That Match Jobs

The vast majority of test items that are created are written at the memorization level. In contrast, the vast majority of jobs require performance above the memorization level. This disconnect between testing practice and job performance is what often leads management to question the value of training and turns testing into a misleading indicator of performance. Furthermore, memorization-level tests encourage weak, irrelevant instruction rather than challenge training to rise to the level the job requires. ''How come they passed the course but can't do the job?'' is a common summary of the problem. Matching the job to the objective to the test item seems to be common sense among trainers, but the execution of this idea has proven harder to accomplish than we first expected. Most test writers create tests that look like the tests that they took in school. Because the vast majority of these tests measured only memorization and used conventional paper-and-pencil formats, this habit is tenacious and pervasive.

In order to write tests that match or approach the knowledge and skill levels of jobs, test writers must be able to write items that require cognition more sophisticated than memorization and place the test-taker under the stimulation conditions typical of job performance. We are not suggesting that memory is unimportant to learning; in fact, it is essential. However, we are suggesting that testing job-related competencies that require higher cognitive levels will also capture simultaneously the prerequisite remembered content as well.

When you design your test, first consider the job, and then consider the level of learning your test assesses in light of this job performance. In Bloom's (1956) terms, design your test items above the ''knowledge'' level; in Gagné's (1985) language, above the ''verbal information'' level. It is usually not productive to worry about precise classification of items beyond memory level; the critical distinction is between memorization and everything else above it. Let the job drive the cognitive level of the test items and their corresponding objectives. It is essential that you distinguish between the tasks that require simple memorization as opposed to those requiring higher cognitive levels because test items that measure above the memory level will require previously unencountered examples, scenarios, or simulations that show what the performer must face on the job.

How Many Items Should Be on a Test

There comes a point in the test-planning process when developers need to decide how many items will appear on the test. This is a question that, unfortunately, does not have a simple, numerical answer. It is an extremely important question, however, because the length of the test has a direct relationship with the test's reliability and, therefore, with its validity as well. In general, longer tests are more reliable than shorter tests. The question of test length for any specific test turns on at least four factors:

- Criticality of the mastery decisions made on the basis of the test;
- Resources (time and money) available for testing;
- Domain size described by the objectives to be assessed; and
- Homogeneity or relatedness of the objectives to be assessed.

Criticality of Decisions and Test Length. We know that test reliability is a function partly of test length. Therefore, when trying to decide how many items to put on a test, it makes sense to ask the question, ''How reliable does the test have to be?'' Sometimes errors in master/non-master classification of test-takers can be tolerated. It is very useful to do a systematic analysis of what the consequences are of both types of errors that can be made by unreliable criterion-referenced tests. Ask yourself and others who are knowledgeable about the responsibilities of the target test-takers, ''What are the costs to this organization of erroneously classifying a non-master as a master?'' (This error is sometimes called an error of acceptance or a false positive error.) Undeserved bonuses? Poor work performance? Lawsuits from clients? Deaths? And ''What are the costs to the organization of erroneously classifying a master as a non-master?'' (This error is sometimes called an error of rejection or a false negative error.) Denial of deserved bonuses? Demoralized employees? Lost talent from the organization? Lawsuits from employees?

The point here is that, to the extent that errors in classification can be tolerated, tests can be shorter. However, if the consequences of classification errors are severe, the tests used to make master/non-master decisions will have to be longer as well as meet other conditions required for reliable and valid tests. Interestingly, norm-referenced tests can usually reach acceptable reliability with fewer items than criterion-referenced tests can. This result is likely because criterion-referenced tests typically include items that cover all important competencies relevant to a particular job or task, even those test items that most test-takers answer correctly. Test items like those do not separate test-takers' scores from one another; in fact, they restrict the variation in scores. Any restriction in the range of the observed scores will depress the reliability estimate calculated for a test. This characteristic of CRTs makes it particularly important that their developers write items of excellent quality. Because CRTs tend to be long anyway in order to provide adequate assessment of necessary competencies, it is essential that test-taker time and organization resources not be wasted on low quality items that don't require genuine test-taker competence to answer correctly.

Resources and Test Length. It will come as no surprise that the creation of tests takes time and, therefore, costs money. The longer the test, the greater the development costs and the greater the cost of compensating test-takers during assessment. There are also costs, of course, associated with maintaining and scoring tests. Test designers are perhaps less inclined to realize that tests also incur other costs to the organization—some dollar costs and others in the form of what are called ''opportunity costs.'' Chief among these latter costs is instruction time lost to test-taking time.

Domain Size of Objectives and Test Length. The number of items required for a test is also influenced by the size of the content domain represented by the objectives that the test is designed to assess. (*Note:* In the NRT world, the term ''domain'' is often linked to a test construction strategy that requires sampling from within some prescribed parameters defining a knowledge base, for example, adding all two two-digit numbers that require carrying.) We are using the term ''domain'' in CRT to describe the breadth of tasks within the job task hierarchy wherein we seek to assess all knowledge within the domain to determine competence—rather than to sample and then to infer competence.

In general, the smaller the domain of content described by the objective, the fewer the items required to assess the objective adequately. For example, consider an objective such as, ''Without assistance, list the six steps required to make a milk shake using the Presto-Malt machine.'' This objective describes a small content domain; in fact, it is difficult to imagine how one could write more than one item to assess this objective. Most objectives, however, require more than one item—parallel items—to assess them adequately.

For example, consider the following objective: ''Given pertinent data and access to all essential technical manuals, diagnose the source of radiation leak in a nuclear reactor.'' This objective describes a far larger content domain and, as you can imagine, would require far more items to instill confidence that it had been adequately assessed. Objectives that describe behaviors that must be performed under several different conditions on the job should be assessed by several items reflecting those different conditions. This discussion should make clear why specific objectives are so important to test creation. It is very difficult to decide the issue of test length if the objectives are ambiguous.

Homogeneity of Objectives and Test Length. Another characteristic of the assessed objectives that influences test length is the homogeneity of the objectives, that is, their relatedness to one another. Consider these two objectives: (1) ''Without access to references, describe the steps in conducting a performance appraisal'' and (2) ''Without access to references, describe the four stages of interpersonal confrontation.'' These two objectives are related in that the content they cover is similar. In fact, the second objective is very likely a prerequisite objective to the first. As a result, test-takers are likely to perform the same way on the test items written for these two objectives; in more technical language, responses to items covering these two objectives will be positively correlated. If objectives are homogeneous to the extent that they result in test items to which test-takers respond similarly, fewer items need be included to assess each objective independently.

It is important to realize that it is very frequently difficult to tell simply by looking at objectives whether or not responses to their corresponding test items will be similar. This conclusion can only be drawn after actual test results are available, and you can determine for certain how similar the responses were. You may be able to reduce the numbers of items included for each objective on a test if you can confirm sufficient homogeneity of the underlying objectives. On the other hand, if the objectives covered by the test are largely unrelated—heterogeneous—you can expect that the test will have to be considerably longer since several items will probably be required for each objective.

Research on Test Length. Research into the accuracy of assessments as a function of numbers of test items per objective indicates that more items result in greater test reliability. However, the improvement in reliability tends to level off, generally between four and six items per objective. It should be noted that such research does not address the necessary levels of reliability essential for the safe or proper assessment of objectives of varying *criticality*. The accuracy achieved with four to six items may not be good enough for some critical objectives. In other words, the assessment of objectives that describe behaviors essential to health, legal requirements, and organizational survival should be assessed by more than six items, possibly by as many as twenty items, and may need to be assessed several times, especially if the content domain of these critical objectives is large.

Multiple assessments are more frequently used in performance testing than in cognitive testing; however, for some essential skills, multiple cognitive assessments may be appropriate. Often when multiple assessments are used, the standard for passing is extremely high; frequently no errors are allowed.

Table 15.1 presents a heuristic decision table relating the criticality, domain size, and relatedness of objectives to approximate numbers of necessary test items.

Summary of Determinants of Test Length

The number of items that should be included on a test is primarily a function of the criticality of the master/non-master classifications that will be made based upon the test results. This consideration is prime because test length is directly related to test reliability. The more costly the consequences of classification errors, the longer the test should be. Time and money, of course, are always limiting factors. Objectives that specify small content domains and that are

Table 15.1 Decision Table for Estimating the Number of Items Per Objective to Be Included on a Test

If the Objectives Are:	Critical	From a Large Domain	Unrelated	10–20
			Related	10
		From a Small Domain	Unrelated	5–10
			Related	5
	Not Critical	From a Large Domain	Unrelated	6
			Related	4
		From a Small Domain	Unrelated	2
			Related	1

Source: From Shrock and Coscarelli, 2007, p. 171

correlated with other objectives require fewer items than those that describe large content domains and are essentially unrelated to other objectives covered by the test. Research suggests that the balance between effectiveness and efficiency in item numbers is achieved at four to six items per objective, but we know that more items will be required for some critical objectives.

CREATE RATING INSTRUMENTS

Performance tests seek to provide an objective rating of either a behavior or a product. Using them can help an organization in a number of ways. Valid performance tests:

- Provide an objective and reliable measure of the trainees' actual ability to perform a task, distinguishing those who can meet the standards from those who cannot. At the same time, they allow trainees an additional opportunity for practice.
- Provide an observable standard for performance against which all trainees can be evaluated consistently.
- Reveal whether a trainee can deal with the stress and pressure of task performance under actual or closely simulated work conditions.
- Indicate whether the instructional program is successful in producing workers whose performance meets job requirements.
- Provide authoritative information on the maintenance of quality instruction and program effectiveness. (Campbell & Hatcher, 1989, p. 2)

It is not a conceptually difficult process to develop good performance tests. A valid performance test is based on a detailed and thorough analysis of the skills required for the behavior or the desired characteristics of the product, or both. While creating performance tests and establishing their validity are often a straightforward part of the test development process, establishing the scoring and the reliability of the raters who will use the observation instruments is usually the real challenge.

Product Versus Process in Performance Testing

There is an essential distinction in performance testing: assessing the outcome of a procedure or process—the product—or assessing the way in which the outcome was achieved—the process. You may need to emphasize one aspect over the other or consider some combination of both as the job dictates. (It is worth noting that process assessment usually requires ''real time'' observation, and, therefore, is often more expensive than product assessment.)

- The nature of the performance frequently dictates where the emphasis should be placed. Some types of performance do not result in a tangible product. Activities such as these require that the performance be evaluated in progress, special attention being paid to the constituent movements and their proper spacing.
- In some areas of performance, the product is the focus of attention and the procedure (process) is of little or no significance. Judging the quality of the product is typically guided by specific criteria that have been prepared especially for that purpose.
- In many cases both procedure and product are important aspects of a performance. For example, skill in locating and correcting a malfunction in a television set involves following a systematic procedure (rather than using trial and error) in addition to producing a properly repaired set. (Gronlund, 1988, pp. 86–87)

Again, how you manage the combination of process and product will be determined by the demands of the job.

Types of Rating Scales for Use in Performance Tests

Once the behavior (or final product) has been analyzed to define the essential characteristics of worthy performance, the next step is the creation of a rating scale to support a final evaluation. There are basically four types of rating scales, only two of which should be used:

- Numerical scale
- Descriptive scale
- Behaviorally anchored rating scale
- Checklist

Of these four, we do not recommend the use of numerical or descriptive scales. Numerical scales provide a continuum of numbers, for example, 1 to 5, undefined except for the high (best) and low (worst) ends of the scale. Descriptive scales provide a continuum of evaluative words such as ''good,'' ''fair,'' ''poor,'' etc. Both of these scales allow for too much rater subjectivity. Both behaviorally anchored rating scales and checklists are acceptable approaches to assessing a skill or product, but of these two, the checklist is generally more reliable.

Behaviorally Anchored Rating Scales. A behaviorally anchored rating scale (sometimes called BAR scales) uses both words and numbers to define levels of performance. However, the words that are used are not vague value labels, but terms that describe specific behaviors or characteristics that indicate the quality of the performance or the product. The use of specific descriptions tends to make these scales more reliable than the unanchored numerical or descriptive scales. Figure 15.2 provides an example of this type of scale. As you can see, the more specific the behavior interpretation, the more reliable the scale will be.

One issue that often arises with the use of these scales is, ''How many points should there be on a scale?'' While the selection of points is tied to the behaviors required for the task, research suggests that raters can reliably distinguish among five levels of performance. More than seven such points may stretch the limits of the rater's ability to quickly and accurately discriminate behaviors.

Checklist. Checklists are constructed by breaking a performance or the quality of a product into specifics, the presence or absence of which is then ''checked'' by the rater. Checklists may also have what are sometimes termed ''negative steps'' in them. These negative steps represent what should not be found, for example, ''no extraneous holes in the wall'' when evaluating hanging a picture.

Behavior	Performance	Rating
I. Response to directory assistance request	1. Curt voice tone; listener is offended	1
	2. Distant voice tone; listener feels unwelcome	2
	3. Neutral voice tone; listener is unimpressed	3
	4. Pleasant voice tone; listener feels welcome	4
	5. Warm, inviting voice tone; listener feels included	5

Figure 15.2 Example of a Behaviorally Anchored Rating Scale.
Source: From Shrock and Coscarelli, 2007, p. 171

Checklists tend to be the most reliable of all rating scales because they combine descriptions of specific behaviors or qualities with a simple yes-or-no evaluation from the rater. The checklist radically reduces the degree of subjective judgment required of the rater and thus reduces the error associated with observation. Remember, however, that while the checklist increases the reliability of the raters, a careful task analysis is required to assure the validity of the scale.

ESTABLISH CONTENT VALIDITY OF ITEMS AND RATING INSTRUMENTS

Validity means that the test measures what it is supposed to measure. It is the fundamental assumption of criterion-referenced testing that the test matches the objectives you have established. The underlying process of establishing the validity of the test is conceptually quite simple. In most training settings, the process will be one of showing a logical link between the job, the objectives of instruction, and the test items. In developing a certification test for which there may be no instruction offered by the sponsoring organization, the link will be between the job and the test. This process applies to both cognitive items and rating instruments used for performance assessment.

Why Content Validity?

Earlier in this chapter we introduced four types of validity—face, concurrent, predictive, and content. While all of the latter three can be formally determined and assess different but legitimate evidence of test validity, in the CRT creation process, you must establish the content validity of a test. Content validity is the first evidence the courts will look for in any challenge to a test. It is also the primary property of a test that indicates the test's ''job relatedness.'' Calculation of concurrent validity (the test's correct classification of previously determined masters and non-masters) or predictive validity (the test's correct determination of future job success or lack of) are regarded as optional determinations, and are typically reserved for high-stakes tests.

The Content Validation Process

A test possesses content validity when a group of recognized content experts or subject-matter experts (SMEs) has verified that the test measures what it is supposed to measure. Note the distinction between face validity and content validity; content validity is formally determined and reflects the judgments of experts in the content or competencies assessed by the test, whereas face validity is an impression of the test held among non-experts.

The steps in the process of determining content validity are described below:

1. The first step in establishing content validity is to select three to five judges who are experts in the competencies assessed by the test. If the test covers sufficiently unrelated objectives, you might have to have a panel of judges for subsets of the items. You might have to have more judges if the test covers sufficiently general objectives. For example, if the test were an assessment of management skills, you might have to have judges who could represent the major divisions of the organization—technical, operations, sales, etc.—to ensure that the test will be acceptable to managers throughout the organization. The identity of the judges and their credentials for serving as judges should be recorded for documentation purposes. This information could be important if the content validity of the test is ever challenged, especially legally.
2. The judges are presented with the objectives the test is supposed to assess and the items corresponding to each of these objectives. For each item, the judges must decide whether or not the item assesses the intended objective. We recommend asking judges to make a yes/no decision regarding whether or not the item matches the objective rather than asking them to rate the objective on a scale. This recommendation simplifies the process for the judges, improves the reliability of their judgments, and facilitates the aggregation of the judges' opinions. Judges should also be asked whether they see any technical problems with the item—any cueing of the correct answer, more than one possible correct answer, and so forth. Judges should also be provided with space to make any additional comments about the item that they think test developers ought to know.
3. It is suggested that judges review and rate the items independently first, then debrief their results together with the assistance of one of the test's writers. The test writer should be there to hear first-hand the judges' remarks and concerns; this person can also facilitate the reaching of consensus among the judges regarding the acceptability of each item.

It is important that the objectives given to judges be based on an accurate job analysis. Since judges are only matching items to the objectives presented to them, they cannot be expected to discover a faulty job analysis. If the job analysis reveals more skills than the planned test can assess, it is important that the objectives chosen for inclusion be representative of the job and that the procedure used to select the objectives be documented in the event of legal challenges to the test's validity. Establishing the content validity of the test is the single most important step you can take in developing your test. It is neither a difficult nor a time-consuming process.

CONDUCT INITIAL TEST PILOT

Just as any systematic approach to course design includes a formative evaluation or course pilot, systematic test development means piloting your test. However, when designing a test, piloting is absolutely *essential* because the detection of faulty items—either test questions or points on a rating scale—*requires* real test data. The piloting process should identify potential problems with test organization, directions, logistics, and scoring, as well as with items, and lead to their correction. Additional test data gathering will also be required in order to establish the cut-off score that defines mastery and to establish the reliability and validity of the test. *The single most important purpose in the initial piloting of the test is to gather feedback for improvement of the test, not to rate the pilot test-takers*. Remember, almost any testing situation can be personally threatening. As you conduct the test pilot, you need to be particularly supportive and emphasize that your purpose is evaluation of the test, not the test-takers. However, you should take the test pilot process very seriously if you expect to create a test with significant consequences for the test-taker. Courts are very likely to support a challenge to a test if persons are not treated equally during its administration. The test pilot is the most informative opportunity to resolve any issues of fairness before the test becomes operational.

Six Steps in the Test Pilot Process

The pilot test is a formative evaluation process that will parallel the course pilot process. You will need to: determine the sample test-takers, orient the participants, give the test, analyze the test results, interview the test-takers, and synthesize the results. You may not be able to do all of these steps as the result of a single pilot test administration. Many test developers plan two formative test pilots. The first is conducted with a small sample (perhaps ten to twenty test-takers, depending on the stakes associated with the test) in order to gather qualitative interview data. The second involves a larger sample of test-takers (as many as possible, but ideally at least sixty) in order to gather more stable item analysis data. Following these six guidelines as you pilot your test will improve its quality.

Determine the Sample. Your test pilot test-takers should mirror your intended test audience. Don't rely on a ''sample of convenience'' where you grab three people who are around the office and between projects. Nor should you be satisfied with just anyone sent to you from the field. If the pilot is to have meaning, the sample test-takers must be representative of future test-takers.

The size of the pilot test sample will depend on the scope of the test. A small sample will be useful primarily to gather qualitative reactions from the test-takers, that is, verbal comments about how the test might be improved. If you are designing a test for a limited-run workshop, then you could work with a smaller sample—even as small as three. If you are designing a high-stakes test that will be used on an ongoing, company-wide basis, then you should invest the time and resources for a full pilot, including careful examination of the item analysis statistics described below and in the next section of this chapter.

Finally, be sure to document the characteristics of the pilot test group. The sampling decisions you make for the pilot should be noted in a memo or other written form in case questions arise later about the test.

Orient the Participants. Because your goal is to evaluate the test, not the test-takers, you need to make it clear to your pilot sample that it is the test that is being evaluated. Since the test-takers are, in effect, your colleagues in the test development process, they should be so informed and treated as such.

Give the Test. When you give the test, give it *exactly* as it would be given in the field. This means you should give the directions verbatim, adhere to the time limits, and avoid any hints, apologies, or interpretations of the test or any of its items. Any intervention on your part during the pilot may jeopardize your understanding of how the test will work later in the field. Smaller errors such as typos in the test should be corrected just as you would during a field administration. Gross errors may require immediate modifications to the test in order to allow the pilot test-takers to proceed. In either instance, it is important that you document your changes and the reasons for the changes.

While the test is in progress, you should take careful notes to document what the test-takers are doing. Watch for nonverbal cues such as head scratching or frowning that might indicate anger, confusion, frustration, and so on.

Analyze the Test. In an ideal setting right after you administer the test, you would complete the statistical analysis of the results. The test item analysis process allows you to identify any items that might be a problem; for example, if nobody selected three of the five distractors on an item, the question is effectively a binary-choice item (presenting a fifty-fifty chance of correctly guessing the answer), rather than a true multiple-choice item (with a 20 percent chance of correctly guessing the answer). Briefly, the three major statistics most commonly employed in an item analysis are:

- *Difficulty index.* This is simply a report of the percent of test-takers who answered an item correctly.
- *Distractor pattern.* This statistic is a report of the number of test-takers who selected each alternative option for each test item.
- *Point-biserial correlation.* A more sophisticated technique, the point-biserial correlation requires computer support. It is, however, a very powerful tool that easily allows you to identify items that test-takers with

the highest scores consistently missed while low scoring test-takers consistently answered correctly. Such items are generally poorly written and require modification.

These techniques take very little time to complete with computer support, and the first two can be done by hand with a little advance planning. The results guide your interviews with the test-takers; for example, ''Why didn't you select any of these three options? What would be a better choice for a distractor?'' or ''Why was this item so easy? Is there a cue to the correct answer elsewhere in the test?''

Interview the Test-Takers. After the test is administered, you should interview the test-takers, preferably individually, in private and shortly after the test is taken. You should plan your interview based on two sources of data: your observations during the test and the test analysis data. When you begin the interview, first remind test-takers about the formative nature of this experience and thank them for their cooperation. Then continue with your questions. Referring to the testing session, you might ask about difficulties they may have had; for example, ''I saw you scratching your head at question nine. Was that one a problem? Was it wording or content?'' If your analysis has identified problems with specific items, ask test-takers how they interpreted the item. As they talk, take careful notes. Don't concentrate only on their performance either; be sure to explore their feelings about the test; for example, ''How do you feel about this test? Would it be a fair test?''

A formative evaluation of the test should be pursued in the same manner as any other formative evaluation. Don't use a series of closed-ended questions; probe on responses, and summarize test-takers' comments to make sure you understand their thoughts.

Synthesize the Results

While your impressions are still fresh, you should synthesize your findings and document them. If your organization uses a standard form for course pilots, you might adapt it to meet your needs for testing. Some of the standard information that you may want to include would be:

- Time of test
- Location of the test
- Administrator
- Description of the participants
- The range of times it took to complete the test and the average time
- Any instructions or procedures that need to be modified
- Any test items that need to be modified

- Any format changes
- Any materials that need to be added or are unnecessary
- Overall impressions
- The item analysis report

HONESTY AND INTEGRITY IN TESTING

Honesty and integrity in testing have more than an ethical significance. Their lack contributes to the error component of test scores. We realize that testing can be a stressful event, the results of which may well affect employment or promotion. As a test administrator, you may find yourself in the uncomfortable position of resisting pressures to ''help'' a friend, colleague, or even worse, an important executive. Clearly, a testing system needs to be insulated from these pressures for ethical, political, and legal reasons. You should take precautions to protect the integrity of the test not only during a training-testing sequence but on an organization-wide basis.

Security During the Training-Testing Sequence

On a day-to-day basis when testing follows training, there are three safeguards you should use to protect your test from the error associated with cheating and a lack of test security.

Test Item Security. Allowing some test-takers to see the items in advance is obviously unfair to the others. An instructor might feel that it's ''okay'' to talk about the test to a whole group, since that ''wouldn't be unfair to any individuals,'' but this practice is still inadvisable for two reasons.

First, such an action will be unfair to other groups of test-takers who were not allowed to preview the test. Second, a preview may well destroy the validity of the test items. Providing answers to items means that test-takers only have to *remember* answers during the test; recalling memorized answers is not the same cognition as classifying new examples, applying principles to new situations, analyzing new information to identify patterns, and so forth, without help.

Interaction Among Test-Takers. Conversation or any other form of interaction among test-takers should not be allowed. Unless the test requires group interaction (for example, a cockpit flight simulation), the test-takers should not be allowed to talk during the test.

Test-Takers Must Do Their Own Work. In many instances, a test may be offered at the test-taker's work site rather than in an instructional or formal test setting. Usually these tests involve an assessment for course equivalency, entry, or prerequisite skills. When these tests are offered at a work site, there is a real opportunity for supervisors or others to provide assistance to the test-taker. It is most important that you communicate with the test-taker and his or her supervisor to emphasize the importance of achieving valid test results for placement purposes.

Organization-Wide Policies Regarding Test Security

Most organizations that are concerned about test policies concentrate on three points: security, access, and destruction.

Security of the Test. Test materials should always be inventoried and kept in a physically secure area. Obviously, tests should not be left out where they could be seen. Any requests for tests that are to be administered outside of the immediate training or testing areas (for example, an equivalency test mailed to a regional training office) should be logged and transmitted in a secure manner.

Access to the Test. A clear policy should be established and adhered to regarding who will have access to a test. A log system should be established that will provide for documentation of access to the test. If the test is available via computer, standard security measures (such as passwords) should also be implemented to limit access. Tests shouldn't be made available to anyone whose name is not on the authorization list.

Destruction of Tests. An organization should have a policy about retention of tests; for example, how long should an individual's test be kept in the event you need to provide evidence of performance in a legal challenge or a grievance? Tests need to be destroyed when test forms or answer sheets have been written on, test forms are worn out, or test copies are defective or incomplete.

PERFORM ITEM ANALYSIS

As described above, item analysis results are likely to be more reliable when based on data from large numbers of test-takers. While these large numbers are logistically often difficult to obtain, two points are worth remembering in this regard. First, data for item analysis need not be collected in a single test administration; assuming reasonable stability in the relevance of the items, the test administration procedures, and representativeness of the pilot test-takers, the raw data for the item analysis could be collected over multiple occasions and then collapsed into a single item analysis. Second, item analysis statistics are used differently and perform a less critical function in the creation of CRTs as opposed to NRTs; item analysis statistics are examined to *improve* questions for a CRT, but to *select* questions for an NRT.

Deciding Which Item Analysis Techniques to Use

The item statistics you gather and the subsequent item analysis will allow you to refine the accuracy of your test. The first point to remember, though, is advice from Hambleton (1994):

> ''Criterion-referenced test items are only deleted from the pools of test items measuring objectives when it is determined that they are not consistent with the content specifications, violate standard principles of item writing, or if the available item statistics reveal serious noncorrectable flaws. Item statistics can be used to construct parallel forms of a criterion-referenced test or to produce a test to discriminate optimally between masters and non-masters.'' (p. 1184)

Again, keep in mind that your most accessible staff psychometrician is most likely to give you NRT advice rather than guidance appropriate for CRTs.

With the distinction between CRTs and NRTs in mind, we would suggest that you proceed as follows:

- Always review the difficulty index and distractor pattern for each item. These two techniques will help you to find quickly items that need attention. For example, you need to be sure that items that everyone or nearly everyone answers correctly do not contain flaws such as cues or unrealistic distractors that allow non-master test-takers to answer correctly. An item that no one answers correctly should prompt a check of the scoring key and other assessment.
- You can use the point-biserial correlation for any low- or medium-stakes test with a fair degree of confidence. Again, an item with a negative point-biserial correlation coefficient is a candidate for reexamination because it seems to favor lower test scorers. However, reexamination is the proper response by a criterion-referenced test developer rather than abrupt elimination of the question, the efficient course of action for a norm-referenced test writer.
- If you are designing a high-stakes test with serious consequences, you may need to consider a more sophisticated approach to item analysis. Oshima, McGinty, and Flowers (1994, p. 196) point out that with a criterion-referenced test, ''developers should be extremely concerned about DIF (differential item functioning) that occurs near the cutoff score but much less concerned when it occurs considerably above or below the cutoff score.'' What this statement means is that techniques such as the upper/lower index will show you whether an item is discriminating between masters and non-masters in general. However, such a technique is too

coarse an approach for dealing with test-taker scores that are near the cut-off point; for example, is there a real difference in mastery classification when a test-taker fails a test with a score of 84 percent when the passing score is 85 percent? The good news is that corporate CRTs, based on a solid job task analysis, are unlikely to require such precise item discrimination analysis—but we thought we'd at least give you a ''head's up'' here.

Garbage In/Garbage Out

Finally, remember that an item analysis, like any numerical technique, is only as good as what goes into it. As they say, garbage in/garbage out (GIGO). The item analysis won't tell you anything about the quality or the job relatedness of the objective being measured by the item. It doesn't tell you how accurately an item assesses a given objective. You could have an item with a high point-biserial, full use of the distractors, and a high difficulty index on a test supposedly assessing supervisory skills, when in reality the item is measuring reading comprehension. Once again, there is no substitute for competent professional judgment in the testing process. An item analysis package will analyze response patterns for the most trivial as well as the most crucial questions. It is essential that you not forget the purpose for creating a CRT in the first place when surrounded by these quantitative indices so powerful to the creation of the more familiar NRT.

CREATE PARALLEL FORMS

When a test is expected to be used a number of times or in a number of places throughout its life, the issue of parallel forms inevitably comes up. Parallel forms are different versions of a test that measure the same objectives and yield similar results. Whether you are thinking about two different forms of the same test in the tradition of Form A and Form B in the university large lecture hall exam, or creating tests randomly created from a computerized item bank, parallel forms can be useful. Some arguments for creating them are:

- Parallel forms can be important if the security of a test is breached; the loose or circulated form can be destroyed and a parallel form placed into immediate service.
- Parallel forms are helpful in case an employee scheduled for group testing has to cancel and take the test at a later date; such an employee can be given a parallel form of the test without fear that the answers to the test may have been shared.
- Parallel forms allow for retesting of individuals who score too close to the master/non-master cut-off score to be classified with confidence.

Creating parallel forms of a test requires very careful matching of items in terms of the objectives they cover and the ways in which test-takers respond to them; matched items should have approximately the same difficulty levels and discrimination indices or point-biserial correlations.

SET THE CUT-OFF SCORE

One of the most difficult, yet critical, tasks required in CRT development is to determine the standard for passing, that is, the cut-off score that separates masters from non-masters. Since 1989 there have been a number of refinements and additions to standard setting processes, with the most recent work attempting to refine decisions made at the borderline between mastery and non-mastery. Kane (1994) offered a discussion about selecting passing scores, and Hambleton provided an update on criteria for evaluation of performance standards in 2001. Donath (2005, pp. 6–7) lists eight techniques that are currently used in various settings and situations: Angoff, Modified Angoff, Nedelsky's method, Ebel's method, Contrasting Groups, Jaeger's method, Bookmark, and Direct Consensus.

A number of these discussions are highly theoretical, but the three techniques—the informed judgment, contrasting groups, and conjectural methods (specifically the Angoff technique)—we discuss in this section are still the most viable for corporate settings. In fact, the Angoff technique is probably the most popular of all the cut-off score procedures for corporate CRTs (Impara & Plake, 2000).

However, before we look at these procedures, we want to make you aware of a number of considerations that affect the standard-setting process, regardless of the method used.

The Outcomes of a Criterion-Referenced Test

Following the assumptions of criterion-referenced testing, the true status of every test-taker is either a master or a non-master. A reliable and valid test will lead to a mastery judgment that matches the test-taker's true status. If the test-taker is a non-master and is classified as such, or the test-taker is a master and is classified as a master, then we have made the correct decisions. However, if the master is judged to be a non-master, we have made an error of rejection (also called a false negative). If a non-master is judged to be a master, then we have made an error of acceptance (also called a false positive). The only way to minimize these errors is to ensure that your test is reliable and valid.

The Substitutability Issue

When establishing a cut-off score for a test, a common procedure is to set the score based on the test-taker's ability to pass a certain number of items (for example, 85 out of 100). In doing so, however, a critical assumption is being made—one that you *must* consider if you are to develop a valid test. The assumption is that failure to perform correctly on one item can be compensated for by success on another item. For example, in our test of one hundred items, if all items are equally substitutable, it doesn't matter which of the eighty-five are answered correctly. However, if there are certain items that must be answered correctly for mastery performance, then failing any of those items should mean failing the test—even if only one item is missed; for example, in constructing a test to assess skills in utility-pole climbing, a test-taker who fails to wear safety glasses fails the test. Efficient climbing, wire splicing, and descent cannot overcome the single failure to follow an essential safety rule. When using any of the three techniques covered in this chapter, you may determine:

- That all items are substitutable and a single cut-off score is acceptable, or
- That you can separate the non-substitutable skills and establish a two-tiered scoring system where a score of 100 percent is required on the non-substitutable items, and a given percentage is required for the remaining items.

Again, we want to emphasize that the chosen cut-off score should not be considered absolutely final. The operation of the chosen cut-off score should be monitored periodically to make sure that it is rendering decisions that are satisfactory to those involved and facilitating the achievement of the company's objectives. In the end, there is never any substitute for good professional judgment.

Three Procedures for Setting the Cut-Off Score

With these thoughts in mind, we are almost ready to turn to three different but complementary methods of determining the cut-off score for a criterion-referenced test:

- The *informed judgment method* draws primarily on perceptions of various stakeholders in the organization;
- The *conjectural methods* base cut-off scores on content expert projections of competent performance on each test item; and
- The *contrasting groups method* uses performance data of masters and non-masters to establish the level of mastery.

We recommend, and research shows (Impara & Plake, 2000), that you should use as many of these standard-setting methods as you can to establish the test cut-off score through a triangulation process. However, we will only discuss the most commonly used technique of these three—the conjectural approach presented by Angoff (Zieky & Livingston, 1977). The other two are discussed by Shrock and Coscarelli (2007) and in other measurement texts.

Using the Angoff Technique

The Angoff method is one of a general class of techniques for estimating the cut-off score of a test by determining estimates of success for a minimally competent performer on each item. These are the steps in the Angoff method that we recommend you follow to enhance the validity of the results:

1. The first step is to identify judges who are familiar with the competencies covered by the test and with the performance level for masters of these competencies. The number of judges you select will depend on availability of judges, criticality of the performance, etc. However, we think you would rarely need more than five, with three being the more typical number.
2. Plan a face-to-face meeting with the judges to set the cut-off scores. We recommend you conduct an Angoff session face-to-face instead of virtually, as critical subtleties of the process seem to be lost without the human interaction. We have had some success with virtual meetings, but only when all the judges had been engaged in an initial face-to-face meeting.
3. Have the judges come to a consensus regarding the definition of the borderline or minimally competent performer. It is important to point out that a ''minimally competent'' person is not an incompetent one. A minimally competent pilot, for example, is still very competent. Another common discussion at this stage focuses on whether the organization is comfortable with mastery levels of performance as they currently exist in the organization or as they might exist if higher standards prevailed. The outcome of that discussion will have an important impact on the definition of the minimally competent performer.
4. Print a copy of the test on a different colored paper for each judge; this will facilitate sharing estimates later in the meeting. Test items should also show the correct answer to each question.
5. Hold a discussion with the judges about the estimation process that addresses the following issues *before* beginning to assign estimates:
    - Make sure the judges understand that a probability estimate should never be lower than the probability of getting the item correct by chance; for example, if there are four alternatives in a multiple-choice item, the estimate should not be lower than 25 percent. These estimates are expressed as percentages and assigned a corresponding decimal value. For example, if a judge thinks there is a fifty-fifty chance of the minimally competent test-taker getting a given item right, that item is assigned a value of .50. If the judge estimates that an item is either so simple or represents content so critical that the minimally competent test-taker will almost surely get it right, then the item would be

assigned a value of 1.0. Estimates need not be in increments of 25 percent or 20 percent; their value can vary anywhere from chance level (the probability of guessing correctly among the possible answers) to 1.0 (certainty that the item will be answered correctly).

- Tell the judges to examine each answer and the distractors. Decide how many distractors the minimally competent test-taker could surely eliminate. This would provide a floor to the rating of the item; for example, if one option on a four-choice item can almost certainly be eliminated by any competent performer, then the lowest estimate for this item will be 33 percent. If the choice between or among the remaining distractors is truly random, then estimate the Angoff weight as the chance probability between/among the remaining option(s). For example, if there are two choices remaining from four distractors, then the probability would be 50 percent or .50. If the choice among the remaining distractors is NOT random, then the decision is based on how far the judge believes the deviation from random should be.
- Judges usually need coaching at the outset of this process. They tend to go through three stages while learning to estimate the appropriate probability for items: Their first reaction is to estimate the difficulty of the item for themselves. Since the judges are rarely at the minimally competent level, using this estimate will produce an artificially high estimate. Judges then tend to estimate the difficulty for the typical test-taker. Finally, after these first two misconceptions have been corrected, judges see that they are setting a standard for how the minimally competent performer should be able to answer.

6. Start with a single item and have the judges independently estimate the probability that a minimally competent test-taker would get the item right.
7. Have the judges share their privately derived estimates and discuss why they decided as they did.
8. Clarify the process and content questions as needed until consensus is reached or until each judge feels comfortable with his or her estimate on the first practice item.
9. Pick a second item; again, have the judges rate it privately and then discuss their estimates. Continue this process until the judges feel comfortable with their role.
10. Send the judges to private spaces to rate the entire test. When they are finished, have them return to the meeting room and have them dictate their estimates (to speed the process) for entry into a spreadsheet program.

11. Review the ratings and have judges discuss any Angoff weight for an item that varies by more than .20 among the judges' estimates. Judges can choose to change the weight for an item in any direction based on the results of their subsequent discussion or maintain their original estimates.
12. The chosen cut-off score is the sum of the probability estimates. If more than one judge is used, the cut-off score is the average of the individual judges' sums.

ESTABLISH RELIABILITY OF COGNITIVE ITEMS

As described in the early part of this chapter, test reliability refers to the consistency of test scores. Noting that there are several different kinds of reliability—several different ways in which test scores can be consistent—we introduced conceptually the three most common types: equivalence, test-retest, and inter-rater. Inter-rater reliability typically refers to performance tests, and it will be discussed in the next section of this chapter.

Statistical Procedures

When one considers how to calculate a reliability coefficient for a cognitive test, it is useful to know that some reliability calculations are based on two test administrations, while others are based on only one.

Reliability estimates based on two test administrations are

- Equivalence reliability
- Test-retest reliability

Reliability estimates based on a single test administration are

- Internal consistency
- Squared-error loss
- Threshold-loss

Recommendations for Choosing a Reliability Estimate

The test designer needs to make a decision about how to proceed based upon which type of reliability is of greatest concern and the resource and logistical limitations of the organization. As we have said on a number of occasions, there is no substitute for professional judgment—and you will have to interpret your needs in light of your own context to decide which technique is best. For most corporate tests, we would recommend your considering these options:

- Use phi to calculate reliability using the two-test administration technique, correlating either two parallel forms of the test (for equivalence reliability) or two administrations of the same test (for test-retest reliability over time). Phi is easy to calculate and is probably the easiest technique to explain. It would be the most powerful evidence of a test's stability should you find yourself in a high-stakes testing situation (Shrock & Coscarelli, 2007).
- To estimate a CRT's reliability using the easiest logistics, use Livingston's kappa (for squared-error loss), which can be determined from traditional test statistics in a single test administration (Shrock & Coscarelli, 2007).
- If you can't use a two-test design, and if the test is being used to make a significant decision, then you may wish to consider using S_C (Coscarelli & Shrock, 2002)—a split half technique that first stratifies the test by objectives and then correlates the consistency of test decisions. The S_C estimate will likely be lower than either Cronbach's alpha (internal consistency) or Livingston's kappa (squared-error loss) as it reflects the dual effects of removing chance agreement from the estimate as well as the variation of performance around the cut-off score (Coscarelli & Shrock, 2002, p. 85).

Summary Comment About Reliability and Validity

It should be apparent that it requires an investment of time and resources to establish the reliability and validity of a test. One way to streamline the process is to compose the groups of sample test-takers for the concurrent validation process—the group of known masters and non-masters—and arrange to have them take the test twice or take both forms of the test in order to establish the test's equivalence or test-retest reliability. It should also be noted, however, that it is difficult to overstate the importance of reliability and validity for tests of any kind. Tests that are not reliable and valid are worse than useless; they are misleading and risky.

ESTABLISH RELIABILITY OF PERFORMANCE TESTS

The reliability problems associated with performance testing are different from those associated with cognitive tests composed of closed-ended items. In some ways, the reliability problems associated with performance tests are similar to those posed by essays and other types of open-ended questions; unlike closed-ended assessments that are even machine scorable, the test-taker's creation in response to an essay test or behavior during a performance test must be rated or judged by an observer. Therefore, the locus of reliability shifts from the test itself to the consistency of the judges' observations.

As described earlier, it is best to use a checklist or a behaviorally anchored rating scale for performance tests. Error is reduced primarily by the precise specification of criteria and careful preparation of raters.

Procedures for Training Raters

For professional and legal reasons, raters need to make consistent and accurate ratings. Good judges have a combination of experience and training that can serve as a model for any rating situation. There are some simple steps you can follow to train your raters to these high standards:

1. Bring together those people who are familiar with the skill or product to be rated and who will later be asked to serve as raters.
2. Plan a rater training session at which you will have available a sample of performances or products that are to be rated. In an ideal setting, you would have a model case performance (or product) where all the attributes of a correct performance are present, a clear non-example of the performance (or product), and a range of stimuli between these two extremes—perhaps with the most common errors illustrated. This training session can be based either on a live performance or on a high-fidelity media simulation, for example, a videotape of an assembly process. If the rating is of a product, the product itself should be present.
3. The raters should be presented with the first stimulus, usually the model case performance. All raters then use the checklist to review the performance or product. If the performance is mediated, the tape can be stopped and an action discussed. If the performance is live, plan to record the actions (for example, a tape recording of an air traffic controller's interchange), so that raters can discuss specific behaviors that may not be clear.
4. Provide the next stimulus, often the non-example, and have the raters assess the performance (or product) as they would during an actual testing session. Again, record the activities if they are live.
5. Ratings for each behavior should be tabulated as a percentage of agreement and posted for the group to see. Raters then share their assessments, step by step, with the other raters. Points of contention (low percentages of agreement) should be reviewed on the tape and discussed until all the raters understand the reasoning behind the correct assessment.
6. A new stimulus is presented, ratings tabulated and shared, and then discussed. This cycle is continued until the judges have reached a high degree of consistency (90 percent).

7. Ten final trials are then presented to the group and the subsequent group decisions used to calculate the inter-rater reliability estimate.
8. Document the final inter-rater reliability estimate and collect the stimulus materials.

Calculating Inter-Rater Reliability

It is important to establish the degree of consistency with which raters are judging the performances of test-takers. The mastery/non-mastery decision made about each test-taker should be determined by what the test-taker does, not by differences among the judges, either in regard to what they see or the value they place on what they see. Remember that, since reliability is a prerequisite for validity, if the judges are inconsistent, the performance test decisions cannot possibly be valid; the value of a performance test rests on determining inter-rater reliability.

There are two conceptually simple methods for assessing inter-rater reliability that yield comparable results. One is based upon a corrected percentage of agreement figure (kappa or κ), while the other is a correlation coefficient (phi or ϕ). The values resulting from the two different calculations are similar. Which one you decide to use really depends on which of the two statistics you find the easiest to understand or perhaps which of the two you think other interested parties will find the easiest to understand.

The percentage of agreement coefficient, kappa coefficient, and phi coefficient and how to calculate them are explained in great detail in Shrock and Coscarelli (2007, pp. 322–344).

REPORT SCORES

Now that the testing process is complete, you need, finally, to turn your attention to reporting the scores. Reporting the test scores is a politically volatile issue in the testing process that will probably be raised at the beginning of the test-development effort, if the stakeholders believe the test will actually be used to make decisions.

Reporting CRT Results Versus NRT Scores

As you enter into what we hope will be constructive discussions about the use of test scores in the organization, please keep in mind the fundamental assumption of criterion-referenced testing—you can only make one judgment about a test-taker's performance: either he or she was classified as a master or was not. This assumption is easily lost on managers who want to use the test scores as a means of sorting people for merit or other purposes. Neither raw scores of test-takers nor class averages should be reported following the administration of a criterion-referenced test.

You cannot use the CRT process to make any distinction beyond the master/non-master decision. Much to the consternation of many managers, there is no acceptable legal, professional, or ethical argument that someone with a score of 82 should be treated differently from someone with a score of 95 if both are above the cut-off point. But there can be a big difference in the consequence to two people if the cut-off is 80 and one scores 78 and the other 82.

What Should You Report to a Manager?

This question is actually one that should be addressed at the beginning of the test development process so that all the test's stakeholders have a shared understanding of how the scores will be used. In general, though, we think there are some minimal types of information that would help the organization make decisions surrounding the test-taker's performance:

- Report the test-taker's score as mastery or not.
- Describe the test measures that were used to assess the skills or knowledge.
- Indicate what, if any, remediation options are available for test-takers who do not meet mastery. Typically, remediation means on-the-job supervision and coaching, repeating a course, or individual study of areas that need improvement, followed by retesting.
- Provide the name of the person the manager can call if he or she has questions about the test.

A FINAL THOUGHT ABOUT TESTING AND LEARNING

In the tension that can exist between the test designer trying to do a job well and the manager who also wants to succeed, we feel there is usually common ground to be shared. While a test score often brings with it the baggage of early schooling experience, it need not divide the organization (for example, subordinate versus supervisor, management versus union), but rather create a bridge toward common organizational goals. We think, in the best of all worlds, the common good can be found in understanding and supporting the fundamental concept that necessitated the invention of criterion-referenced assessment—mastery learning. Criterion-referencing was created to support learning. Often the resistance one finds to testing comes when people feel there is an imbalance in what should be complementary processes—challenge through good testing and support through good instruction. We think organizations that bring both factors to the table are the ones that will prosper.

References

AERA/APA/NCME Joint Committee. (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives. Handbook I: Cognitive domain*. New York: David McKay.

Burton, S. J., Sudweeks, R. R., Merrill, P. F., & Wood, B. (1991). How to prepare better multiple-choice test items: Guidelines for university faculty. Provo, UT: Brigham Young University Testing Services and the Department of Instructional Science. Retrieved November 23, 2006, from http://testing.byu.edu/info/handbooks/betteritems.pdf.

Campbell, C. P., & Hatcher, T. G. (1989). Testing that is performance based and criterion-referenced. *Performance and Instruction*, *28*(5), 1–9.

Case, S. M., & Swanson, D. B. (2001). Constructing written test questions from the basic and clinical sciences. Philadelphia: National Board of Medical Examiners. Retrieved November 23, 2006, from http://www.nbme.org/about/itemwriting.asp.

Coscarelli, W., & Shrock, S. (2002). The two most useful approaches to estimating criterion-referenced test reliability in a single test administration. *Performance Improvement Quarterly*, *15*(4), 74–85.

Donath Group. (2005). Test item writing best practices: Useful and practical considerations. Waltham, MA. Retrieved December 14, 2006, from http://e-learning.umh.ac.be/file.php/1/Guide_pratique_pour_la_conception_de_tests.pdf.

Eyres, P. (1997). *The legal handbook for trainers, speakers, and consultants*. New York: McGraw-Hill.

Gagné, R. M. (1985). *The conditions of learning* (4th ed.). New York: Holt, Rinehart, and Winston.

Gronlund, N. E. (1988). *How to construct achievement tests*. Englewood Cliffs, NJ: Prentice-Hall.

Hambleton, R. K. (1994). Criterion-referenced measurement. In *The international encyclopedia of educational evaluation* (pp. 1183–1189). New York: Elsevier.

Hambleton, R. K. (2001). Setting performance standards on educational assessments and criteria for evaluating the process. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 89–117). Mahwah, NJ: Lawrence Erlbaum Associates.

Hambleton, R. K. (1994). Criterion-referenced measurement. In *The international encyclopedia of educational evaluation* (pp. 1183–1189). New York: Elsevier.

Impara, J. C., & Plake, B. S. (2000). A comparison of cut scores using multiple standard setting methods. Paper presented at the Large Scale Assessment Conference. Snowbird, UT. Retrieved December 14, 2006, from http://www.education.ualberta.ca/educ/psych/crame/files/RickerCSSE2003.pdf.

Kane, M. (1994). Validating the performance standards associated with passing scores. *Review of Educational Research*, *64*(3), 425–461.

Mager, R. F. (1962). *Preparing instructional objectives*. Belmont, CA: Fearon.

Oshima, T. C., McGinty, D., & Flowers, C. P. (1994). Differential item functioning for a test with a cutoff score: Use of limited closed-interval measures. *Applied Measurement in Education*, *7*(3), 195–209.

Popham, W. J. (1978). *Criterion-referenced measurement*. Englewood Cliffs, NJ: Prentice-Hall.

Shrock, S. A., & Coscarelli, W. C. (2007). *Criterion-referenced test development: Technical and legal guidelines for corporate training* (3rd ed.). San Francisco: Pfeiffer.

Thompson, D. E., & Thompson, T. A. (1982). Court standards for job analysis in test validation. *Personnel Psychology*, *35*, 865–874.

Uniform guidelines on employee selection procedures (1978). *Federal Register*, *43*, 38290–38309.

Uniform guidelines on employee selection adoption of questions and answers to clarify and provide a common interpretation of the Uniform Guidelines on Employee Selection Procedures (1979). *Federal Register*, *44*, 11996–12009.

Zieky, M., & Livingston, S. (1977). *Manual for setting standards on the basic skills assessment tests*. Princeton, NJ: Educational Testing Service.

CHAPTER SIXTEEN

The Role of Evaluation in Instructional Design

Steven M. Ross
Gary R. Morrison

As described in the preceding chapters, quality instructional design is partly science, partly art, but always a reflective, systematic application of research-based principles of teaching and learning. Most educators and training professionals, however, know the field only through its products—instructional programs, units, mediated units, courses, and interventions that help learners (the consumers of instruction) achieve desired educational goals. But the development of successful products requires more than simply diligent effort, professional experience, and good faith by designers. As is (or should be) true for all fields that deliver products to consumers (automotive, medical, nutrition, sports, and others), *evaluation i*s an essential component of the instructional design process: (a) as the products are being developed, (b) when they are in completed form and ready for release, and (c) as they are being employed over time.

Komoski (1974) described the lack of evaluation in the development of educational materials in the late 1950s to mid-1960s as the go-go years marked by instructional materials in which instructional effectiveness was not a major factor. Komoski proposed the use of learner verification for instructional materials directed at K-12 education. Obviously, in the absence of evaluation, the instructional materials made available to trainers, teachers, and learners would be unrefined and untested based on valid and reliable empirical data, a potential disappointment and loss to the consumer and a black mark against the designer.

WHAT EVALUATION ENTAILS

Although the basic meaning of ''evaluation'' is familiar to all readers, there should be definite benefits in starting with a common, more formal definition. Simply put, *evaluation* is using assessment or measurement to judge the worth or value of something (Morrison, Ross, & Kemp, 2007). This rather straightforward definition is also useful in differentiating ''assessment'' from evaluation. Assessment involves collecting data or information about people or events, whereas evaluation uses such data to make a judgment. For example, a measurement of test performance in a class on sexual harassment reveals the average is forty out of fifty items correctly answered. The *evaluation* of this performance, however, might be the judgment that it falls below standards, thereby suggesting a weakness in the instruction.

Evaluation can and should take place at different phases of the instructional design process. As reflected in the title of this chapter, these evaluation applications have names—formative, summative, and confirmative (see Table 16.1). *Formative evaluation* is used to provide feedback to designers as the instruction is ''forming'' or being developed. Consequently, the designer has an objective basis for knowing what works as planned and what needs to be improved *before* the product is made available for usage by practitioners. *Summative evaluation* is conducted to determine the degree to which a *completed* instructional product produces the intended outcomes. Thus, the designer and other stakeholders can judge the effectiveness of the instruction for present and future uses. *Confirmative evaluation* examines the success of instruction as it is used *over time*. The rationale is to determine whether instructional outcomes continue to meet goals or vary due to changing applications or environmental conditions. In the next section, we will examine the purposes and practices of each type of evaluation in more detail.

Formative Evaluation

The concept of formative versus summative evaluation was formalized over 40 years ago by Michael Scriven (1967; 1991). Using Robert Stake's classic cooking metaphor (Scriven, 1991), formative evaluation can be described as the cook tasting the soup and summative evaluation as the guests tasting it. In an instructional design context, we can easily substitute the designer for the cook and the learner (or instructor) as the guests. A designer who fails to make necessary improvements before the instruction is released may well end up ''in the soup'' with project stakeholders, consumers, and, especially, his or her boss. In other words, the ''guests'' in Stake's metaphor are not just the immediate diners but many others who may be impacted by the recipe, to include learners, the client wanting a solution to a business problem, instructors, training managers, and others.

Table 16.1 Summary of Key Features of the Three Evaluation Modes

Evaluation Mode	*Key Questions*	*Timing*	*Orientation*	*Methods*
Formative	How well is the product working? What can be done to improve its effectiveness?	Development Stage	Informal, ongoing, as-needed, designer usually involved	Expert (connoisseur) review; one-on-one trials; small-group trials; field trials; designer as evaluator; measures of student learning, attitudes, completion time, implementation requirements
Summative	How well did the product work? Which objectives or goals were or were not satisfied? Should usage be continued?	Completion Stage	Formal, often ''high stakes,'' could involve replication studies at different sites, usually third-party evaluator	Experimental or quasi-experimental; pre-test/post-test design with user group; rigorous qualitative analyses (observations, interviews, student performance assessments; mixed methods (quantitative and qualitative); measures of student learning, attitudes, completion time, implementation requirements
Confirmative	How well is the product continuing to work with new learners? What are the product's long-term impacts on former students/trainees? Which objectives or goals are still being satisfied? Which are not? Should usage be continued?	Maturity Stage	Same as Summative Evaluation	Same as Summative Evaluation

Types of Formative Evaluation

There is no single best approach or specific model for conducting formative evaluations. Given the basic purpose of obtaining feedback for improving instructional design, alternative ways of gathering the data needed for decision making exist. According to Flagg (1990), there are four basic types of formative evaluation. First, a *connoisseur-based study* largely depends on expert review of the instructional program as opposed to collecting data from potential users (see Morrison, Ross, & Kemp, 2007). This approach will be illustrated in the case study presented at the conclusion of this chapter section on formative evaluation. A second category is the *decision-oriented study*, which is designed to provide data related to specific questions of interest to the designer in collaboration with stakeholders, such as ''Would a pre-test be used? Is there adequate time to administer and score one?'' Third, an *objectives-based study* most closely resembles summative and confirmative evaluations by assessing outcomes for specific instructional objectives, often using pre-test/post-test designs in which learner progress after receiving instruction is gauged. Fourth, *public relations-inspired studies* are designed largely for the purpose of making preliminary evaluation results known to targeted individuals, such as key stakeholders and funders. We have used the latter type considerably in our work with organizations that have sponsored the development of a particular instructional design product and want to know well in advance of product completion how well the work is progressing.

One or more of these formative evaluation approaches may be used at different stages of the design project. For example, conducting a decision-oriented evaluation of the first few instructional modules developed would provide useful feedback for developing design standards for those units and for the project in general. Later, expert (that is, connoisseur-based) review of subsequent modules might be used to corroborate the standards and suggest further refinements. Objectives-based and/or public relations–inspired studies would generally prove most valuable when the design project is in a sufficiently completed stage to determine its success in achieving objectives and disseminating evaluation outcomes to key stakeholders

Stages of Formative Evaluation

Perhaps with the exception of connoisseur-based evaluations, which rely exclusively or predominantly on expert review of the instructional program, formative studies will often be conducted in stages of increasing formality and intensity resembling the following (Dick, Carey, & Carey, 2005; Morrison, Ross, & Kemp, 2007):

- *One-on-one trials.* In this ''developmental'' stage, the designer will try out the newly created instruction with individual learners, using observations and interviews to obtain impressions of strengths and weaknesses.

- *Small-group trials.* In this second ''preliminary'' stage, the designer expands the try-out of the instruction and associated evaluation to small groups of eight to twenty individuals.
- *Field trials.* In a third stage, the formative evaluator examines a ''completed'' version of the instruction using, if possible, classes or learning contexts similar to the ones for which the product is intended. Assessments of performance and attitudes are typically made to provide impressions of effectiveness before the product is released.

Constructivist-Oriented Approach

Recently, Kay and Knaack (in press) have criticized traditional formative evaluation approaches as too limited and behaviorally oriented to assess effectively and meaningfully contemporary web-based instructional tools, which they generically label ''learning objects.'' Such support tools, which encompass such supports as help systems, prompting, hierarchical navigation aids, program-generated questions, graphics, and animations, are not simply designed to increase the amount of content remembered (Friesen & Anderson, 2004; Krauss & Ally, 2005; Nurmi & Jaakkola, 2006), but to engage students in higher-order learning and problem solving. The formative (and summative) evaluation model they propose uses a combination of the connoisseur-based and decision-oriented approaches to obtain *valid* triangulated data addressing the three factors of *learning construct* (interactivity, feedback, graphics, and so forth), the *quality* of the instruction (help supports, instructions, organization, and so on), and *engagement* (interest, motivational, enjoyment, and so forth). Data sources are students participating in field trials who complete a survey, make open-ended comments, and complete pre-tests and post-tests on the content of the instruction. Teachers, in turn, implement the field trial, while answering questions on the degree of student learning, instructional quality, and student engagement.

Using this multi-faceted model, Kay and Knaak extend the traditional objectives-based approaches to formative evaluation by examining broader, constructivist properties of the instruction. Depending on preferred theoretical and operational preferences of the designer (see Morrison, Ross, & Kemp, Chapter Thirteen), this approach may be considered as a useful paradigm for the increasing design work being performed to accommodate open-ended and web-based learning environments (Clarebout & Elen, 2008; Saab, Gijlers, Van Joolingen, & van Hout-Wolters, 2008).

Formative Evaluation Methods

There is no simple ''cookbook'' approach to conducting a formative evaluation or, for that matter, any type of educational study. Numerous ''models'' exist, ranging from classic behavioral approaches to newer constructivist paradigms (as in Kay & Knaak, in press, above), and from the simple to complex (see Stufflebeam, 1983; Weiss, 1998). In our opinion, the most skilled and successful evaluators are not only scientific and artistic but also highly pragmatic. Although the latter practical orientation receives scant attention in the formal literature on evaluation methods, in real life it becomes essential in balancing the collection of useful and valid data with available resources (budget, data sources, participants) and stakeholder interests.

A scientifically valid but also highly practical formative evaluation model that we have implemented many times consists of the following processes (Gooler, 1980; Morrison, Ross, & Kemp, 2007):

Step 1: Determining the Purpose. The ''purpose'' of a formative evaluation will be context-specific, depending on the nature of the instructional design project, the resources at hand, expectations of stakeholders, time frame, and other factors. Identifying the purpose should also set the groundwork for later determining which type of evaluation approach (connoisseur-based, objectives-based) is most apt.

For example, one possible purpose might be to determine the readiness of the project for release to consumers, whether or not there is an opportunity to revise it at the current time. A related purpose with a different goal, however, would be to gather data for improving the product, whether or not it is considered ready for release. Rapid prototyping, a fundamental part of some instructional design models, provides continuous formative feedback as the instructional materials are being developed. A third purpose might be to gather data so that specifications can be identified and documented regarding completion time, normative performances, administration requirements, and so on. A fourth purpose might be to address the project sponsor's organizational requirements for subjecting new products to ''quality control'' testing (formative evaluation) prior to release. Note that quality control, which attempts to fix errors in production, is less efficient (less economical) than ''quality management,'' which prevents the error in the first place through process improvement. Overall, the primary purposes of the formative evaluation are to determine how the product is likely to be used and received by learners, its present strengths and weaknesses, and ways that it can be improved to increase effectiveness.

Step 2: Identifying the Audience. This step involves determining who will be the major consumer of the evaluation—instructors, managers, funders, or perhaps *you* as the instructional designer. Depending on the audience (consumer), the focus of the evaluation (instructional properties, costs, program logistics, and others) and the way it is reported (level of technicality) are likely to vary.

Step 3: Defining the Evaluation Questions/Objectives. What specific questions need to be answered by the evaluation as influenced by the major purposes (Step 1) and the primary audience (Step 2)? Professional evaluators consider it essential to define the evaluation questions or objectives explicitly as a framework for informing subsequent steps of determining the needed resources, types of required evidence, data-gathering techniques, analyses procedures, and reporting needs. Common types of questions deal with learning outcomes, completion time, and student interactivity and interest/motivation, as illustrated below:

- What percentage of students attain ''proficiency'' (80 percent correct or higher) on the biology tutorial?
- How much time does it take the average learner to complete the unit?
- What are student perceptions of the unit with regard to difficulty, interest, and instructional value?
- What is the quality of the instruction with regard to organization and learner support?
- How frequently do students select the various types of help tools provided?
- What is the level of student engagement with the software?

Step 4: Determining Resource Needs. This step is directly linked to the evaluation questions/objectives defined in Step 3. Simply put, it involves asking what equipment, instrumentation, facilities, and human participation (subject-matter experts, students, instructors) are needed to address each question.

Step 5: Defining Evidence Needs. What types of data are needed to answer each evaluation question sufficiently and credibly? With regard to the six sample questions presented in Step 3 above, likely choices would be, respectively: (1) student performance outcomes from the ''proficiency test,'' (2) completion time data for each student participant in the formative evaluation trials, (3) interview or questionnaire data collected from students, and (4) online recording of student interactions with the instruction (see Chapter Four in this volume of the Handbook on data collection methods).

Step 6: Collecting Data. This step involves deciding how to collect the evidence needed for the formative evaluation. In the example above, four types of evidence—student achievement, completion time, student perceptions, and online data—were identified. Data collection decisions encompass questions such as: What instruments will be used? From how many students is data needed? Who will collect the data, at which times, and in what contexts? Answers to these questions, when documented, will form the data collection plan to be followed by the evaluator(s).

Step 7: Analyzing the Data. The target audience for the evaluation (Step 2), the evaluation questions/objectives (Step 3), and the type of evidence gathered (Step 5) will strongly influence the data analysis techniques employed. The analytical methods appropriate for formative evaluations typically are simpler and more informal than those used for summative evaluations, given the purpose of obtaining useful information for improving instruction (as opposed to testing the effectiveness of a completed product). Accordingly, descriptive statistical analyses such as frequency distributions, computation of means and medians, and graphs and charts will often suffice. When analysis needs require greater sophistication or complexity, subcontracting with a statistical expert is usually well-advised to ensure accurate results.

Step 8: Preparing and Disseminating a Report. For several reasons, it is almost always desirable to prepare a written report of the formative evaluation, even when the study is conducted informally and internally (that is, by the designers for their own needs). One benefit is documenting that the instruction was subjected to formative study for the purpose of identifying and making needed improvements. Another is to provide concrete directions for revision as guidelines that the original or newly involved designers can follow. When the designer or evaluator prepares the report, a third advantage is more active analysis and synthesis of the data obtained.

There is no established format for writing formative evaluation reports. However, a reasonable generic orientation would be modeled on conventional educational research and evaluation study reports:

I. Executive Summary. One- to three-page summary of the study purposes, methodology, and main findings.

II. Purposes. What were the evaluation questions/objectives?

III. Methodology. Who were the participants? What were the data collection instruments? What was done to collect the data?

IV. Results. What analyses were conducted? What were the findings?

V. Conclusions and Recommendations. Based on the results, what are the strengths and weaknesses of the instructional product? What improvements are suggested in what order of priority?

Case Study

Recently, one of us was asked to design a "Guidebook" to assist state directors of "Supplemental Educational Services" (a federal program that provides free tutoring to disadvantaged students) to conduct state-wide evaluations of the effectiveness of the program in raising student achievement (Ross, Potter, & Harmon, 2006). As a first step, a draft version of the guidebook was created based on the three primary authors'

(*Continued*)

(*Continued*)
knowledge of the content and instructional design principles. At the same time, the authors possessed contextual knowledge that enabled them to evaluate what the decision risks and information needs actually were, and thus the value to the project of obtaining that information.

Because the main goal of the project was to provide basic guidance to the state directors when and if they chose to conduct their studies, there was no expectation for them to "master" particular content objectives or complete the instruction within a particular time frame. Accordingly, drawing on the procedural steps summarized above, the purpose (Step 1) of the formative evaluation was defined as determining the reactions of experts in the content area (in SES and educational evaluation) and of potential users (state SES directors) to the quality and utility of the guidebook. The approach selected, therefore, was the *connoisseur-based study* in which expert feedback and suggestions were solicited and analyzed using a systematic process.

An important aspect of determining the formative study's purpose and general orientation was to involve the key stakeholders of the project in the decision-making process. In this case, there were two main stakeholder organizations: One was a federally funded center, The Center for Innovation and Improvement (CII), whose role was to provide research-based information to regional centers that assisted member states in improving their districts and schools. CII solicited and funded the present project. A second stakeholder was the SES directors from the U.S. Department of Education (USDE), who were responsible for approving any documents or information released under federal auspices. Without involving these stakeholders, the design team could have spent substantive time and resources on an evaluation approach that did not meet expectations in either its design or the evaluation questions identified. A tradeoff, however, is the possibility of receiving feedback that is contradictory or ill-advised, and incurring uncontrollable delays in completing the product.

With the assistance of CII and USDE, the study participants were identified and their assistance solicited. The experts specifically consisted of (a) an experienced instructional designer not connected with the project, (b) an educational consultant highly familiar with federal policies and the target SES program, (c) a "panel" of four educational researchers specializing in the evaluation methodologies proposed in the guidebook, (d) three SES state directors, and (e) the two federal directors of the SES program.

The audience for the formative evaluation (Step 2) was the design team and the two stakeholder groups identified above. The evaluation questions (Step 3) focused on dimensions of quality concerning the guidebook's accuracy of content, scope of coverage, readability, attractiveness, and utility or value to consumers. As indicated above, it was important to ensure that these questions were acceptable and of interest to the project stakeholders. The resources required (Step 4), evidence needs (Step 5), and data collection procedures (Step 6) were simple and straightforward, involved providing each expert with a copy of the guidebook draft and a series of questions to answer in writing or via a phone interview. Similarly, data analysis (Step 7) involved systematic compiling and informal coding of the feedback received relative to the evaluation questions posed in Step 3. Reporting (Step 8) involved documenting, primarily for internal use by the designers and sharing with the project stakeholders, the major results of the expert review and plan for revising the guidebook accordingly. The revised draft, in turn, was subjected to reviews by a smaller expert team. Based on the second round of feedback, the "completed" guidebook was released to all state directors.

Once an instructional design product is released, the need for evaluation continues and, preferably, never ends. That is, *summative* evaluation (see next section) will be needed to determine whether the product truly achieves its objectives when implemented with actual consumers. If not, use of the product should be discontinued unless improvements are made. Even if the product proves successful in the short term, *confirmative* evaluation will be needed to ensure that it is meeting objectives over time. In the case of the guidebook, its popularity among consumers spurred their increased involvement in conducting the required state SES studies. But the studies themselves revealed the need for greater assistance in certain areas (for example, determining which tutoring providers should be removed from approved listings) as well as more flexibility with federal expectations. These factors, constituting a naturalistic type of confirmative evaluation, resulted in a request by CII and USDE to supplement the guidebook with an updated brief on best practices.

SUMMATIVE EVALUATION

Relative to formative evaluation, summative evaluation shifts the focus to instructional products that are considered to be in ''relatively completed'' stages (see Table 16.1). However, we purposely equivocate in defining a product as ''developing'' or ''completed'' by recognizing that, in education and training, there would rarely be cause to define instruction as fixed and unchangeable. Even mature instructional products that have been popularly used over the years may lose relevancy and effectiveness as changes occur in educational contexts and consumer needs. In this sense, aren't virtually all instructional products constantly in ''developmental'' stages and, therefore, in need of *formative* evaluation (Komoski, (1974); Komoski & Woodward, 1985; Misanchuck, 1976; Weston, 1986)?

The answer to the somewhat rhetorical question above is probably ''most of the time,'' but differentiating between formative and summative (and confirmative) evaluations is more than semantics. An objective distinguishing factor is the defined status of the instruction by the designer, distributor, and/or user. Another factor concerns the stakes of adopting the product over other options, because once it moves from the designer's desk to the classroom, actual learners and instructors will be directly impacted. Stakeholders, including designers, funders, consumers, and administrators, will question the product's effectiveness and efficiency for the educational purposes intended. At this point, interest in product improvement (formative evaluation) becomes subjugated to evidence-based decisions (summative evaluation) regarding continued use of the product with regard to such factors as:

- Effectiveness in improving learning,
- Efficiency of learning,
- Cost of implementation,
- Ease of implementation, and
- Positive acceptance by learners, administrators, and other stakeholders.

These interests naturally make summative evaluations more formal and higher stakes than are formative evaluations. If a summative evaluation is part of a high-profile study, as was Borman, Hewes, Overman, and Brown's (2003) evaluation of comprehensive school reform models or Borman and D'Agostino's (1996) evaluation of Title I programs, a negative or positive outcome could have strong implications for a product's reputation and marketability in the eyes of many potential consumers.

Attributes of Summative Evaluation Methods

The methods used in a summative evaluation often resemble those of formative evaluation, but are generally more rigorous, credible (externally based), and oriented to specific questions related to product effectiveness.

Rigor. A summative evaluation is not necessarily a highly controlled ''research study'' using, for example, random assignment of learners to treatments to ensure equivalence of groups (Shavelson & Towne, 2002; Slavin, 2008). Such studies are often intended to support causal inferences about the effect size of programs or treatments on the outcome of interest. Summative evaluations are unlikely to afford or even need such high levels of precision. However, in the vast majority of instances, a summative evaluation should involve a ''research-type framework,'' so that the outcomes obtained, whether quantitative, qualitative, or mixed methods, constitute credible evidence for making judgments relative to instructional objectives or goals. Therefore, in general, summative evaluations will be more rigorous than formative evaluations with regard to study design, measures, and data analyses.

Credibility. For evidence to be credible, it must be judged as reliable and, above all, valid. As implied above, validity is increased by methodological rigor that helps ensure that results are attributable to properties of the instructional program rather than to extraneous factors, such as sampling bias or other learning experiences. For example, if positive learning outcomes were attributable potentially to the treatment group (that is, program users) being higher in ability than the control group, the summative evaluation might well support the erroneous conclusion that the instructional product (rather than student characteristics) engendered improved learning.

Another aspect of credibility concerns the identity of the evaluator. Whereas there are good reasons for designers to be involved directly in conducting *formative* evaluation (that is, to experience first-hand early product results), concerns about the objectivity of evidence might be raised if the designer had a central role in his or her product's summative evaluation. The higher stakes of summative relative to formative evaluations have already been discussed several times. Therefore, we'd typically expect a third-party (independent) evaluator to be the sole or primary investigator of the study.

Orientation to evaluation/research questions. Summative evaluations should be explicitly oriented to address specific evaluation questions concerning product effectiveness. Examples of such questions will be reviewed below in a discussion of summative evaluation processes and description of a case study. Although formative evaluations are also oriented around specific purposes and questions, they tend to be more open-ended and exploratory in nature.

Exemplary Summative Evaluation Approaches

Where instructional goals focus on objective or quantitative outcomes, such as achieving proficiency on an achievement assessment, expressing positive attitudes toward learning on a rating scale, completing a lesson within a certain time limit, or implementing a training course within a prescribed budget, the summative evaluation study will most commonly draw from two types of designs. One type involves mostly *descriptive* analyses for which the required data are collected (on cost, completion time, student/instructor attitudes, and so forth) from representative application contexts without necessarily requiring a comparison (control) group.

Example. A designer develops a web-based tool that offers advisement (such as help with interpreting figures and graphs) to managers as they complete an online course on financial operations of the company. The summative evaluation collects data during an eight-week period on cost, instructor perceptions, and student usage of the support using a random sample of twelve offerings of the course across international boundaries of the corporation. Results show that set-up and implementation costs meet the corporation's expectations and needs. However, in questionnaire and survey responses, instructors questioned the value of the support, feeling that it was more confusing than helpful. Student usage data, in turn, showed that only a small number of the students actually selected the support. On the basis of the summative study, it is decided to reconsider use of the program beyond the current year.

A second common type of summative study uses a quasi-experimental design to determine learning outcomes for treatment (product using) students relative to comparison students. In highly rigorous research studies, the ''gold standard'' design is commonly considered to be a randomized experimental study in which learners are assigned to treatment and control groups at random (Boruch, 2006; Shadish, Cook, & Campbell, 2002). In applied contexts, which provide the desired real-world venues for summative evaluations, random assignment may not be feasible or may even change the natural conditions being sought. That is, it is much more common for training programs, schools, or individual teachers to elect to purchase or try out a new instructional program than to volunteer to be randomly selected to use or not use it. Quasi-experiments involve comparing pre-existing or self-selected user groups to similar control groups.

*Exampl*e. An instructional design team creates a hybrid fifth-grade mathematics course that combines classroom instruction by the regular teacher with online learning. In a large school district, twenty fifth-grade teachers volunteer to implement the new course. Over a two-year period, the evaluators compare the achievement of students taking the hybrid course to similar students taking the regular course. Based on significantly higher scores by the hybrid group on the state math assessment, along with positive reactions by teachers, students, and parents, the district decides to continue the course and commission hybrid designs for additional grades.

Summative Evaluation Processes

Exhibit 16.1 summarizes basic steps or processes used in summative evaluations. In brief, similar to a research study, the summative evaluation revolves around specific evaluation questions, typically mirroring program objectives. That is, for each evaluation question, the evaluator will need to identify an appropriate design (descriptive or experimental) and instrumentation (achievement tests, questionnaires, interviews, etc.) that are consonant with time constraints, budget, resources, and conditions. Data collection methods will often depend on what information sources are accessible. According to Morrison and his colleagues, the following data sources often prove available and valuable:

- *Direct Testing.* Collect new data by administering appropriate assessments to learners and other participant groups (instructors, administrators, parents).
- *Analysis of Naturally Occurring Events.* Collect data that routinely becomes available from learners or other participants, such as attendance records and state-mandated test scores from K-12 students, plant safety records, accounting records, number of traffic violations by driver safety students, number of medical insurance claims processed in a thirty-day period, etc.
- *Direct/Indirect Observations.* Collect data via observations of learner performance or use of the instruction conducted by the evaluators or by

individuals naturally present in the application context. For example, supervisors use a systematic observation protocol to observe and rate trainees' skill in applying new strategies for recording inventory.

- *Portfolios/Exhibitions.* Data is collected from collections of work that demonstrate the learners' or trainees' skills associated with the instructional product; for example, budgets prepared by company accountants who completed an online training unit are examined for accuracy.

Exhibit 16.1 Major Steps in Summative Evaluations

1. Specifying evaluation questions

   Example. Did student learning increase as a result of the situated prompting?

2. Selecting the evaluation design for each question

   Example. Quasi-experimental comparison between classes of students using the prompted program and similar classes using the regular program.

3. Selecting/designing data collection instruments for each question

   Example. An achievement test on the material covered by the prompted and regular programs.

4. Conducting the evaluation

   Example. Students in ''program'' and control classes are pre-tested on the instructional unit. Following completion of the units, post-tests are administered to all students, along with surveys and interviews to those in the program group.

5. Analyzing results

   Example. Achievement test results are compared for program and control students, using analysis of covariance. Interview responses are analyzed via qualitative methods to extract key themes.

6. Interpreting results

   Example. Based on the achievement outcomes, students in the program group demonstrated superior learning, thereby supporting the situated prompting.

7. Disseminating results

   Example. A final report is prepared and distributed to the school district, program funders, and the instructional designers.

Adapted from Morrison, Ross, and Kemp, 2007

Depending on the information source and type of data collection instrument employed, a variety of types of measures may be used. Examples consist of objective test questions (such as multiple-choice), constructed-response or essay questions, checklists, rating scales, rubrics, or qualitative (subjective) impressions based on interviews or observations. Results from each data source will be analyzed and interpreted using rigorous quantitative or qualitative methods. Dissemination will almost always take the form of a written report that is broadly distributed, as in a national study of a product's effectiveness (Borman, Slavin, Cheung, Chamberlin, Madden, & Chambers, 2005) or restricted to a particular consumer (perhaps a school district) for review and decision making (Dessinger & Moseley, in press).

Case Study

Recently, the first author served on a team of evaluators contracted to evaluate a statewide technology integration program, labeled "Freedom to Learn (FTL)" (Lowther, Strahl, Inan, & Bates, 2007), implemented in Michigan. This evaluation could be considered summative because the program was in its final year of funding and included established professional development and implementation materials and procedures.

In accord with the procedural framework presented in Exhibit 16.1, the evaluation was oriented around multiple evaluation questions dealing with various expected outcomes of the program: (a) increasing students' skills in applying technology to learning; (b) increasing teachers' skills and comfort in integrating technology with classroom learning; (c) shifting pedagogy from teacher- to student-centered learning; (d) engendering positive attitudes toward the program by technology coaches, students, and parents; and (e) improving student achievement on state assessments. For examining the student achievement and student skills outcomes, the evaluators used a quasi-experimental design in which student performances on associated tests were compared to those of matched comparison students enrolled at non-program schools. For examining pedagogical change, they employed a modified quasi-experimental design that compared the frequency of observed teaching strategies to national norms compiled for the same observation instruments. For assessments of participant attitudes, they implemented descriptive (non-experimental) designs involving the administration of interviews and surveys.

Data analysis, in turn, encompassed (a) inferential statistical comparisons (for example, analysis of variance and analysis of covariance) for examining student and teacher outcomes relative to control groups or norms; (b) descriptive summaries and bar charts of survey responses; and (c) qualitative analysis to identify major themes expressed in interviews and open-ended survey items. Results, in brief, confirmed the success of FTL in achieving objectives of increasing (a) students' skills in using technology as a learning tool; (b) teachers' positive attitudes toward, and skills in, teaching technology integration; and (c) the frequency with which student-centered and higher-order learning occurred in classrooms. However, the program failed to attain the objective of raising student achievement in either reading or mathematics on the Michigan Educational Assessment Program (MEAP).* With additional time and budget, an enhanced study could have focused over a multi-year period on how the nature of student learning and performance changed, particularly in the direction of enhanced problem solving, self-efficacy, collaborative skills, and, in general, ability to use technology on educational tasks. Dissemination of results took place via a formal evaluation report and several informal presentations to different stakeholder groups. In fact, based on the evaluation results, interest developed among state legislators in seeking funds for its continuance.

* Although raising student achievement on the state assessment was considered an educationally and politically necessary "objective," most stakeholders (program administrators, teachers, and principals) did not expect technology integration to have direct immediate effects on high-stakes standardized tests (see Baker, 2007).

CONFIRMATIVE EVALUATION METHODS

General Characteristics

Confirmative evaluations use essentially the same orientations and methodologies described for summative evaluations (see Exhibit 16.1), but their purpose and timing differ (Dessinger & Moseley, 2004; Misanchuck, 1976; Moseley & Solomon, 1997). Simply put, the purpose now is no longer to test the effectiveness of a recently developed or implemented instructional product. Presumably, a summative evaluation has already been completed and yielded positive results. But once effective doesn't mean always effective, as is attested by the fact that most of us no longer write with manual typewriters, cut our lawns with push mowers, or watch our favorite shows on black-and-white TVs. Nor may we demonstrate excellent retention of those periodic table elements from chemistry or the Revolutionary War battles that we mastered back in high school. Confirmative evaluations, therefore, examine the effectiveness of instruction over time.

There are two basic types of situations that warrant the conduct of confirmative evaluation studies. One type is *learner-oriented* and concerns the degree to which, as time passes, consumers of the instruction retain the skills and knowledge needed to perform at desired levels. A second type is *context-oriented* and concerns the degree to which the instructional product remains effective as conditions (such as policies, politics, resources, technological advances) change over time. Each of these applications is examined in greater detail below and in Exhibit 16.2.

Exhibit 16.2 Two General Approaches to Confirmative Evaluation: Focus and Questions

---

Learner-Based Approach

Focus: What are skills, motivation, or interests of learners who received the instruction following the passage of time?

Exemplary Questions

- What percentage of learners demonstrate proficiency sixth months following training?
- What are learner reactions to the instruction after returning to the workplace?
- Do students who received the instructional program enroll in the advanced algebra course?
- How do teachers rate the abilities of students who completed the special course last year?

Context-Based Approach

Focus: What is the effectiveness of the instruction in achieving objectives following possible changes in conditions or policies as time passes?

Exemplary Questions

- How economically and efficiently can the instruction be implemented in the present context?
- Do teachers and students still view the instruction as valuable for learning?
- Do student learning gains associated with using the instruction continue to meet expectations?
- Are current organizational policies for employee development still supportive of the training approach employed?
- Do instructional outcomes continue to meet the needs of learners?

Learner-Oriented Confirmative Evaluations

Several years ago, the two of us were hired by a large chemical corporation to evaluate its employee training courses in areas such as communication, public speaking, and interpersonal relations. Although the corporation routinely conducted follow-up evaluations of employee skills and attitudes soon after participants completed the courses, there was strong interest in assessing longer-term impacts. For example, after completing the course on making oral presentations, the employees might do well on an immediate performance or knowledge test, but would they demonstrate the desired skills on-the-job a year later? If not, the need to strengthen the original course or offer refresher training would be implied. The confirmative evaluation that we conducted involved asking managers, peers, and subordinates to rate on-the-job performance on target skills of former trainees. A second measure was asking the trainees to rate retrospectively the helpfulness for fulfilling their job requirements of various course content and activities. From this data, the evaluation study identified course components that were successful for long-term achievement of objectives and those that were not.

Note that the measures employed in the above example and in many confirmative evaluations were fairly low-cost and coarse-grained indicators of continued program success. Had the results indicated major weaknesses in these continuing training courses, refinements in the course designs accompanied by more granular formative evaluation would have been implied. Again, formative, summative, and confirmative evaluations are not independent entities but operate in a mutually supportive interactive fashion.

Context-Oriented Confirmative Evaluations

But what if the desired objectives of instruction or training change over time? Obviously, a course or unit that was formerly judged effective (in formative and summative evaluations) would no longer be optimum for the educational needs concerned. The fact that learners might retain the skills taught over long periods of time would hardly matter if the skills mastered were no longer the ones desired. Thus, as summarized in Exhibit 16.2, another type of confirmative evaluation examines whether the instruction is achieving objectives following changes in conditions and policies as time passes.

For the most part, the context-oriented confirmative evaluation would use the methodology as outlined on Exhibit 16.1 for summative evaluations. In a prototypical case, the study would compare learners who receive the instruction to a control group several years after the instruction was originally released or adopted in the particular context. While stakeholders in the instruction would hope for continued supportive results, the evaluation may uncover that the program's effectiveness has diminished due to such factors as:

- Learner characteristics have changed. For example, students may no longer possess the prerequisite knowledge that former students had in the past due to changes in the school curriculum or enrollee ability levels.
- The curriculum or performance expectations have changed. For example, a program to help students prepare for taking a Microsoft certification exam may emphasize topics that are no longer covered.
- Technology has changed. For example, the program may require students to work with outdated computers or lab equipment.

- Budgetary support for the program is reduced. For example, due to reductions in funding, offering of the program is restricted with regard to time available per student.
- The original need addressed by the training no longer exists. For example, management identified a need for all employees to take a popular psychology course learning to classify individuals into one of four personality types. After several years, a new management team emerges and no longer places value on the process, but the course is still offered. A confirmative evaluation identified the lack of continued need for the course.
- Teacher support or preparation has changed. For example, given their concern about students' achieving proficiency on the state assessment, teachers are reluctant to use the instructional program due to the perception that it detracts from teaching to the test.

Case Study

Several years ago, the first author was part of a team asked to evaluate the continuing effects in the Toledo Public School District of two school reform models being implemented for improving teaching and learning (Ross, Nunnery, Goldfeder, McDonald, Rachor, Hornbeck, & Fleischman, 2004). The two designs—Success for All and Direct Instruction—have undergone extensive formative and summative evaluation over several decades and, in fact, are considered among the most thoroughly researched models in the history of education (Borman, Hewes, Overman, & Brown, 2003). Accordingly, this evaluation need was not to demonstrate that the models were ready for release or prove they could work, because both of these conclusions were already well-supported. Rather, consistent with the purposes of context-oriented confirmative evaluation, the important question was whether teachers and learners, within this particular setting and time, were benefiting from the models at the levels expected.

As desired for both summative and confirmative evaluations, credibility and rigor were promoted by employing third-party evaluators to conduct a comprehensive mixed-methods quasi-experimental design, examining data from test scores, surveys, interviews, and observations. Results confirmed what the school district expected. Due to funding reductions, diminishing teacher support, and political vying by certain community members and teacher union leaders, neither program was being strongly implemented and, not surprisingly, demonstrating clear or consistent success. Based on these results, the district decided to give the program supporters another year to show progress or the programs would be discontinued.

SUMMARY

Uses of evaluation in instructional design were discussed and analyzed in this chapter. Key ideas are summarized below.

- Evaluation is using assessment or measurement to judge the worth or value of something. Three types of evaluation are formative, summative, and confirmative.
- Formative evaluation is used to provide feedback to designers as the instruction is ''forming'' or being developed.
- A connoisseur-based formative evaluation collects impressions and feedback from experts regarding the quality of the instruction. A decision-based study is directed to answer questions that are of specific interest to the designer. An objectives-based study assesses outcomes relating to specific instructional objectives. Public relations–inspired studies make preliminary evaluation results known to targeted individuals, such as key stakeholders and funders, to raise interest.
- Formative and most other evaluations are systematic studies that generally include the steps of (1) determining the purpose, (2) identifying the audience, (3) defining the evaluation questions/objectives, (4) determining resource needs, (5) defining evidence needs, (6) collecting data, (7) analyzing data, and (8) preparing and disseminating a report.
- Summative evaluation is conducted to determine the degree to which a *completed* instructional product produces the intended outcomes.
- Relative to formative evaluations, summative evaluations are more formal and higher stakes, with greater emphasis placed on rigor, credibility, and orientation to specific evaluation or research questions.
- The steps involved in summative evaluations include specifying the evaluation questions, selecting the evaluation design for each question, selecting/designing data collection instruments for each question, conducting the evaluation, and analyzing, interpreting, and reporting results.
- Data sources for summative evaluation may include direct testing, analysis of naturally occurring events, direct/indirect observations, and exhibitions and portfolios.
- Confirmative evaluation examines the success of instruction as it is used over time.
- Learner-oriented confirmative evaluations examine the degree to which, as time passes, consumers of the instruction retain the skills and knowledge needed to perform at desired levels.
- Context-oriented confirmative evaluations examine the degree to which the instructional product remains effective as conditions (that is, policies, politics, resources, technological advances) change over time.
- Formative, summative, and confirmative evaluations, although differentiated with regard to their timing, design rigor, and complexity, and the

maturity of the target instructional product, use similar evaluation methods. Because instructional products rarely attain completed, unchangeable forms, and because educational conditions are continually evolving, these major evaluation categories are often overlapping and applied simultaneously to both test and improve products.

References

Baker, E. L. (2007). The end(s) of testing. *Educational Researcher*, *36*(6), 309–317.

Borman, G. D., & D'Agostino, J. V. (1996). Title I and student achievement: A meta-analysis of federal evaluation results. *Educational Evaluation and Policy Analysis*, *18*, 309–326.

Borman, G. D., Hewes, G., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research*, *73*(2), 125–230.

Borman, G. D., Slavin, R. E., Cheung, A. C. K., Chamberlin, A. M., Madden, N. A., & Chambers, B. A. (2005). The national randomized field trial of Success For All: Second-year outcomes. *American Educational Research Journal*, *42*(4), 673–696.

Boruch, R. (2006, April). Ethical standards, evidence standards, and randomized trials: Error flees, but slowly. Paper presented at the annual meeting of the American Educational Association, Chicago.

Clarebout, G., & Elen, J. M. (2008, March). Help-seeking and advice: A complex relationship? Paper presented at the annual meeting of the American Educational Research Association, New York City.

Dessinger, J. C., & Moseley, J. L. (2004). *Confirmative evaluation: Practical strategies for valuing continuous improvement*. Hoboken, NJ: John Wiley & Sons.

Dessinger, J., & Moseley, J. (Eds.) (in press). *Handbook of performance measurement and evaluation*. San Francisco: Pfeiffer.

Dick, W., Carey, L., & Carey, J. O. (2005). *The systematic design of instruction* (6th ed.). Boston: Allyn and Bacon.

Flagg, B. N. (1990). Formative evaluation for educational technologies. Hillsdale, NJ: Lawrence Erlbaum Associates.

Friesen, N., & Anderson, T. (2004). Interaction for lifelong learning. *British Journal of Educational Technology*, *35*(6), 679–687.

Gooler, D. D. (1980). Formative evaluation strategies for major instructional development projects. *Journal of Instructional Development*, *3*, 7–11.

Kay, R., & Knaack, L. (in press). Assessing learning, quality and engagement in learning objects: The learning object evaluation scale for students (LOES-S). *Educational Technology Research & Development*.

Komoski, P. K. (1974). An imbalance of product quantity and instructional quality: The imperative of empiricism. *Audio Visual Communications Review*, *22*(4), 357–386.

Komoski, P. K., & Woodward, A. (1985). The continuing need for the learner verification of textual materials. In D. H. Jonassen (Ed.), *The technology of text: Volume 2* (pp. 396–415). Englewood Cliffs, NJ: Educational Technology Publications.

Krauss, F., & Ally, M. (2005). A study of the design and evaluation of a learning object and implications for content development. *Interdisciplinary Journal of Knowledge and Learning Objects, 1*. Retrieved March 15, 2008, from http://ijklo.org/Volume1/v1p001-022Krauss.pdf.

Lowther, D. L., Strahl, J. D., Inan, F. A., & Bates, J. (2007). *Freedom to Learn: 2005-2006 evaluation report*. Memphis, TN: Center for Research in Educational Policy.

Misanchuck, E. (1976). Descriptors of evaluations in instructional development: Beyond the formative-summative distinction. *Journal of Instructional Development*, *21*(1), 15–19.

Morrison, G. R., Ross, S. M., & Kemp, J. E. (2007). *Designing effective instruction* (5th ed.). San Francisco: Jossey-Bass.

Moseley, J. L., & Solomon, D. L. (1997). Confirmative evaluation: A new paradigm for continuous improvement. *Performance and Instruction*, *36*, 12–15.

Nurmi, S., & Jaakkola, T. (2006). Problems underlying the learning object approach. *International Journal of Instructional Technology and Distance Learning*, *2*(11). Retrieved April 1, 2008, at http://www.itdl.org/Journal/Nov_05/article07.htm.

Ross, S. M., Nunnery, J. A., Goldfeder, E., McDonald, A., Rachor, R., Hornbeck, M., & Fleischman, S. (2004). Using school reform models to improve reading achievement: A longitudinal study of direct instruction and Success for All in an urban district. *Journal of Education for Students Placed at Risk*, *9*(4), 357–388.

Ross, S. M., Potter, A., & Harmon, J. (2006). *Evaluating supplemental educational service providers: Suggested strategies for states*. Memphis, TN: Center for Research in Educational Policy, American Institutes for Research, and Academic Development Institute.

Saab, N., Gijlers, H., Van Joolingen, W. R., & van Hout-Wolters, B.H.A.M. (2008, March). Using or not using the provided tool during collaborative inquiry learning. Paper presented at the annual meeting of the American Educational Research Association, New York City.

Scriven, M. (1967). The methodology of evaluation: In R. Tyler, R. Gagné, & M. Scriven (Eds.), *AERA monograph series on curriculum evaluation* (pp. 39–83). Chicago: Rand McNally.

Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Thousand Oaks, CA: Sage.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Shavelson, R. J., & Towne, L. (Eds.). Scientific research in education. Washington, DC: National Academy Press.

Slavin, R. E. (2008). What works? Issues in synthesizing educational program evaluations. *Educational Researcher*, *37*(1), 5–14.

Stufflebeam, D. L. (1983). The CIPP model for program evaluation. In G. F. Madaus, M. Scriven, & D. L. Stufflebeam (Eds.). *Evaluation models* (pp. 117–141). Boston: Kluwer-Nijhoff.

Weiss, C. H. (1998). *Evaluation* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall.

Weston, C. (1986). The importance of involving experts and learners in formative evaluation. *Canadian Journal of Educational Technology*, *16*(1), 45–58.

PART FIVE

MANAGEMENT

This is another part of this handbook that experienced instructional designers are likely to come to first. Its focus goes beyond doing ID to managing the doing of ID—designing and delivering the course.

As anyone who ever tried to implement all the great ideas this volume has presented in Chapters Four through Sixteen knows, the road to implementing ID principles is anything but a smooth one. Even if ID is a set of principles rather a process, the project to develop a course is indeed a process and, as such, must be managed like any other business process. Running throughout the principles (from analysis through evaluation) is the need for the ID to work with other people in accomplishing the development of a course. And, to paraphrase the philosopher Bishop Berkeley, ''If a course is developed but no one takes it, is it really a course?'' Once the course is developed, the logistics of having learners interact with and learn from the course must be dealt with.

Chapter Seventeen: Managing ID/Training Development and Delivery. This chapter begins with the authors describing a process for managing an ID project. Using a unique and interesting visual approach, as well as traditional project management techniques, they describe what project management is, how it relates to ID, and what the various tasks and roles that must be performed are. They walk through the management of an ID project step-by-step, providing guidelines to follow and tools to use.

Chapter Eighteen: Managing Relationships in the Performance Improvement Process. This chapter focuses on the people side of an ID project. Although everyone knows ID projects do not happen in isolation, but rather involve interaction among many different people playing many different roles, there are very few ID books that even address this topic at all. Ranshaw combines her own unique insights into people and relationships with those of Peter Block to explain the do's and don'ts of working with others. She explains the relationship issues that arise, how to prepare yourself for them, and how to handle them when things go—well, let's just say amok.

Chapter Nineteen: Managing ID in the Context of a Training Organization. The authors of this chapter provide a detailed, step-by-step blueprint for managing a training function in an organization and delivering those courses that have been designed. From finding instructors and facilities, to duplicating course materials, to hotel facilities, to the future of the training function in the age of e-learning, this most comprehensive chapter provides details of how to implement every aspect of the training organization.

CHAPTER SEVENTEEN

Managing ID/Training Development and Delivery

Deborah Singer Dobson
Michael Singer Dobson
Ted Leemann
Kevin Forsberg

INTRODUCTION

The purpose of this chapter is to help you apply the tools and methodologies of project management to the process of training design and delivery, based primarily on Kevin Forsberg, Hall Mooz, and Howard Cotterman (2005).

Projects are characterized as work that is temporary and unique. *Temporary* in this context means that the work has a definite beginning, middle, and end; *unique* means that a particular project differs in important ways even from projects that are similar in subject matter and work process. Operations, on the other hand, are ongoing, with no planned end-point. Work is the sum of operations and projects.

Operations and projects often live in symbiosis. A training design and delivery program is an operation. There's no planned point at which you intend to stop designing and delivering training products. Each individual training product, however, is a project. The goal is to finish. In fact, most projects have no significant value until they are finished. Different values, methods, and processes apply, depending on whether you're managing an operational environment or a specific project.

The project management process is a highly developed, standardized, and validated approach that allows people managing projects in very disparate fields to approach the complex management issues involved in a way that is both systematic and effective. However, a ''one size fits all'' approach is doomed to failure. Project management carries an overhead cost, and must therefore be tailored to the size, complexity, and risks of a given category of projects.

The process described in this chapter is our attempt to ''right-size'' project management for the training environment. Of course, the size of training organizations and the complexity of needs vary, so it's important to state up-front that you may need a bit more or a bit less than presented here. In addition to understanding the fundamental project management approach, you should be able, after reading this chapter, to analyze your own project management issues and make appropriate choices to gain maximum benefit while minimizing overhead cost.

THE PROJECT ENVIRONMENT

Project management complexity and difficulty should be thought of separately from technical complexity and difficulty. Work can be enormously complex and difficult from a technical perspective and yet be relatively straightforward from a project management perspective, or vice versa.

In the training environment, project management complexity and difficulty arise from various causes. In determining what type, level, and degree of project management process is desirable for a given function, you must determine specific challenge areas and their associated issues. The factors cited in Table 17.1 complicate any attempt to perform formal project management, adding overhead and cost when our objective must be to keep project management overhead and cost at the lowest effective level.

Is project management scheduling software, such as Microsoft Project®, necessary or appropriate? If your training program is at any given moment running thirty projects scheduled over a staff of twenty people, with multiple deliverables and numerous tasks per project, you may find that using software and even employing a full-time scheduler will improve your operation sufficiently to warrant the substantial costs involved (Dobson, 1999, pp. 69–83).

If, on the other hand, you're looking at one or two projects at a time, scheduled across three people, with six or seven work packages per project, graph paper and sticky notes are more appropriate technology. One size definitely doesn't fit all.

Both projects and operations (work) normally take place within the context of an organization. Perhaps you're in charge of the training department for a corporation, or perhaps you're managing a training consulting firm. In the first

Table 17.1 Project Management Complexity and Difficulty Factors in the Training Design and Delivery Environment

Challenge Area	*Project Management Issues*
Single vs. Multiple Projects	Classical project management involves optimizing a single project. Multiple projects must be prioritized within a limited resource environment.
Intact Work Teams vs. Shared Resources	Classical project management assumes all key team members are available exclusively for project needs. In the more common shared resource environment, people may be supporting multiple projects simultaneously.
Known Duration vs. Variable Duration	If you need to know how fast a brick wall can be built, you can determine the speed of individual bricklayers and add additional bricklayers to the project to make it go faster. The length of time necessary for training concept design, however, is much less easily quantified, and adding people has been known to increase, rather than decrease, duration.
Emergent Work	Although a training department may construct an annual plan for the training programs it plans to design and deliver, user needs and emergencies commonly trump the plan, necessitating major shifts in priorities and planned resource allocation.
New Technologies/Tools	In the fast-changing world of instructional design and training, new technologies, new tools, and new methodologies must be integrated into existing operations, even though you and your staff may not yet have substantial familiarity and experience with them.
Size/Priority/Importance	Other important variables to consider are relative size of the project (greater/smaller than normal), priority (strict deadline, unusual urgency), and organizational importance/visibility.

instance, training is normally only one of many organizational functions being performed, and usually exists in a staff, rather than line, role. In the second instance, training is the main product and is considered a line activity. However, the consulting organization still performs marketing, financial management, procurement, and human resource functions. Regardless of whether training is the central product or serves to support line operations, it must share resources and priorities with other organizational functions. Projects may have the right of way, but they don't ever have exclusive use of the entire road.

Integrating project management thinking into an environment that thinks primarily in terms of operations requires all levels of management to visualize the big project picture. Only then can they meaningfully apply the tools and techniques to your individual projects and achieve the desired results.

PROJECT MANAGEMENT IN TRAINING ORGANIZATIONS

The mechanics of a project don't, obviously, exist in a vacuum. The Project Management Body of Knowledge (PMBOK) process identifies organizational issues affecting the project, but its primary focus and classic tools are concerned with the traditional project.

For reasons of space, we cannot delve sufficiently into all relevant aspects of management, politics, motivation, negotiation, and human behavior. Suffice it to say, as challenging as are the operational details of project management, it's the people and politics that get you every time.

Visualizing Project Management (Forsberg, Mooz, & Cotterman, 2005, pp. 23–24) adds the necessary big picture of how projects work and are managed within organizations, balancing strategic and tactical considerations to develop an overall systems approach, and allows addition of advanced techniques such as agile development.

Too many organizational decisions consider only tactical issues. To achieve organizational balance, a systems approach is essential. Projects have numerous stakeholders, each with varying interests, desires, needs, and concerns. Projects always and necessarily involve risks, both positive (opportunities) and negative (threats). Successful projects require five essential elements, as shown in Exhibit 17.1.

In thinking about projects and project management, it's valuable to distinguish between practices that are continuous throughout the project cycle and those that are situational, limited to specific periods, phases, or activities. These essentials are always present, as distinct from the specific tools and techniques (Gantt charts, the Work Breakdown Structure, Microsoft Project) that most people associate with project management. All the Gantt charts in the world won't make up for a lack of leadership and teamwork.

You can visualize the relationships among these project management essentials by using a wheel and axle model (Forsberg, Mooz, & Cotterman, 2005, pp. 23–24). The axle contains the three elements of the project cycle (Exhibit 17.3) in its center, and the phases of the projects along its length, as shown in Figure 17.1.

As shown in the figure, in a project, the business, technical, and budget aspects form the core. The phases of a project typically include a study period,

Exhibit 17.1 Five Essentials for Every Project

The process model for a successful project team is based on these five essential elements.

1. **Organizational commitment.** The foundation for the project. Includes culture, team charter, financial and other resources, and tools and training to support effective and efficient execution.
2. **Communication.** Ability to achieve understanding within a group, enabling team members and stakeholders to interact effectively and function as a team.
3. **Teamwork.** Efficiently working together to achieve a common goal, with acknowledged interdependency and trust, acceptance of a common code of conduct, and a shared reward.
4. **Project cycle.** The process for managing the project, from both a strategic and tactical perspective, performed in periods and phases punctuated by decision events. The project cycle usually starts with the identification of needs and ends with the disposition of project products. Includes the three aspects of business, budget, and technical.
5. **Management elements.** The ten categories of interactive management responsibilities, techniques, and tools that are situationally applied throughout the project cycle by all stakeholders.

an implementation period, and an operations period, though the detailed project cycle can vary.

The wheel consists of the ten management elements (Item 5 in Exhibit 17.1). These elements are listed in Exhibit 17.2.

Nine of the management elements are depicted as the spokes of a wheel, the tenth (project leadership) being the rim that holds the structure intact, as shown in Figure 17.2.

Techniques and tools are located within the element where the benefit is most significant. Although not all ten elements are in active use at any given

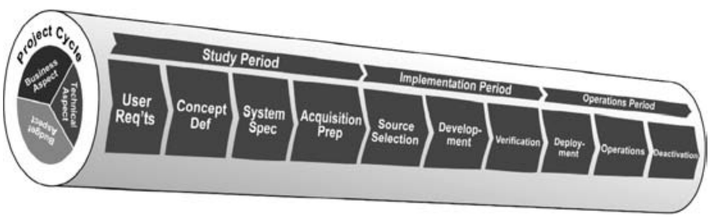


Figure 17.1 The Project Cycle Portrayed As an Axle.

Exhibit 17.2 Ten Management Elements

The ten management element groups are applied situationally through the project cycle. They include the hundreds of tools and techniques of project management and provide a structure for where and when those tools and techniques apply.

1. Project requirements
2. Organizational options
3. Project team
4. Project planning
5. Opportunities and risks
6. Project control
7. Project visibility
8. Project status
9. Corrective action
10. Project leadership

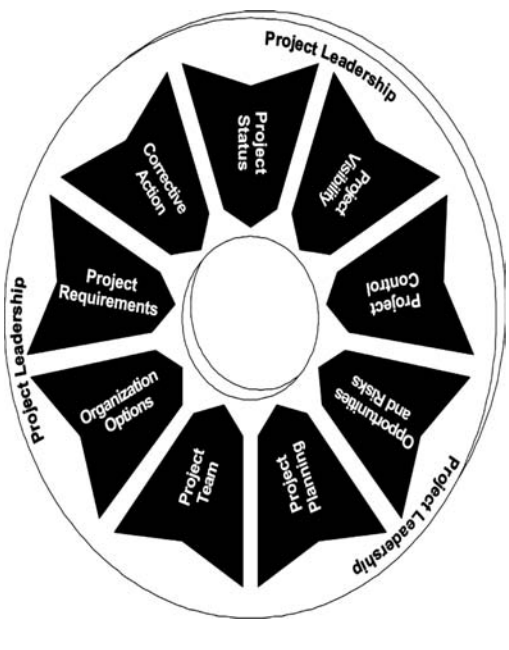


Figure 17.2 Management Elements Depicted As a Wheel.

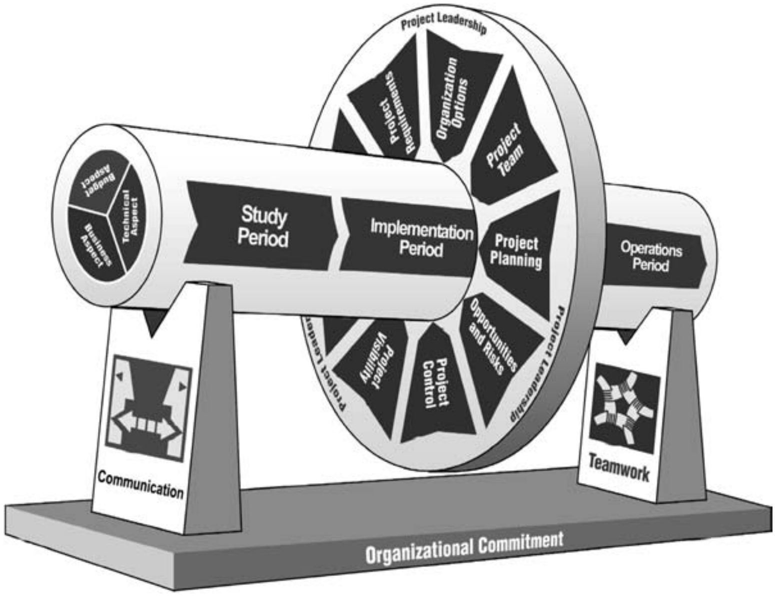


Figure 17.3 The Wheel-and-Axle Model of Project Management.

project moment (except for project leadership, which is an ever-present necessity), they are all part of the project cycle from beginning to end.

Communication and teamwork support the axle, and a base of organizational commitment makes everything else possible, as shown in the final wheel and axle model illustrated in Figure 17.3.

The wheel progressing along the axle represents the project's logical sequence of events. Rotating the wheel represents the dynamic selection and application of the techniques and tools appropriate to the project situation at a given time.

This sequential project cycle and situational management wheel in Figure 17.3 are supported by the ever-present piers of communication and teamwork, resting on a foundation of organizational commitment. Without a solid foundation, this model collapses, just as do real projects when management support and structure are inadequate.

Agile Development/Iterative Development

There are different approaches to managing projects, shown in Figure 17.4. Traditional project management is built around an approach known as the ''waterfall'' model. In the waterfall approach, Task A has a discrete beginning and a discrete end, and Task B requires the completion of Task A in order for it to start, and so forth. To put a swimming pool in my back yard (simplified), I must

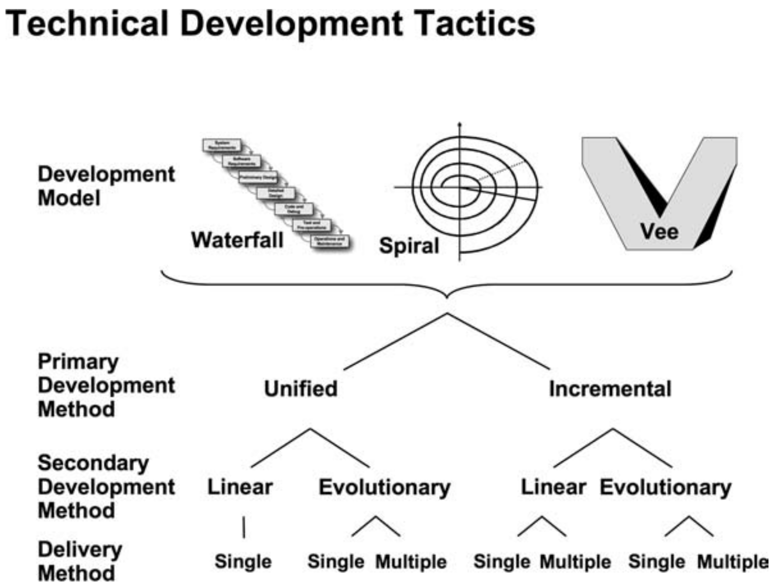


Figure 17.4 Different Approaches to Project Management.

(a) dig a hole, (b) pour the concrete, (c) fill the pool with water. I can measure start and stop of each task, and the sequence of actions is non-negotiable.

Traditional project management approaches are known as ''waterfall'' projects, because the output of each step serves as the input to the next step, as a waterfall takes the output of one body of water and feeds it into another body. Modern agile development uses Spiral, Dual Vee, and other approaches.

Classic engineering and construction projects fit well with traditional project management tools and the sequential ''waterfall'' approach. Workers are specialized, and many jobs cannot begin until some predecessor job is completed.

Waterfall is an example of a *linear development* technique. Linear development is a single path approach for which the requirements and the solutions are sufficiently well understood to allow straightforward design and implementation without iteration or experimentation, such as installing electrical and plumbing in a home. Some categories of technical or skills training may be appropriate for linear development, but other approaches are more common and usually more effective in the training development environment.

Creative and collaborative projects, from software to advertising to training, don't fit as well into the waterfall model. Do we really finish all the research before we start writing? Are all the learning objectives set in final form before we design the training materials? Will everything be complete in time for the beta, or will we have to make do? Do we wait until everything's completed before we present and review?

We know the answer: In every case, it's ''no.''

These processes have unclear starts, unstable definitions of ''finished,'' and overlap in ways that cannot be predicted or set in planning concrete. ''Spiral'' and ''Vee'' approaches are examples of *agile development*, in which work overlaps, circles, and interacts in ways difficult to depict using traditional project management tools. Agile development practitioners have developed a comprehensive set of tools for faster development in these flexible environments, focusing on the key attributes of *velocity* and *adaptability*. The wheel-and-axle process model presented here provides a tailoring framework to adapt traditional project management to modern processes.

The training environment often uses an agile approach known as *evolutionary development*, appropriate when experimentation or investigation is necessary to determine the best solution. It works well when requirements are uncertain, when you might end up pursuing opportunities or alternate concepts, or when you have many business risks (opportunity + threat) to consider.

The key consideration in selecting the appropriate methodology is whether you'll make a *single* delivery, turning over an actual final product and being done, or whether you'll make *multiple, iterative* deliveries. In the first case, you do the iterations inside the project, and in the second, each iteration turns over a usable product to the customer, which will be improved for the next edition.

Visual models exist for these approaches, but they're designed to help you understand more than to help you manage. If you feel you need tools such as Microsoft Project, how will you define the tasks crisply enough to give meaningful information without pretending to a rigidity that doesn't exist? Here are a few suggestions.

If the deadline is reasonably strict, or if certain events (a beta test) have to be fixed in the calendar, you can organize the work around milestones. ''From January 21 to February 15,'' you might say, ''We'll make the revisions based on the results of the first beta. I've booked the training room for the 16th, so we have to be ready to teach a class of sixteen. That means materials have to go to production by the 12th, giving us the 13th and 14th for lesson planning and the 15th for insurance in case anything slips.''

To organize the seventeen available work days, prepare a list of the deliverables—the actual items that must be complete for the next beta test. Workbooks, PowerPoint slides, handouts, toys for the group activity, evaluation forms—make sure the list is complete. Make someone on the team the accountable lead for each piece, and keep the master checklist for yourself.

Easier said than done? Absolutely. As important as an organized timeline is, people and organizational issues form a greater concern. That's why agile development organizations must adopt some key practices and values to make it work effectively.

Seven Key Practices of Agile Development

There are seven key practices an organization must follow to use an agile approach successfully, shown in Exhibit 17.3.

In practice, this requires excellent management discipline. People tend not to check their humanity at the door when they punch out on the time clock. Personalities, egos, preferences, power imbalances, conflicting goals, and personal agendas are hard to keep completely at bay. Operationally, the

Exhibit 17.3 Seven Key Practices of Agile Development

By adopting the following seven key practices of agile systems engineering, any organization can improve its velocity (speed + direction) to customer satisfaction.

1. The project team understands, respects, works, and behaves within a defined development process, which is systemic in the organization and implicit to all participants.
2. The project is executed as quickly as possible with minimum downtime or staff diversion during the project. Every opportunity is exercised to move the project forward, especially for the critical path activities.
3. All key players are physically or electronically co-located. Other contributors are available online 24/7.
4. There is a strong bias for automatically generated electronic documentation. Artifacts for operations and replication are done only if necessary—not to support an existing bureaucracy or policy. Documentation, formal and informal, are team property and available to all.
5. Baseline management and change control are achieved by formal, oral agreements based on "make a promise, keep a promise" discipline—participants hold each other accountable. Decision gate agreements are confirmed with a binding handshake. Formality relates to the binding of the action, not to the amount of documentation.
6. Opportunity exploration and risk reduction are accomplished by expert consultation and rapid model verification, coupled with close customer collaboration. Development is done in a rapid deployment environment. There is no resistance or inertia to securing expert help; it is sought rather than resisted.
7. A culture of constructive confrontation pervades the project organization. Issues are actively sought. Anyone can identify an issue and pass it on to the most likely solver. No issue is left unresolved. The team takes ownership for success; it is never "someone else's responsibility."

From Forsberg, Mooz, and Cotterman, 2005, pp. 352–355 project manager or department head behaves like a cowboy of old, gently pushing around the edge to keep the herd together and moving in the right direction. A gentle touch generally works better than a six-gun when you want people to be open to each other. Your own ability to let go of ego needs is a huge indicator of how successful you're likely to be.

THE PROJECT LIFE CYCLE

Although process details and the sequence of project management activities necessarily vary by project, the wheel-and-axle model (Figure 17.3) provides an overall context applicable to virtually every project environment.

As the wheel moves down the axle, the project moves from life cycle phase to life cycle phase, with three main sections.

During the *study period*, the project team determines user requirements, develops and defines an overall concept, writes specifications, and (if the project involves major outside procurement) prepares an acquisition strategy. In this initial phase, the project manager and team determine the preliminary scope of the project. Sometimes, customers and stakeholders provide the scope; other times, the project management team must survey stakeholders and users to determine preliminary scope. Either way, it's not uncommon for the preliminary scope to be incomplete, vague, or even wrong. Through the process of *progressive elaboration*, project team members and stakeholders provide mutual feedback to establish a definition of project scope that is both complete and correct. High-level planning begins in the study period and continues into the following phase, the implementation period.

The *implementation period* is when the project is developed, verified, and (if there's major outside procurement) the source is selected. In agile development, planning continues, becoming more refined as the work progresses.

Most of the traditional tools of project management are used during the implementation period. Planning involves identifying and scheduling the work packages or activities necessary to perform the work; constructing estimates of duration, cost, and resources; developing plans to ensure quality, manage risks, and control scope; establishing communications and reporting strategies; acquiring resources, both people and contracts; and integrating these elements into a comprehensive planning document.

These planning activities overlap with project execution through the development cycle. Activities include acquiring and building the project team, performing the work, and producing deliverables. The project manager and team must monitor and control the project, tracking conformance to plan, identifying discrepancies, handling change management, and providing feedback to update and progressively elaborate the plan.

Finally, the *operations period* takes the development from the project (building) stage through deployment, operations, and (when the training has been delivered) deactivation. As the project moves from phase to phase, the project team transfers deliverables to the next stage. This sometimes involves turning them over to customers or users, and other times involves operating or using the deliverables themselves. Either way, the project is completed, and the remaining work of the project team involves releasing resources to other projects, developing lessons learned for future improvement, and closing out the administrative elements of the project from an organizational perspective.

Study Period

The study period typically determines the scope, feasibility, and funding of a project: in other words, the project is defined and established. This process may be extremely formal or comparatively casual, but at some point on the road from ''We're just talking about maybe doing something at some point'' to ''Yes, we have a project! Why aren't you done yet?'' you need a signpost: ''You are now entering Project Land.'' (Some project managers would add a second sign, ''Abandon hope, all ye who enter here,'' but that's what you get in the absence of the five project essentials.)

From Problem to Project. In a training and development environment, a client might express concern that his or her team members are experiencing an unusual degree of conflict. This is not a project; it's a problem. It may become a project, but there are numerous decisions that must be made first.

After you have completed front-end analysis and other performance analysis, you will have a good perspective on what the organization needs. But now you face the problem of who and how that work will be done. Table 17.2 shows some of the factors you will need to consider.

The Triple Constraints. Because projects must always end, they are constrained in ways that don't neces-sarily apply to operations work. The project environment is bounded by the Triple Constraints (Dobson, 2004, pp. 7–12), shown in Figure 17.5. The balance among time, cost, and performance informs virtually every important decision about your project.

1. **Time Constraint.** How long do we have? The time constraint can be expressed as a specific deadline (before the end of the fiscal year), an event trigger (before the failure rate hits 10 percent), or a degree of urgency (''I need it yesterday!'').
2. **Cost Constraint.** How much can we spend? The cost constraint can be expressed as an amount of money, a number of person-hours, use of

Table 17.2 Defining the Initial Project Space

Is it in our jurisdiction?	Are we the appropriate organization to solve the problem? Are there other organizational units that have superior jurisdiction over this case?
Is it within our capabilities and expertise?	Do we have the knowledge? Do we have the skills? Can we take on this workload?
Does the client want us to solve the problem?	Does the client trust us? Is there a benefit to the client or other key stakeholders for keeping the problem alive rather than solving it? Does the client believe we have the capabilities and experience to solve this problem?
Is there funding to solve the problem?	Do we have the resources in our budget? Is the client expected to pay us to perform the necessary service? Is the price affordable, competitive, and proportionate?
Are we prepared to handle the risks involved?	If we've identified specific areas of risk and concern, can we handle them? Do we need to modify other project elements to cope with identified risks? Are there specific areas (see Table 17.1) that must be taken into account?
Should we accept this assignment?	What are the consequences for us, both positive and negative, in taking on this project? If the financial consequences are not attractive, are there long-term considerations that should influence our decision? Is the decision in our hands, or is it made by our organizational superiors?
Should we do this work in-house or contract it to someone else?	Is it faster, cheaper, or better to choose one option over the other?
To whom should we assign it?	Who has the capability, expertise, and capacity to take on the responsibility for managing this project? Can we supply the necessary management support to someone who may not have all the desired qualifications?

equipment, consumption of supplies, or intangibles (political capital, goodwill).

3. **Performance Criteria.** What does the product have to do? Performance criteria can be expressed as functional and technical requirements, the

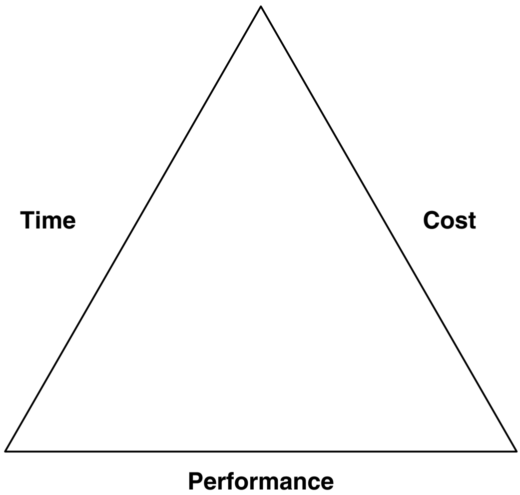


Figure 17.5 The Triple Constraints.

project's purpose or desired end state, evaluation criteria, or the establishment of the ''good enough'' point.

The three dimensions of the Triple Constraints exist in a hierarchy, based on the particular priorities or goals of the project, known as Driver/Middle/Weak Constraint.

Imagine the regulatory environment of your organization has changed drastically, and you must produce and deliver the training necessary to ensure five thousand people are in compliance with the new laws within no more than thirty days. If you're late, the organization will receive substantial fines for each day it fails to be in compliance. Time, in such a situation, would clearly be the project's driver. Performance—the need for the training content and methodology to ensure the trainees are able to understand and apply the new rules—would be the middle constraint. Cost is left as the weak constraint, more flexible than the other two. You would naturally plan the project using far more resources than usual, because the additional cost to achieve the deadline is a lot less than the cost of the fines associated with lateness.

Time/Performance/Cost is only one of the six possible hierarchies of constraints. Any order is possible. You've probably experienced situations in which cost considerations drive the project, with time pressure in the middle, and performance the weak constraint. Even if performance is the weak constraint, there's still a minimum acceptable level of performance you must meet (just as in the previous example, there's a maximum cost allowed, even though it's more flexible than usual)—but it may not be possible to achieve much more than the minimum.

To determine the right hierarchy of constraints, you have to understand the underlying ''why'' of the project. Notice in the first instance that the ''why'' is driven by the change in regulations, and the hierarchy flows logically and necessarily from that change. In the second instance, where cost is evidently the driver, you may have to do some detective work. Do the cost pressures flow from the financial condition of the organization, or do they reflect a lack of confidence on the part of senior management that the training program will in fact address the problem to be solved? If the financial condition is perilous, you must find creative ways to work within it. If, on the other hand, there's doubt about your capabilities or effectiveness, you need to find ways to understand and address that doubt.

The severity of any given risk is affected by its position in the hierarchy. For example, on a time-driven/cost-weak project, risks that impact schedule are automatically more serious than those that impact cost. In fact, you may choose to increase cost risk if by doing so you lower schedule risk.

User Needs and Requirements. In other chapters of this volume, you've learned methods of understanding and analyzing user needs and requirements. From a project management perspective, this process begins before the project is established and continues throughout the initial parts of the project. The preliminary study of user needs and requirements determines first whether there is a project and whether you and your organization should do the project, and once those questions are answered, determines the requirements your project must meet to be successful.

This is another example of progressive elaboration in project management. At the beginning, the understanding of user needs and issues tends to be general, and through continuing investigation the understanding becomes deeper and deeper. Depending on the nature of your organization and your relationship with your client or customer, user needs and requirements may be thoroughly developed before you accept the project officially, or it may be more appropriate to accept a general idea of user needs and requirements at the time the project is authorized and develop a more detailed understanding in later phases.

Project Charter. The project charter is the piece of paper that marks the official beginning of the project. Before the project charter, the project is potential; afterward, the project is real. The least important characteristic of the project charter is that it contains the words ''project charter.'' In some training organizations, the project charter takes the form of a signed contract;

Exhibit 17.4 Project Charter

A number of important project decisions must be made at the very beginning of the project. The project charter is a tool to ensure that these decisions are made and that they are appropriately documented. All the items below should be part of the project charter, or at least documented as a supplement to the project charter.

- A statement of the project commitment;
- Identification of the project manager, along with a statement of the project manager's authority and responsibilities;
- A preliminary statement of the scope of the project and its major objectives;
- The business need the project was undertaken to address;
- Assignment of other major project roles and responsibilities;
- Those responsible for review, approval, and oversight of the project, and the specific roles they are to play, including sponsors, customers, and other key stakeholders; and
- A process for change control, budget, and project status reporting.

in others, a memo may be sufficient. A handshake or oral commitment is, however, insufficient, and opens you, your project, and your organization up to major potential risks. If your organization does not provide you with some documentation to certify that a project commitment actually exists, you should at a minimum confirm any oral understanding or handshake in writing yourself.

There are a few essential details that must be established at the beginning of the project. If the document that represents your project charter doesn't contain all the elements in Exhibit 17.4, you need to get these details in writing as well.

TOOLS FOR PROJECT PLANNING

The robust set of project management tools comes into play during all three periods of the project, although in different ways and with different levels of detail. Therefore, we'll cover the tools here, then return to a discussion of the implementation period.

Where and when to begin using these tools is often inexact, depending on the circumstances of your own individual project. Sometimes the analysis of user needs and the preparation of functional and technical requirements is performed completely before the project is accepted and a project charter (or contract) issued. Other times, only a preliminary problem identification is made before the work becomes a project. The project manager and project team, along with the customer and user, must then perform the detailed analysis of user needs and requirements as part of the work of the project. In the first case, formal planning can start as soon as the project is established. In the second, planning becomes more of a process of discovery, and neither customer nor project team quite knows at the beginning what it will take to accomplish the project.

Don't confuse planning with scheduling. A schedule is, of course, an important element of any plan, but the plan doesn't end there. A project plan must take into account resources, risks, quality, communication, procurement, and many other areas of the project. It must provide a map and a blueprint of what the project is, how it will be done, who will do it (and with what resources), how problems will be addressed, how progress will be reported, and how the project team will interact with the customer, the performing organization, and other stakeholders.

It's worth noting that the majority of these tools were originally developed for a linear ''waterfall'' development approach. In an agile environment, the tools don't work quite as crisply, but they're still of great value.

Statement of Work

As noted, sometimes all user needs and requirements are developed prior to accepting the project. Other times, only a preliminary project scope has been established, and the project manager/project team must develop the full scope of work.

The ''Statement of Work'' is a narrative summary of the project, with the detailed project requirements forming a supplement. If it hasn't already been developed, now is the time. It's essential that all key stakeholders agree to the Statement of Work before the project moves forward. Failure to gain agreement opens the project up to significant risks from miscommunication and misunderstandings.

A Statement of Work should be short and clearly written. When there are numerous details, develop those in a requirements format, using numbered statements organized by category, similar in structure to training terminal objectives.

Work Breakdown Structure (WBS)

Although the Statement of Work is essential to ensure agreement among key stakeholders about the scope and objectives of the project, more detail is needed to enable the project team to manage the project. The Work Breakdown Structure, commonly known simply as a WBS, uses a graphic approach to organize project scope and break the scope into manageable work packages. For each work package, the project team will develop estimates for duration and resource usage, identify risks, assign responsibilities, and develop the other important project management tools.

Think of the WBS as the foundation and framing of a house. Even though the foundation and framing of the house are not very visible in the final construction, it's clear that the quality of the house will surely be no better than the quality of the foundation and framing. The quality of your overall plan cannot be better than the quality of the WBS that underlies the other planning elements.

The WBS organizes and defines the scope of the project and serves as the underpinning of the other planning tools. The better the quality of the WBS, the better the quality of subsequent steps in the planning process. This WBS is displayed in ''org chart'' format; you will also see the WBS presented in outline form. Use sticky notes on a whiteboard or flip chart to create the WBS so you and your team can work collaboratively and explore different options.

In Figure 17.6, the training project is organized by phase. From a management perspective, this implies that the work takes place largely within a single department, with a project manager from that department overseeing the project from start to finish. This, of course, is not the only option.

The WBS for your specific project should reflect the management and organizational structure that will actually be used for that project. For comparison purposes, assume that the company has an M&OD department and a separate training department, and will be using an external ISD vendor for the development of materials and train-the-trainer resources. This gives you the WBS approach shown in Figure 17.7.

The same project can be organized different ways with equal validity. To be useful, the WBS must be organized in the way you actually plan to manage the project. In Figure 17.6, the project is organized by phase, suggesting that a single team will manage the project from start to finish. In this version, two departments and an outside contractor have responsibility for different aspects of the project, which will require a very different approach to managing the project.

The role of the project manager is going to be very different between the two approaches shown in Figures 17.6 and 17.7. In the first case, the project manager is a hands-on team leader. In the second, the project manager will most likely be someone of a higher managerial level, shepherding the project from department/vendor to department/vendor but not providing hands-on direction. The project manager will need to oversee the transitions when primary operational responsibility moves from work group to work group, but hands-on direction of the details will necessarily be provided within each work group.

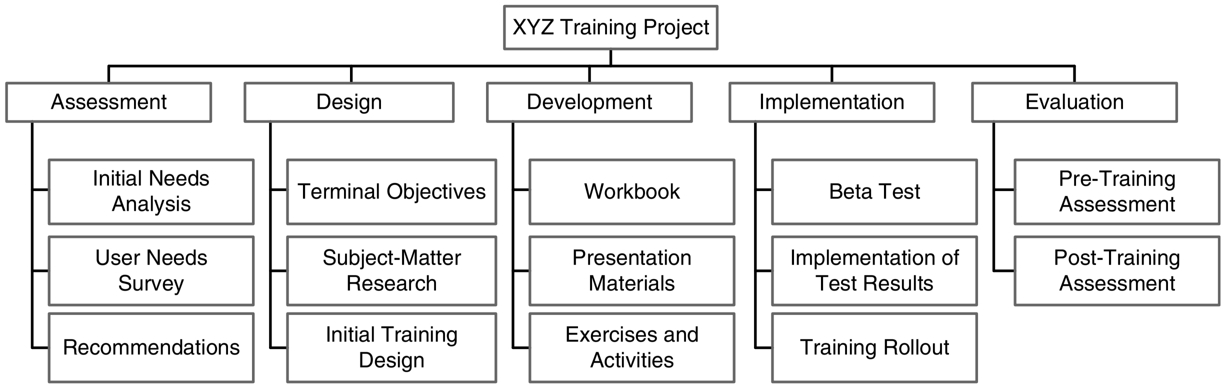


Figure 17.6 WBS Organized by Phase.

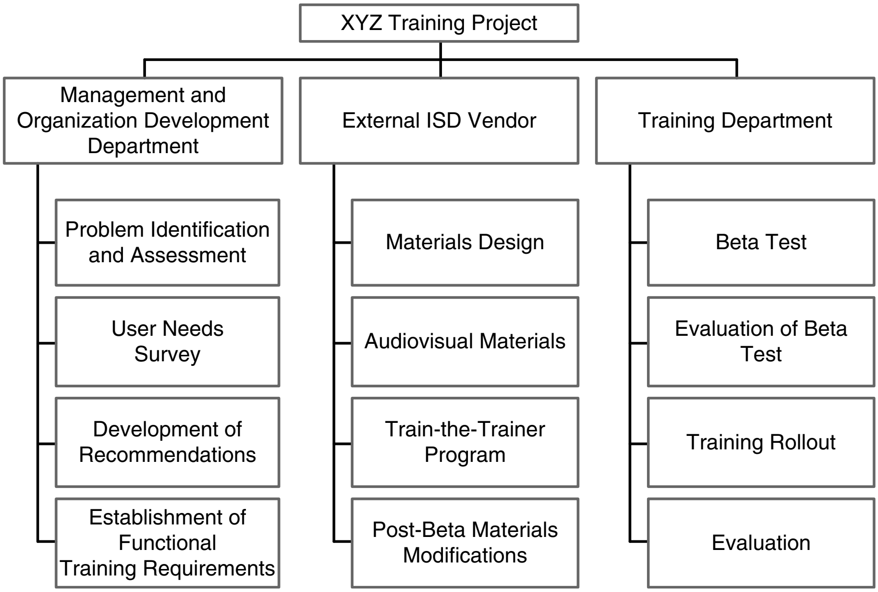


Figure 17.7 WBS Organized by Functional Work Group.

Each approach has advantages and risks. The approach you will follow may be determined by existing organizational structure and political realities of the organization. To be useful and relevant, a WBS must always reflect the actual management approach that will be followed.

Develop the WBS using sticky notes on a whiteboard or flip chart, rather than by using project management software. Although project management software has numerous advantages that increase along with the size of your project, the software is generally not conducive to group brainstorming. It's important to stay flexible and explore different options for laying out the project.

Sticky notes with subordinate activities are known as ''control accounts,'' and they represent the management and organizational structure of the project. Sticky notes without subordinate activities are ''work packages.'' It is at the work package level that resources are spent and project work is accomplished. Control accounts allow you to ''roll up'' information to facilitate reporting and oversight of the project.

When you are certain the WBS is complete and reflects the management structure you will follow, then it's time to enter the data into your project management program, if you choose to use one. Most project management software will display the WBS in the ''outline'' format rather than the ''org chart'' format used in Figures 17.6 and 17.7, as shown in Exhibit 17.5.

Exhibit 17.5 WBS in Outline Format

In project management software, the WBS normally appears as an outline, rather than the ''org chart'' approach shown in Figures 17.6 and 17.7. WBS numbering normally follows an outline structure, but it's most convenient for you to use whatever numbering system provided by your project management software.

WBS for XYZ Training Project (Organized by Phase)

1.0 XYZ Training Project

1.1 Assessment

1.1.1 Initial Needs Analysis

1.1.2 User Needs Survey

1.1.3 Recommendations

1.2 Design

1.2.1 Terminal Objectives

1.2.2 Subject Matter Research

1.2.3 Initial Training Design

1.3 Development

1.3.1 Workbook

1.3.2 Presentation Materials

1.3.3 Exercises and Activities

1.4 Implementation

1.4.1 Beta Test

1.4.2 Implementation of Test Results

1.4.3 Training Rollout

1.5 Evaluation

1.5.1 Pre-Training Assessment

1.5.2 Post-Training Assessment

WBS for XYZ Training Project (Organized by Functional Work Group)

2.0 XYZ Training Project

2.1 Management and Organization Development Department

2.1.1 Problem Identification and Assessment

2.1.2 User Needs Survey

(*Continued*)

Exhibit 17.5 (*Continued*)

2.1.3 Development of Recommendations
2.1.4 Establishment of Functional Training Requirements
2.2 External ISD Vendor
2.2.1 Materials Design
2.2.2 Audio-Visual Material
2.2.3 Train-the-Trainer Program
2.2.4 Post-Beta Materials Modifications
2.3 Training Department
2.3.1 Beta Test
2.3.2 Evaluation of Beta Test
2.3.3 Training Rollout
2.3.4 Evaluation

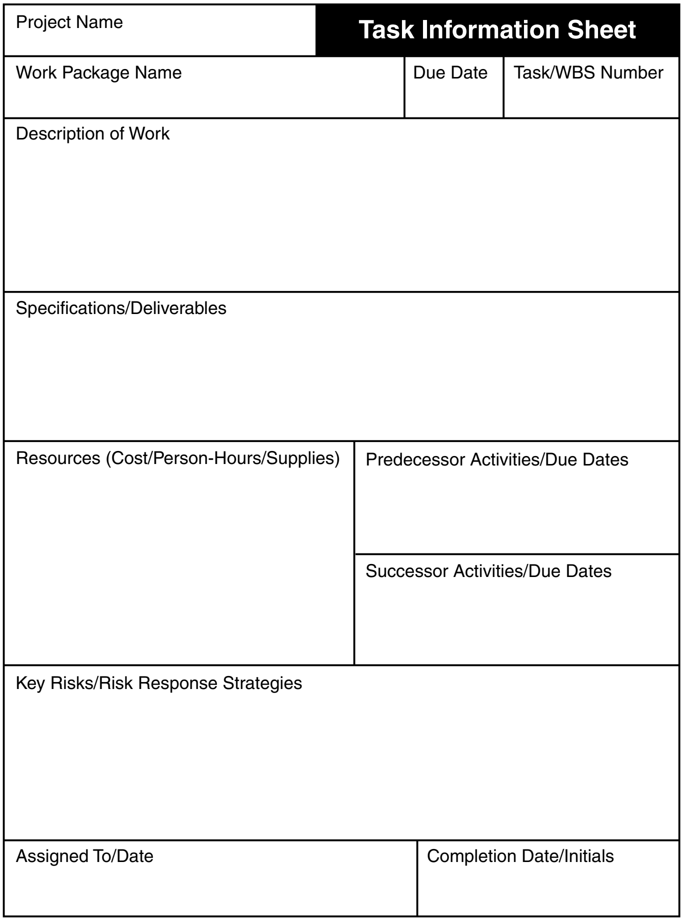
Task Analysis Worksheet

Each work package in your project needs to be developed. What is the work that must be accomplished? When is it due? What must be completed to allow this work package to proceed? What can't start until this work package is completed? These are only a few of the important questions you need to answer, and the answers must be written down so neither you nor the members of your project team overlook important work. The Task Information Sheet shown in Figure 17.8 is a tool to help you organize this critical information.

The task information sheet captures all the vital information needed to ensure each work package in the project is performed correctly.

You will normally have to develop the information for this sheet piecemeal. At the beginning of the process, you may only be able to identify the names of the individual work packages. Due dates, WBS numbers, detailed work description, and other information is added as developed. To save time and improve performance, consider having the team members who will perform each task do most of the work of developing the form (subject to project management approval, of course) and recycle task information sheets from previous projects.

Effective delegation is one of the keys to successful project management. When the task information sheet is fully developed, it makes an excellent tool for delegation. The technique of *kanban* task management (Dobson, 2003, pp. 311–312), shown in Exhibit 17.6, helps you control activities on multiple projects.


Project Name | **Task Information Sheet**

Work Package Name | Due Date | Task/WBS Number

Description of Work

Specifications/Deliverables

Resources (Cost/Person-Hours/Supplies) | Predecessor Activities/Due Dates | Successor Activities/Due Dates

Key Risks/Risk Response Strategies

Assigned To/Date | Completion Date/Initials

Figure 17.8 Task Information Sheet.

Exhibit 17.6 Kanban Task Management

Use the Task Information Sheet to control project work, especially in a multiple project environment in which shared resources are common, by following the steps listed below. (This process envisions paper, but you can easily implement this on a spreadsheet or using other software of your choice.)

1. Create a binder for the project that contains *two* copies of each Task Information Sheet.
2. To assign each work package, give one copy of the sheet to the person assigned. Write that person's name in the ''Assigned to'' space along with the date the work was assigned.

(*Continued*)

Exhibit 17.6 (*Continued*)

3. Write status information and updates on your copy of the form, adding additional sheets as necessary.
4. If the scope of the work changes, revise the Task Information Sheet as necessary. Give a copy of the revised sheet to the assigned person and collect the previous sheet. Do *not* allow multiple versions of the same assignment to circulate.
5. The assigned person completes the work package, verifies with checkmarks that specifications and deliverables are achieved, places his or her initials and date on the form, and returns it to the project manager. The work package is not considered complete until this step is done.
6. Keep both copies of the sheet, notes, changes, and other information in the master project binder to be reviewed during project evaluation and ''lessons learned.''

Schedule Development

Two scheduling tools are common in project management: the Gantt chart, which is essentially a bar graph over a calendar; and the network diagram, which shows the sequence in which activities will be performed. For small and medium-sized projects, the Gantt chart is the most common and easiest scheduling tool; for very large projects, the network diagram is more appropriate.

Even when the Gantt chart will be used as the scheduling tool, it's often better to lay out project activities and work packages as a network diagram first, then convert the information to the Gantt chart format. Here's how to develop both tools, using the WBS ''organized by phase'' version from Figure 17.6.

Network Diagram. A network diagram resembles a computer flow chart. To build a network diagram, use the ''work package'' sticky notes from your WBS. Do not use any ''control account'' sticky notes in making the network diagram.

The first step is to create a ''Start'' milestone for your project. A milestone is a work package that has zero duration and no associated work or resource consumption. In other words, a milestone is simply a signpost. Traditionally, a milestone is represented by the shape of a diamond. Turn a sticky note 45 degrees to indicate that a given work package is a milestone.

Next, lay out the subsequent work packages in the order they are to be performed. Activities can be *dependent* (following a predecessor activity) or *parallel* (performed at the same time as other project activities. Dependent activities are sometimes required by the logic of the work (that is, you can't conduct a beta test of the training unless the training materials have been developed), and are sometimes driven by resources or other factors (that is, the same person can develop the workbook and the exercises, but not at the same time). If no particular order is demanded by logic, you can choose whichever order you prefer.

When all activities have been placed and connecting lines drawn, create a ''Finish'' milestone. Connect all unlinked activities to Finish so that every work package has at least one predecessor and at least one dependent activity—don't leave any orphaned work packages.

Normally, more than one sequence of activities is possible. As with the WBS, the correct order for your project is the one that represents how you and your team plan to approach this project. Figure 17.9 shows a sample network diagram.

The network diagram shown in Figure 17.9 reflects the order in which work packages will be performed. Note that estimated durations have been assigned to each activity.

Here's how to read the network diagram. Task A is a milestone and serves as the start of the project. Both tasks B and C are dependent on the start milestone. Task D is dependent on both tasks B and C; task E is dependent only on task C. Task F is dependent on task D; task G is dependent on both tasks D and E. Task H, the finish milestone, is dependent on both tasks F and G.

So how long will the project take? With dependency relationships crossing from top to bottom and back again, the answer takes a little bit of calculation. You need to find the longest path through the project network to determine the length of the project. The longest path is commonly known as the *critical path*. If you use project management software, it will determine the critical path for you automatically. To calculate the critical path manually, you must perform a *forward pass* followed by a *backward pass*.

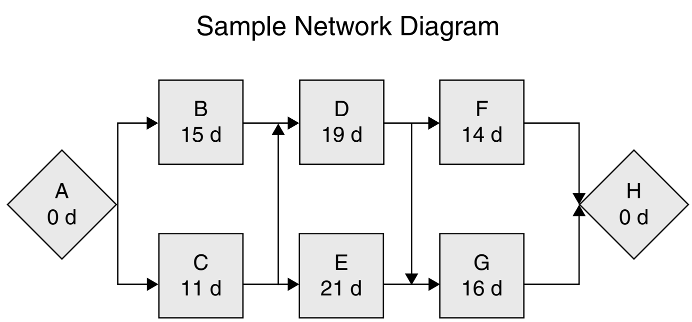


Figure 17.9 Network Diagram.

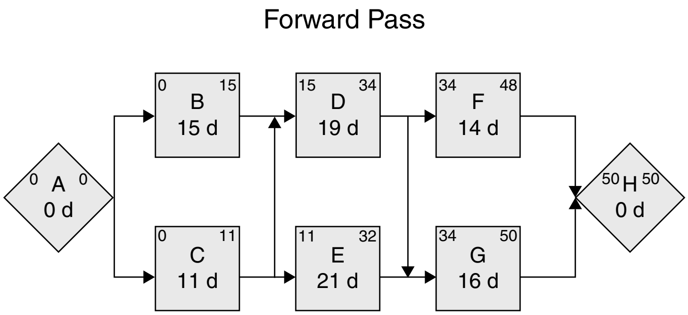


Figure 17.10 Forward Pass.

Forward pass. In the forward pass, you calculate the *early start* and *early finish* of each activity, starting with the first task. Figure 17.10 shows the forward pass. The forward pass calculates the early start and early finish for each activity. The final number in the forward pass calculation is the planned duration of the project.

Starting in the upper-left corner of the first work package, you enter a zero, which represents the start of business of the first day of the project. Add the task duration to the start, and write that number in the upper-right corner. In task A, the early start (0) is added to the duration (0) to get the early finish ($0 + 0 = 0$). Copy the early finish number into the connected tasks B and C, and add the respective durations ($0 + 15 = 15$ and $0 + 11 = 11$).

Task E, you'll note, is dependent only on task C, and can start on day 11. Task D, however, is dependent on both tasks B and C. Because task D cannot begin until both predecessor tasks are complete, it takes the *larger* of the two early finish dates, or day 15. Similarly, task G takes the larger of the early finish dates of its two predecessors, tasks D and E.

At the end of the forward pass we know the duration of the project. However, we aren't done yet.

Backward pass. If we want the project to finish within its allotted fifty days, we will now calculate the *late finish* and *late start* of each activity, shown in Figure 17.11. The late finish is the latest an activity can be completed while still achieving the overall deadline. The late start is the late finish minus the task duration.

The backward pass calculates the late finish and late start of each activity, showing the latest any activity can be performed while achieving the original deadline.

As seen in Figure 17.11, tasks F and G can both finish as late as day 50. Task F can therefore begin as late as day 36, while task G cannot begin any later than day 34. Task D must finish in time to allow both tasks F and G to finish no later than day 50. This means that the *lower* of the two late start numbers (34 in this

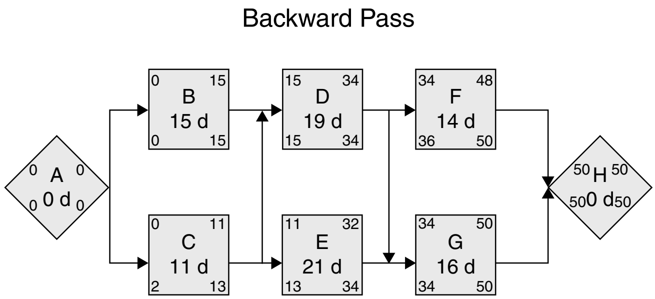


Figure 17.11 Backward Pass.

case) is the late finish of task D. The backward pass must end in zero when you reach the beginning of the project.

Critical path and float. The critical path is the longest path through the network, shown in Figure 17.12. On the critical path, there is no difference between the early start (or finish) and the late start (or finish) of each activity. In other words, any delay of a critical path activity results in immediate danger of a late project.

A task is *critical* if there is no difference between the early start (or finish) and the late start (or finish). Any delay in a critical path activity results in a delayed finish of the project. If the late start (or finish) is greater than the early start (or finish), the difference is called *total float*. A task can have delay equal to its float without affecting the project's deadline. *Free float* is the amount of delay before the task forces a delay in any subsequent activity; float that is not free is *shared* with other activities.

Notice that in the case of task C, it can start as early as day 0 but can finish as late as day 13. It has two days of *total float*, extra time before lateness jeopardizes the project deadline. The same is true of task E. However, this

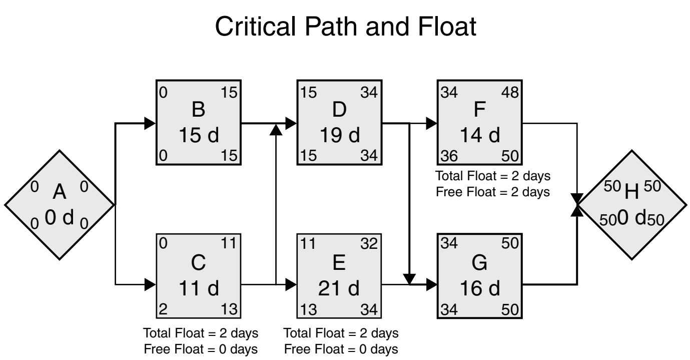


Figure 17.12 Critical Path and Float.

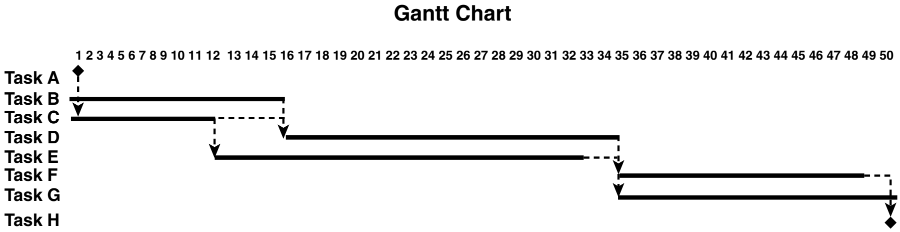


Figure 17.13 Gantt Chart.

does not mean that both tasks C and E can be two days late. The float in task C is *shared*, because any delay in task C, even if it doesn't affect the deadline, reduces the float available for task E. The float in tasks E and F, however, is *free float*, because no other task is affected if those activities use their available float.

Gantt Chart. The network diagram allows you to design the sequence of activities, but it's not a very intuitive way to visualize how long the project will take. For that, the Gantt chart is much clearer. A Gantt chart is essentially a bar graph of the schedule in calendar time. Almost all project management software will display a Gantt chart easily, but you can use graph paper or a spreadsheet program if you don't have a need for specialized project management software. Figure 17.13 shows the network diagram converted to a Gantt chart (this one done in Microsoft Excel).

Gantt charts help you assign resources, track progress and, in a multiple project environment, help you oversee the progress of all projects simultaneously. You can identify how resources are used across multiple project boundaries, and even compare different versions of the same project side by side.

The Gantt chart shown in Figure 17.13 displays the schedule as a bar graph over time. It is the easiest schedule to develop and can be read by people without formal project management training.

Risk Management Planning

Because projects are temporary and unique, risks are inherently part of project management. A risk differs from a problem in tense: problems are present tense; risks are future tense. Every risk in your project will eventually go away—some by not happening at all, and others by turning into problems.

Risks include threats (negative risk events) and opportunities (positive risk events). While threat avoidance and threat management normally occupy the majority of a project manager's focus, it's worth it to spend some time on opportunities as well. Can work done for one project benefit other projects? What if you could find pre-existing training materials that could be adapted, rather than developing the entire program in-house? What additional benefits might result from a given training effort other than those specified originally?

A risk is defined as the probability of a given event happening multiplied by its effect if it does happen ($R = P \times I$). If there's a 10 percent chance of a problem on your project and a $1,000 cost if it happens, the risk score is $100. In other words, if you can get rid of the risk for less than $100, your project is better off. If the cost of risk mitigation exceeds $100, it may be better to accept the risk. However, the project management situation is more complex.

Classical risk relies on the law of large numbers: insurance underwriters can't tell whether you personally will be involved in an automobile accident, but given a large population of drivers, they can determine the likely number of accidents and their associated costs. Project managers usually don't have the law of large numbers working for them, so can't calculate risk scores with the same precision. However, it's usually possible to conduct a reasonable risk analysis without that information. The first step in risk management planning is risk identification, followed by the steps of risk analysis, risk response planning, and risk monitoring and control.

Risk Identification. Through brainstorming, review of lessons learned on previous projects, and analysis of project documentation (requirements, user needs, terminal objectives, contracts, and so forth), compile a list of potential risks for the project. In the initial risk identification process, err on the side of inclusion. Begin with the risks you identified using Exhibit 17.7.

Risk Analysis. The goal of risk analysis is to understand each risk thoroughly, prioritize the risks, and identify those risks that most require attention on the part of the project manager and project team. A number of different techniques exist to allow categorization and analysis of risks, even when hard statistical data is not available to calculate $R = P \times I$ scores. Exhibit 17.8 shows the use of the filtering technique; Figure 17.14 shows a risk priority assessment tool that can be used in the absence of hard numerical information. You may use both techniques in conjunction to perform risk analysis.

Risk Response Planning. Risks can be sorted and ranked, even in the absence of hard numerical data, using Table 17.3. Classify each risk by your estimate of its probability and by your estimate of its impact on project objectives, then

Exhibit 17.7 Risk Identification

List potential risks (both threats and opportunities) that might affect your project. Write each risk in the format ‘‘condition + consequence,’’ identifying the issue of concern and then specifying its potential impact on the project if it occurs. Be sure to identify potential opportunity risks as well as threat risks.

Examples

1. *Threat*. Our ISD specialist is currently working on ABC Training Project. If that project takes longer than expected (condition), she may be delayed in starting on the XYZ project (consequence).
2. *Threat*. The XYZ training requires us to administer the MBTI instrument. Only some of our trainers are certified in MBTI (condition), which may result in the need to shuffle schedules to ensure classes are covered (consequence).
3. *Threat*. The customer for the XYZ training has expressed a desire for staff to spend no more than one day in the classroom (condition). This may require compromises in the quality of instruction or amount of hands-on practice (consequence).
4. *Opportunity*. The JKL training program we designed two years ago addresses issues similar to those identified in the XYZ training. If the detailed analysis of the JKL training shows enough overlap, we may be able to adapt the JKL materials (condition), resulting in a significant reduction in development and design work (consequence).

Exhibit 17.8 Filtering Technique

For each potential risk, establish risk priority by asking a series of filtering questions. ‘‘No’’ answers show a reduced risk priority; ‘‘Yes’’ answers an increased risk priority. The following example uses the risks from Exhibit 17.7.

Filter 1: If this risk occurs, will the project consequences be visible to any customer in terms of the Triple Constraints of cost, schedule, or performance?

Risk 1: Yes, potential schedule delay.

Risk 2: No, all impact is within our department.

Risk 3: Yes, potential performance impact.

Risk 4: Yes, potential cost reduction.

Filter 2: Is there any reason to believe that this risk is likely to happen (past experience, circumstances, other information)?

Risk 1: Yes, past experience shows that ISD work often takes longer than expected.

Risk 3: No, customer will allow additional classroom time if necessary to achieve training objectives.

Risk 4: Yes, some benefits from JKL training are likely to be realized.

Filter 3: Must this risk be addressed immediately, or can we wait until more information is received?

Risk 1: No, we need to wait until the ABC training design is completed.

Risk 4: Yes, we must analyze JKL material quickly in order to make adjustments to the XYZ development plan.

---

cross-reference your estimates to establish priority as High Risk, Medium Risk, or Low Risk. Focus risk mitigation efforts on your higher risk activities to ensure best use of your resources.

When your risks have been identified and analyzed, it's time to figure out what you plan to do about them. There are a limited number of available strategies for managing threats and opportunities on your project. Consider all possibilities before making your final decision.

When determining which strategy to follow, you need to look at two additional issues: *residual risk* and *secondary risk*. Residual risk is the risk left over after you've applied your risk response strategy. Secondary risk is new risk created by your proposed solution. Use Table 17.3 to determine your risk response options and consequences. A limited number of potential risk responses for threats and opportunities are shown in the table. Note that the potential for residual and secondary risk is almost always present.

In the agile environment, don't neglect the risk management advantages of evolutionary development (spiral or vee approaches). Rapid prototyping of training programs helps manage risk by bringing problems to the forefront early

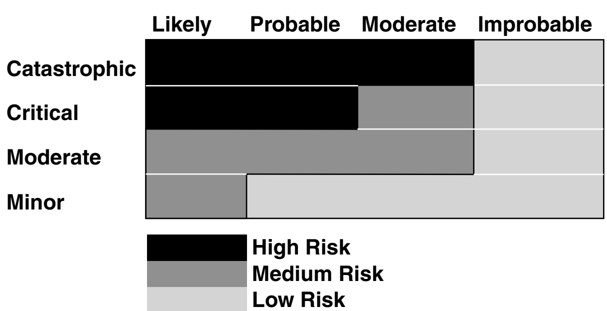


Figure 17.14 Risk Ranking Table.

Table 17.3 Risk Response Options

Strategy	*Description*	*Residual Risk?*	*Secondary Risk?*
Avoidance (Threat)	Change the project so that the risk event cannot occur, or that the project is completely protected from its consequences.	No	Yes
Transfer (Threat)	Give the risk to another party. Financial risk is often transferred using insurance or contracting. Management risk can be transferred to other departments or other managers, as appropriate.	Yes	Yes
Mitigation (Threat)	Reduce the threat by lowering its likelihood of occurrence or by lowering its impact on project objectives.	Yes	Yes
Exploitation (Opportunity)	Use the opportunity to improve the project's timeliness, cost, or quality.	Yes	Yes
Enhancement (Opportunity)	Try to improve the opportunity's probability or impact.	Yes	Yes
Sharing (Opportunity)	Transfer the opportunity elsewhere, such as to another project.	Yes	Yes
Active Acceptance (Either)	Develop a contingency plan to be triggered if the risk occurs or if it appears immediately likely to occur.	Yes	Yes
Passive Acceptance (Either)	Do nothing unless the risk occurs, then figure out a strategy based on the facts at hand.	Yes	No

in the process, when remediation is relatively easy. It's not just how *bad* a problem is, it's also *when* it occurs.

Quality

All the various quality management programs say that quality must be built into a program or project up-front. Inspection and quality assurance activities keep defects out of the hands of the customer, which is desirable. When quality is built in from the beginning, however, defects are removed altogether, which is far more desirable.

The first step in designing quality into a project is defining what constitutes quality in the first place. Quality may be defined as ''exceeding customer expectations,'' in which case you must start first by defining what those expectations are, and then defining what constitutes meaningful improvement. Some quality programs focus on the ''-ilities,'' identifying factors such as manufacturability, repeatability, safety, security, etc., that mostly (but not always) end in ''-ility.'' Some of these are more important than others for a given project. ''-Ilities'' for a training project might include teachability, clarity, or (not ending in ''-ility'') depth. Going for a deeper level of understanding might decrease clarity or ease of instruction. In this particular case, is the tradeoff worthwhile?

One technique common to most quality programs cannot be used within the confines of a project: continuous improvement. While you can continuously improve the way in which you do projects, you must ultimately finish the current project (or project phase) or else you have nothing of value.

Human Resources/Procurement

A key question for project planners is to determine the human resource requirements for your project and how you will obtain and use them. Analysis of the WBS and schedule will determine the people required to do the job; analysis of the subject matter will determine the level and type of skills required.

In training design and development, an associated question is to what extent the work will be performed in-house using available resources or contracted out to training vendors. These essential questions must be asked and answered early in the planning process, because they have a substantial impact on all subsequent project activities.

Communications Planning

The final part of developing a comprehensive project plan is to determine the communications needs. How often will project status be reported, in what form, and to whom? If there will be a project team, should you use a team charter? How will the customer and user be kept informed of project progress? How will risks and other issues be escalated? To improve quality and lower time, project management organizations often establish a standard communications protocol to which all new projects automatically conform.

IMPLEMENTATION PERIOD

The implementation period includes the steps of source selection (if you're using a vendor), development, and verification. It is performing the work of the project, and as a result it is necessarily tailored to the unique environment of that project. The other chapters of this book address most of those issues. In this chapter, we are concerned with the project management aspects.

From the project manager's perspective, there are a few elements that involve classical project management activities. First is acquiring, development, and leading the project team. Second is managing any associated contracts. Third is distributing information to stakeholders according to the communications plan.

Such activities as status meetings, progress reports, and review of work performed to date give the project manager information about the progress of the project. You need to establish in advance how and when you will be gathering this information.

Verification includes all the oversight issues necessary to oversee the project. These activities divide into three groups: quality assurance, which involves applying the quality strategies developed in the planning process; monitoring project progress in the areas of time, cost, and performance; and dealing with changes to the project, both voluntary and involuntary.

Discrepancies between the plan and your project reality result in involuntary change: the project schedule is slipping, costs are increasing, and performance is not meeting expectations. (Or, if you're lucky, the project is early, requires fewer resources than you expected, and the results are much better than anticipated.) Minor discrepancies may not require action on your part, but any large discrepancy requires investigation, at a minimum.

The final part of risk management takes place during this phase: risk monitoring and control. You may find that risks you thought likely and serious turn out to be unlikely and minor, while other risks climb the hit parade with alarming speed. Review and update your risk management plan periodically, and monitor how well your proposed risk strategies are working. Make changes as necessary.

Voluntary change occurs when stakeholders request or require you to make changes to the previously approved project scope. ''Scope creep'' is a traditional enemy of project success. Scope changes by themselves aren't the problem: unmanaged and undocumented scope changes are what you need to worry about.

A good change management system always requires that changes be made in writing, reviewed for project impact, then approved by the appropriate party (which may or may not be the project manager or even someone inside your own organization). The impact to the project of the proposed change is the critical issue. Does the change in scope affect the time, the resources required, or the quality of other project elements? Do you simply want to accept the impact (yes, it will take three more weeks, but that's okay) or, in approving the change, do you need to implement corrective action (three weeks' delay is unacceptable, so the price of the contract must be increased to allow addition of one more staff member or not all training modules will be rolled out in the initial phase of the training program)?

OPERATIONS PERIOD

The study and implementation periods build the first iterations of our training program. At some point, whether there's one beta test or whether there are several, you must roll out the training program to your constituency. You're not done with the project, although you're done with most of the development. You still have to ensure that the user needs—the reason for undertaking the project in the first place—are fulfilled by the solution you have developed.

Often, moving from implementation to operation involves a handoff of project leadership and primary responsibility from one organizational element to another, whether it's from instructional design to training operations, or from a consultant/vendor to in-house resources. It's often useful to think of certain points in the project as *decision gates*, because not only do you move primary responsibility for the project, but you often also make decisions about whether a project should be continued, cancelled, or changed.

DEACTIVATION PHASE

Projects are temporary; therefore, they must end. Ending the project is known as project deactivation. In project deactivation, the product of the project must be moved to the next phase (implemented, given to the customer, published) or terminated; any contracts must be completed; the project team and other resources must be assigned to new work or returned to the labor pool; final reports must be provided; and the project files must be completed and archived.

The two remaining steps in project closeout have more to do with future projects than with the current project, but they are no less important.

First, you need to celebrate success. Even though the next project is calling, it's important to acknowledge the hard work and efforts that made the current project successful. Say thank you in a meaningful way: write a letter, prepare a certificate, throw a party, or take people to lunch. This is not optional. You'll need many of these people to make your future projects a success. If you want a high-performing team for this and future projects, you must always recognize success to build loyalty.

The final step is developing lessons learned, using some of the questions in Exhibit 17.9. Although, as we've learned, you can't do continuous improvement within a project, you can definitely improve the way in which you perform

Exhibit 17.9 Lessons Learned

Consider the following questions in determining how today's experience can make future projects better.

1. What did we do right?
2. What could we do differently or better next time?
3. What did we do on this project that could be recycled for future projects?
4. What did we *not* do that we should ensure we *not* do in the future?
5. What value did we get from the project management tools that we used?
6. What training or skills will we need in the future that were lacking in this project?
7. What surprises did we encounter? Should we have been surprised? Should our risk plan consider these issues for the next project?

future projects. To do that, you must extract the learning value from today's experience. Avoid allowing this to become a ''blamestorming'' session. It's no longer relevant who did what wrong. What's relevant is how future projects can be better.

CONCLUSION

Project management was originally developed for large industrial and engineering projects using mostly linear ''waterfall'' approaches. To make it relevant and useful for smaller projects and for projects using agile approaches, we must be careful to scale our project management to the size of our projects and the risks and complexities we are likely to encounter, and adjust our project management methodology based on the type of development and approach we choose.

If you haven't used formal project management techniques on previous projects, start small and add complexity and sophistication as appropriate. Don't be afraid to decide—after a fair trial—that a given technique may not deliver the project results you desire.

From operational project control to advanced agile methods, the wise project manager recognizes the value of good process and good techniques. Project managers operate within the bounds of the Triple Constraint, within the context of their organizations' and stakeholders' environment, in a world where power is limited but responsibility often isn't, and have accrued practical wisdom over centuries of projects—both successful and not so much.

The wise project manager understands the benefits of process, and knows how to customize it for the project environment at hand. Remember, failure may sometimes be an accident—but success never is.

References

Dobson, M. (1999). *The juggler's guide to managing multiple projects* (pp. 69–83). Newtown Square, PA: Project Management Institute.

Dobson, M. (2003). *Streetwise project management* (pp. 311–312). Avon, MA: Adams Media Corporation.

Dobson, M. (2004). *The triple constraints in project management (project management essential library)* (pp. 7–12). Vienna, VA: Management Concepts.

Forsberg, K., Mooz, H., & Cotterman, H. (2005). *Visualizing project management: Models and frameworks for mastering complex systems* (3rd ed.). Hoboken, NJ: John Wiley & Sons.

CHAPTER EIGHTEEN

Managing Relationships in the Performance Improvement Process

Jane Ranshaw

OVERVIEW: HELPING RELATIONSHIPS

Performance improvement and consulting are helping relationships in which the practitioners must accomplish their goals through influence rather than direct action as a line manager would do (Block, 2000; Schein, 1999). Because they rarely have the authority to implement changes, they must rely on their relationships with line managers. More projects fail because of poorly managed relationships, including poor communication, between the various players than any other single reason. (A simple search of the term ''project failure'' on the Internet will produce more than eight million responses; a more targeted search results in hundreds of articles.) In a July 28, 2008, article for itsmwatch.com, columnist Hank Marquis of Enterprise Management Associates writes, ''A recent survey of over 1,000 IT workers engaged in failed projects found that the reason most IT projects fail was due to poor communications. About 28 percent of respondents said poor communication is the number one cause of project failure.''

While most research deals with information technology and related fields, the principles apply more widely. In fact, the need for managing relationships occurs well before any project. A good business relationship requires trust and knowledge of each other, and these take time to build.

The emphasis in this chapter is on managing relationships within one's own and client organizations. First, we must identify the relationships we want to manage, beginning with ourselves and including the many types of relationships we find in organizations. Next is contracting between the consultant and the client. Everyone can remember a situation that foundered because of unclear expectations. And, regardless of the clarity and specificity of the contract, every performance improvement consultant can expect to encounter resistance on most projects. The later part of this chapter examines the kinds of resistance one should anticipate and how to handle them.

IDENTIFYING RELATIONSHIPS

The first step in managing relationships is to identify the ones you want to manage. Sometimes the reason is obvious: you will have some type of relationship with your boss and immediate co-workers. Others occur because potential clients contact you or your organization about a project. Still others occur because performance improvement practitioners have strategically identified the relationships that will help them add the most value to their organizations.

What Your Intent Is

The first relationship to manage is the one you have with yourself. While you don't have to undergo psychoanalysis, you should assess your own strengths and weaknesses. (See Exhibit 18.7, My Personal SWOT Analysis, at the end of this chapter.)

Keep in mind that sometimes skills involve much more than specialized knowledge. When the author was beginning her career as a technical writer in the early 1980s, one of her first clients was a consulting firm whose client wanted a reference and custom learning program for a new word processor. She was intimidated by the other two writers because they both had considerable experience in the field. The author had a vague idea of what computers were, but she decided to view the project as an opportunity to learn about a new topic. The project encountered challenges typical of any emerging area, and rewrite after rewrite ensued for all the writers. At the end of the project, the client asked her to undertake some additional work, which proved to be relatively easy and fairly lucrative. When all the work was completed, the author asked the client why they had selected her for this final phase. After all, the other two knew much more about writing for computer users. ''You were the only one we were still speaking to,'' said the client. The author's decision to accept the client's sometimes acerbic comments and demands for seemingly endless rewrites as a learning experience deepened the relationship between herself and the client. The experience was also a lesson in contracting, which will be discussed later in this chapter.

Before establishing relationships with others, you must have a reason. Ask yourself, ''Why do I want this relationship?'' ''Is this an opportunity to learn new skills?'' ''Is this a person I can help?'' The best relationships begin with positive intent and involve reciprocity. Positive intent refers to motivation. To be truly successful at managing relationships, you must begin with a genuine concern to be of help and to add value to your organization. Avoid looking as each new person you meet as an instant potential client or road to a client (and ignoring anyone who doesn't meet that criterion). Also, be cautious of seeing every relationship as an opportunity to practice your hard-won performance improvement skills; you run the risk of being a solution looking for a problem. Look at each new contact as someone who has an interesting story to tell before jumping to conclusions. Be prepared to listen well and help the other person identify true needs. Sometimes the best advice is that not given. Being non-judging can help build trust, which is a solid foundation for any relationship.

At the same time, you must consider what's in it for you. Make conscious decisions about what you want from the different kinds of relationships you have. Be prepared to let go of relationships that do not let you meet your personal or organizational goals. You do not serve your organization well if you become a servant of line organizations. Do keep in mind, however, that clients often do know what they want, and your role is to focus their efforts.

The Purpose of the Relationship

Whether the performance improvement practitioner is an internal or external consultant, the obvious purpose for most business relationships is straightforward: work. Those who put too much emphasis on an immediate, direct payoff, however, are overlooking fields rich with opportunities. Other purposes include:

- *Identify needs not currently identified by the persons directly involved*—Successful performance improvement groups often find that a line operation is reluctant to approach them (or the line managers may be oblivious to the possibilities the group presents). Rather than trying to convert the managers in a heretofore resistant group, they choose a less direct approach by partnering with other key support groups such as purchasing or accounting.

  One group of internal consultants visited with supervisors on the third shift (at a little after midnight) to discuss challenges around training schedules. From this initial contact, they learned about organizational issues unique to shift workers, real-life operations in the company, and new areas where they could support mid-level plant managers. Over the next two years, they were able to forge relationships with supervisors on all shifts and found new ways to add value.
- *Learn more about the culture of the organization*—Simply spending time with people from different parts of the organization helps performance improvement practitioners soak up knowledge about differences within

an organization's culture. Accountants really are different from salespeople, and corporate operations rarely parallel their field counterparts. Although the differences are often subtle, they are nonetheless real, and the only way to absorb them is from contact beyond occasional, formal meetings.

Ask people from different departments to explain how their operations work. You could ask them to make a presentation to your group or take a peer from another group to lunch. Keep your antennae out for any issues with which you might help, of course, but don't expect an immediate payoff. In fact, your colleague knows more about his or her organization and may be more likely than you are to see opportunities for you.

- *Find resources that can help other parts of the organization*—Real win-win solutions occur whenever a performance improvement professional is able to introduce other resources to a group needing assistance. One practitioner at a Chicago bank discovered that the compliance department was spending hours each month manually transferring numbers from one report to its own spreadsheet. They had repeatedly asked ''someone in IT'' for help and had been told the effort was too great for the value received. In less than an hour, the practitioner learned they were asking the wrong group for assistance. She referred the compliance people to a different group in the IT department, and the problem was quickly solved. While she had never worked with either group on a project, by involving pure performance improvement principles, she helped her colleagues solve a problem and built friends.
- *Find mentors/be a mentor*—Everyone can benefit from the advice of more seasoned co-workers, especially those whose specializations are in areas other then performance improvement. Be a bit cautious, however, of people who seem a bit too eager to establish a relationship. Mentoring relationships work best if each person keeps a respectful distance, at least in the beginning. Friendships often do develop between co-workers, of course, but they should evolve naturally. Also, be aware that, as in the rest of life, there can be sudden shifts in once-stable relationships. Remember that time in junior high school when your dearest and best friend suddenly switched sides and became buddies with the person you both hated the week before? In some ways, corporate relationships haven't evolved beyond that standard.

With Whom to Have Relationships

The temptation is to say you need relationships with everyone. There are, however, not enough hours in the day to follow through on that type of commitment. You need to learn those areas in which you might provide PI assistance. Look for people who can either hire you or refer you, of course, but also look for those who can champion you. Needless to say, the best referrals are from those who can also support you. Since there is so much emphasis in organizations today on different demographic groups, look for people whose experiences are different from your own. Are you a Baby Boomer? Then look for people who are Gen X or Y or Millennials. Don't worry if you aren't sure what those terms mean; the point is to aim for age diversity as well as ethnic, regional, racial, and others.

Type of Relationship You Want with Clients

You should develop a strategy for achieving the type of helping relationship you want with your internal or external clients. Decisions made at the initial stage have important implications for the future of the relationship. Also, consider what you can offer to and gain from other relationships. Two of the leading experts who have defined consultant/client relationships are Peter Block, author of *Flawless Consulting* (2000), and Edgar Schein, author of *Process Consultation* and *Process Consultation Revisited* (1999). Their works form the basis for the relationships discussed here.

Most performance improvement projects involve more than one type of relationship, and relationships may shift during the course of a project. As Edgar Schein (1999) notes, the consultant ''must choose from one moment to the next which role to be in.''

Pair of Hands. Block (2000) coined this term. In this role, the consultant acts as a direct result of specific instruction from the client and has little influence. He or she is often viewed as a contractor with limited skills. Opportunities to deepen or expand these relationships are rare and control is firmly in the client's hands. The client is definitely in the one-up position and is likely to remain that way. Consultants may work in the ''pair of hands'' on occasion if the tasks revolve around their routine job duties, or if there is an opportunity to learn about a new field. But over time, the most rewarding relationships are those that involve more than your ''hands.''

Expertise. Consultants assume an ''expertise'' role when they offer certain areas of specialization such as workshops or seminars. Quite often the role is similar to many outsourcing relationships today. While they may work closely with the client to present a customized session, they usually have little influence on line managers. Of course, any role can be a platform for developing deeper relationships, but most consultants find it difficult to expand beyond their areas of acknowledged expertise. Clients tend to put the consultant in a box with a special label and call on him or her for limited opportunities. The consultant may even be seen as a peer, but the relationship is definitely at arm's length.

Expert (Block)/Doctor (Schein). In these relationships, the consultant is seen as all-knowing and assumes all responsibility for the project. As Block points out (2000, p. 19), however, purely technical problems are rare. There is a considerable chance that underlying problems will remain unaddressed and client commitment to even identify them, let alone address them, will be minimal. The author once worked on a pro bono consulting project with a group of fellow alumni/ae from the University of Chicago Graduate School of Business. The client was a small non-profit organization facing a serious challenge in fund-raising. The group's board of directors was thrilled to be getting some advice from a group of experienced professionals. Early in the project, the consultants learned they were the third group the non-profit had approached in the last four years; in fact, the organization gave the consultants copies of the previous two efforts. Unsurprisingly, they found the reports contained solid advice the non-profit board had ignored. The consultants realized their recommendations would most likely remain unimplemented as well and withdrew from the project. Their only recommendation was to follow the advice previously given.

In the situation just described, the board of the (now potential) client first tried to explain why it had not been able to follow the solutions outlined in the reports. As any consultant who has served in the expert or doctor role will instantly recognize, the client was trying to blame a refusal to implement an outsider's advice on the poor quality of the previous consultants. The consultants stood their ground, which was made easier because the project was pro bono. Internal consultants or external consultants who have been having some lean weeks or months understand that walking away is often more difficult. In those situations, the best one might be able to do is to begin surfacing the client's resistance as soon as possible. In other words, the consultant begins to act more collaboratively while the client is still thinking he or she is an expert.

Collaborator (Block)/Process Consultant (Schein). In these roles, the client and the consultant form a true partnership, which offers the best chance for constructive change. In this relationship, the client has equal responsibility for the outcome and great interest in successful implementation. They identify and diagnose problems together and develop implementation strategies. There is opportunity for both to grow in the relationship. The ultimate goals are to solve problems so they stay solved and to enable clients to solve similar problems on their own in the future.

Type of Relationship You Want with Peers

Peers can be a significant source of support and opportunity or they can be seen as competitors. In fact, they are often both. The challenge for the consultant is to decide whether to be open in such relationships or to be closed and guarded. The author's experience over many years is to opt for openness when appropriate.

Of course, openness does not mean the consultant should share all sorts of personal information or impose on his or her colleague. Today's peer might be tomorrow's client, and the consultant will want to already have an open, reciprocating relationship.

Advisor/Mentor. One of the most rewarding relationships is that of advisor or mentor. Whether one is the mentor or the protégé, the opportunities for personal growth are wonderful (and sometimes painful). An advisor or mentor is someone who can help the consultant move from a pair of hands, expertise, or expert/doctor role to a more collaborative one.

Listener. Sometimes the best relationship is the one with the person who will listen to you rant (usually for a predetermined time) and say little or nothing. The listener's role is to be supportive and offer advice only when it is requested.

Resource. Some people always seem to know who the best website developer is or where to take a client to a lunch that isn't too expensive yet doesn't leave you looking like a cheapskate. They may not know everyone, but they can get you there in a few steps.

The lesson is that there is some type of constructive, helping relationship opportunity with nearly everyone. Not all relationships need to be as intense as others, but all relationships can be positive. Establishing relationships is often easier than people think. Exhibit 18.1 is a checklist for establishing relationships, and Exhibit 18.2 is a worksheet for identifying target relationships.

Exhibit 18.1 Checklist for Establishing Relationships

1. Learn about the organization's core business.
   - Attend staff meetings, trade shows, and any other gathering of line managers.
   - Shadow managers.
   - Ask to be included on distribution lists.
   - Read the same trade publications managers read.
2. Perform your core functions expertly. Successful consultants find that the best clients are those with whom they have worked closely and who are happy with the results. By proving you know your more immediate field, you will build trust.
3. Avoid being an order-taker by offering suggestions for alternative approaches that will advance business goals. For example, instead of adding permanent staff, might the client contract out work during busy times?

4. Be willing to let go of personal or department-related goals that don't advance business goals.
5. Network assertively to help identify business needs and market yourself and your department. This means attending every company function you can—even the boring ones.
6. Be flexible. Meet one-on-one with managers and listen to their problems before pronouncing your ''solutions.''
7. Ask questions.
8. Approach every situation as having potential for learning and sharing.
9. Attend meetings at lower levels, and let the news ''bubble up.''
10. Anticipate resistance and work to counter it in advance.
11. Publish a department newsletter or write a column for the company newsletter. Use case studies that highlight how different areas of the company benefited from your advice or programs. Use quotes when possible.
12. Write for professional publications. Then share reprints with key internal managers.
13. Speak publicly and at company functions such as sales meetings.
14. Find mentors/coaches. Listen, learn, and leverage your mentors' skills and experience to advance your career.
15. Be a mentor/coach. Be generous in helping co-workers. Most will return the favor.
16. Volunteer—wisely—for projects within your organization. Look for projects that will let you make a real contribution while allowing you to position yourself and your department in the best possible light.
17. Finally, be consistent within your own organization's culture. If yours is more closed, start small and build on successes.

Establishing Relationships

- *Use the direct approach*—Take someone from another area to lunch. For internal and external consultants, consider informational interviews. For example, you might ask a client to review an article you've written for the company newsletter or other publication.
- *Offer to help a colleague with a project he or she is working on*—Your assistance need not be extensive; it can be as simple as giving your colleague a good resource or reviewing a draft of a report.

Exhibit 18.2 Identifying Target Relationships

Think of a current or potential client. Briefly describe your current relationship. Next, think about what your desired role as a consultant would be. Last, create a strategy you can use to begin moving into the desired relationship.

Identify Client Relationship

Client:

____

____

Current Relationship:

____

____

Desired Relationship:

____

____

Strategy:

____

____

Other resources that will help me implement this strategy:

____

____

- *Perfect your elevator speech*—Sometimes the briefest of encounters can lead to long relationships. Remember the strength of weak ties: casual conversations often lead to valuable results.

CONTRACTING

The second step in managing relationships is to contract successfully. You, your client, and, if you are an internal consultant, your boss must have a full, clear understanding of a situation before agreeing to taking on a client's project. Schein (1999) states the basic purpose of any agreement: *make visible that which is invisible*. To ensure that the relationship survives intact, you must begin making everything visible from the very first contact. The maxim has a few limits, of course. Mentioning successful projects can build the client's trust. Describing them in great detail or dwelling on past mistakes can be a tactical error. Also, remember that your first job is to make sure you understand what the client needs. Avoid citing your own experience until you know it is relevant to the current discussion.

Early in her career, the author was asked by a colleague to meet with a client of his who needed documentation for a new computer system. Her colleague had been a successful salesman with IBM and assured her the meeting was probably a formality, but one never knows for sure. The author had several recent documentation projects to show as samples and made sure to bring as many as would fit in a canvas bag. She was concerned about whether the client would trust her experience so, at the first opportunity, she shared the samples with the client. She cleverly pointed out, in some detail, how each one applied to the project under consideration.

She evidently did a masterful job; by the end of the meeting, the client had offered her the assignment. Afterward, she asked her colleague for feedback, and they agreed to meet at a nearby coffee shop. She took a seat, waiting for his praise on her selling technique. Instead, his first comment was, ''Do you have to talk so much?'' Evidently, there is such a thing as too much visibility.

The lesson is to avoid revealing everything in the first meeting. You will seem overeager and unsure. Rather than impress a client, you may convince him or her that you care only about your perspective.

What You Can Give

Performance improvement practitioners must know what they have to offer clients beyond a specific set of skills. Their first responsibility is to discover, or help the client discover, what the challenge is. At the same time, you must have a clear understanding of what you can give to the client—not so you can recite your qualifications at the first opening, but to have the confidence in yourself that you can help the client.

Perhaps you want to revisit your SWOT analysis at the end of the chapter. In addition to strengths such as knowledge and experience, consider your ability to trust others, be honest but tactful, and know how to support others. Then there is the issue of genuineness or, as Peter Block terms it, authenticity. Other strengths include interpersonal skills such as listening or being able to bring disparate groups together.

What You Have a Right to Expect

Block's key message is that you have a right to succeed and, therefore, you should be able to expect honesty, fairness, access to confidential information and people who have been involved on both sides of issues. You must also feel confident that the climate in the department is such that the manager will be able to implement the process you help him or her develop. When you sense that your client is a less-than-willing partner, consider whether you should continue the relationship. (See Handling Resistance later in this chapter.)

BUILDING THE CONTRACTUAL RELATIONSHIP: THE SUBSTANTIVE SIDE

A contract is not just a piece of paper stating conditions each side will meet. Long before the contract is written, the client and consultant must agree on what will happen, who will do what and why. While written contracts are important, they are a reflection of the relationship rather than a simple list of deliverables and fees. Every contract reflects both a substantive, factual agreement and an emotional commitment. Exactly what are you and the client agreeing on? Ask direct questions. You need to show that you want to provide a real solution the client can own. The following questions will help you start. Listen to both the substantive and emotional parts of the answers. You do not necessarily need to ask all the questions yourself, but you do need answers to them. Watch your tone. You are asking substantive questions, not challenging the client. If the client react emotionally, explore any misunderstandings. Don't let steam build up. See the checklist in Exhibit 18.3.

Exhibit 18.3 Preliminary Contracting Checklist

Client's Answer	**Effect on Your Approach**
What created the need?	
What if we do nothing?	
How will you evaluate success?	
Who approves? What are his or her needs?	
What is the schedule?	
What is the budget?	
What is the scope?	
Your Answer	**How to Resolve Differences**
Are the client's objectives realistic?	
Can the client's need be met with the objectives you've agreed on?	
What will realistically be different after your intervention?	

What Created the Need

The start of most helping relationships is to understand where the pain is. When there is no urgent need or when the client has to think too long to answer, you should be concerned about the client's motivation for finding a solution. Either the client doesn't know how the need came about, doesn't care, or doesn't trust you. Most people like to tell stories; in contracting discussions, people who feel a need are eager to share. You should be prepared to listen to several stories. That's OK. Listen to them.

What If We Do Nothing

At some time during the life of any project, someone is sure to ask you, your client, or your client's boss, ''What if we do nothing?'' While there are times when doing nothing may be the best option, you must have an answer. Many projects are derailed because a person cannot answer this question succinctly. In addition to saving a valuable project from the dust heap, you avoid looking like someone who doesn't know what he or she is doing or why you are doing it. Asking about the ''do nothing'' options helps in other ways:

- First, you can *further sharpen the insights* gained from asking about what created the need. The need for a project is often found in understanding what will happen if action is not taken.
- Second, you can get a further *sense of the client's commitment* to the project. If a client isn't committed, your chances of success drop dramatically. Avoid injecting your own scenarios in hopes of motivating your client. The problem may be that he or she knows the probable outcome but either doesn't care or in some way prefers it.
- Third, you help the *manager sharpen his or her responses* to others who will inevitably ask the same question.
- Last, but certainly not least, you will gain many points to *help you identify the benefits* your eventual proposal will offer.

How to Evaluate Success

Some clients will know immediately how they'll evaluate your success and will not hesitate to tell you. This response is especially true if there are readily available metrics such as retention or recruitment. For those who have more difficulty in identifying a way to measure success, consider a balanced scorecard approach. (For more about balanced scorecards, see balancedscorecard.org.)

No matter what metrics you and the client agree on, use caution when agreeing to measures of success. Consider how much your work will affect the final outcome. For example, suppose you are asked to help a client whose department has experienced problems with harassment. While an obvious sign of success would be to reduce or eliminate the problem, there are many factors that may be beyond your or the client's control. Still, fewer complaints or more time between complaints may indicate improvement. Try to find a metric on which both you and the client agree. While attending training on the topic is no guarantee that problems will be eliminated, it may be the only reasonable measure to expect.

Who Approves and Their Needs

Quite often, you may find yourself talking to a proxy for the true client. The proxy may lack even basic information and may or may not fully support the effort. Can you talk to the approvers? If not, do you trust the proxy client to be giving a fair representation of the client's views?

What the Schedule Is

In most situations, neither you nor the client will be able to set a firm, detailed schedule in your initial discussions. You should, however, be able to agree on key milestones. Use a calendar when setting any dates. Be sensitive to religious holidays other than your own. A consultant was embarrassed to learn that he had scheduled a key meeting on the Greek Orthodox Christmas. He had thought everyone would be back from their various holidays by then.

When you prepare a formal proposal, you want to examine all milestones carefully and always allow a bit of room for slippage. If you notice anything that might affect the milestones, even at this early stage, be sure to discuss them with the client. For example, productivity in most organizations falls during the period from mid-November through December as people gear up for the holidays. While you do not want to tell the client you assume low productivity during these times, you can mention the challenge of conflicting vacation schedules.

What the Budget Is

Some clients are reluctant to share budget numbers with anyone, while others give the information freely. Whether you are internal or external, you need at least a sense of the budget to determine the best approach to take. The author has found that, when clients resist sharing the information, she can usually disarm the concern by explaining that she wants to focus on the best solution *and* meet the practicalities presented. Even if clients still refuse to give a firm number, she has found they will usually give some indicators ranging from "Our budget is very tight this year" to "We want to make sure we do this right."

What the Scope Is

Work with the client to understand what the project does and does not cover. The goal of defining the scope is to:

- Clarify the presenting problem. Believe it or not, you may still not have a clear understanding of the client's problem at this point in your process. As you discuss the scope, ask yourself whether both you and the client clearly understand the problem.
- Uncover peripheral issues and address underlying concerns.
- Agree on limitations to the project (if any).
- Identify who owns what parts of the project.
- Confirm objectives.

EVALUATING THE PROJECT BEFORE CONTINUING

Pay attention to any small red flags before continuing with the consulting engagement. Every consultant has stories of not listening to that little voice telling him or her to abandon a project and in a few weeks or months realizing why he or she should have listened. The outcome is often a failed or damaged project and relationship. Before deciding the client's expectations are out of line, make sure you understand all the facts.

- *Are the client's objectives realistic?*—No client wants to hear halfway through a project that you thought his or her objectives were "a bit out there" from the start. If you think the client's objectives are unrealistic, treat the situation as a learning opportunity. Rather than rolling your eyes or otherwise disagreeing, ask a few probing questions. If you decide, after several probes, that the client's objectives are truly unrealistic, you must tell him or her. Avoid taking a judgmental approach. Instead, begin the discussion with something like, "I think your objectives are very admirable, but I am concerned whether we can meet them with the information I've gotten so far. What am I missing?"
- *Can the client's need be met with the objectives you've agreed on?*—The author once sat in on a meeting with two client representatives who were determined that they would have specific, measurable objectives for a program to teach factory supervisors to be more participative. It was a noble goal, but the two employees discarded objective after objective because it was too fuzzy. Rather than look for ways to quantify the objectives, they kept searching for things they could measure. Those items became the objectives. The author reported back to her client that she didn't think she could add any value to the project because the two employees had a very clear vision of the path they felt impelled to take. The author wasn't sure whether it would succeed or not (it might have), but her own reluctance told her she couldn't bring her best work to the project.

- *What will realistically be different after your intervention?*—Clients often think of projects as ''silver bullets'' that will transform their organizations. Such talk is a clear signal that clients might have unrealistic expectations. Someone must take time to clarify expectations so everyone can agree on them. A simple ''Wow, that would be wonderful. Is that your expectation for this project, or do you anticipate more phases? Tell me more.''
- *What specific outcomes do you and the client hope to achieve?*—Remember, you and the client have an equal right to succeed. If you believe that either of you is not going to be successful in the project, now is the time to disengage. Internal consultants often feel caught in such situations. They must keep in mind that they do a disservice to themselves, their clients, and the organizations for which they both work if they proceed with projects that have a poor chance of succeeding.

How You Can Say No or Postpone Demands You Might Not Be Able to Fulfill

While you have a right to succeed, you can jeopardize that right by taking on more than you can accomplish. Before saying ''Yes,'' ask yourself:

- Is this really a request or just a comment? Clients (and consultants) often like to theorize how a perfect situation might evolve—in a perfect world, that is. Less experienced consultants (and even some who should know better) sometimes interpret the comments as real requests and may fret about being able to take on yet another commitment. The first step is to confirm the request.
- Does this request fit with other work/personal/professional goals? Some consultants, internal and external, feel they must say ''yes'' to every opportunity because they must fill every available hour. You don't need a detailed five- or ten-year plan, but you should have a clear sense of when a project is right for you.
- Do you have time or resources available? The old saying is, ''If you want something done, ask a busy person.'' On the other hand, if the busy person is approaching burnout, the results may be less than desired. Every consultant can recall a time when he or she took on too much and both the consultant and the projects suffered.
- Do you want to do this? The point of this question is not that you accept only projects that are high- (or low-) visibility. You want to be honest with yourself about your motivations. If you have concerns about the project, now is the time to raise them.
- What will realistically happen if you say no? The other side of the ''ask a busy person'' adage is the belief that the consultant is the best person for

the job. No one else can do the job nearly as well. In truth, many people could do a good job. Another reason for avoiding a ''no'' answer is the fear that the consultant will be punished in some way for the answer. Saying ''No'' is hard! The following process may help.

- Express a desire to help. (I'd love to help, but . . . .)
- Give facts about other commitments.
- Mention a project your boss just gave you. (You must actually name it.)
- Explain honestly why you aren't the best person for the assignment.
- Negotiate.
  - ''When do you need this?'' (A neutral question buys time and begins negotiations.)
  - Get an extended deadline on another project with the same client.
  - Delay the start date.
  - Agree and give dates that work for you.
  - Lower expectations—for example, a verbal report instead of written.
  - Ask your boss to set priorities.

Evaluate the Engagement

Intelligent questions are as meaningful as answers. Before accepting any engagement, ask the following questions:

- Is this project valuable to my organization?
- Do I understand the client's needs?
- Are the needs realistic?
- Do I already have all the information I need? One characteristic of a good proposal is that it is complete. There is no law that says you can't contact the client to ask for more information or to clarify a misunderstanding before presenting your business case.
- Does the client have the information? If he or she doesn't have the information, do you really need it? If the client insists on information neither of you has readily available, is it a fair request? If it seems excessive, will the client pay for it?
- What type of additional information do I need? Is it numerical or more subjective?
- Is it available within the client firm? Will the client help me find the right resources?
- Is it available from outside resources? Which ones? Who will absorb any costs? You must clarify these questions with the client.

- How precise do you need to be? Most business decisions are made with less than complete information.
- Have I clarified the deliverables? What specific outputs will I deliver to the client? A consultant contracted with a client to develop a custom training program for what she considered a fair fee. At the end of the development phase, she discovered the client also expected her to deliver the program—twice! Fortunately, they were able to compromise, but the relationship was damaged. Other examples of deliverables would include a typed, written final report versus a less formal PowerPoint presentation.
- Have I surfaced all my assumptions? In the previous example, in addition to stating her assumptions, the consultant could have indicated a price for delivery. It is common practice to specify software packages the consultant will use, including version numbers, in a list of either deliverables or assumptions.
- Do I understand the form these deliverables will take: full report, presentation, and other options?
- What items am I not delivering, as agreed on when defining the scope?
- Will the client be sufficiently available? When possible, opt for face-to-face meetings to build stronger relationships. Some clients prefer lots—but that in itself may be a sign of resistance. Beware of clients who never want face-to-face meetings.
- Will I make formal presentations, or will they be informal updates?
- What are the arguments for proceeding with the project (the rationale)? You will need to provide sufficient background information, including present and future operating environments.
- What benefits, both tangible and intangible, will result from the project?
- What is driving the need for this project now (the justification)? Address the key drivers or reasons for your project. Examples are reliability, environmental, regulatory, safety, and financial. Some projects may have multiple drivers, while others have only one.

The Risks and Assumptions

Before concluding contracting discussions, be sure to uncover any significant events that could influence this project. For each event, discuss consequences, financial impacts, and contingency plans. Items to consider:

- Regulatory—Is a deadline or authorization needed?
- Human resources—Are necessary staff available? Or do available staff have the expertise to achieve the goals? What is the current attitude of

Exhibit 18.4 Worksheet for Evaluating the Engagement

Assumptions	**Effect(s) on Project**
I will be allowed to spend at least three-quarters of my time on this project.	Stopping and starting will delay the project.

the workforce? Motivated? What change management plans will be needed?

- Economic assumptions related to the project—Does a weak or strong dollar affect the project? Do interest or unemployment rates matter?

Assumptions

List the assumptions associated with the project and their effects on the project. Use Exhibit 18.4 to evaluate the engagement. Typical assumptions would include staffing levels, past budgeting levels, and corporate structure (no mergers, etc.). You need to clarify as many assumptions as possible. At the same time, you may not need to discuss all of them in advance, depending on your relationship with the client (and your boss, if you're internal).

MANAGING RELATIONSHIPS: THE EMOTIONAL SIDE

The goal in managing the emotional side of relationships is to build trust with the other person. Most of the time, people will meet you at least halfway. Until that sense of trust is firmly established, however, you must take special pains to avoid miscues that can harm it. Quite often, just taking care of the substantive part of your relationships with your clients is so time-consuming that it's easy to overlook subtle shifts in the emotional relationship.

Your first reaction might be to ignore that small voice telling you that something has changed. In many instances, you're probably right in doing so. The client may be distracted with other issues and may not be focusing on your meeting or conversation. But don't completely discount what your body is telling you. In fact, this is the precise time to listen to your body. You might be picking up on early signs of resistance or differences in communication styles.

Make sure you are communicating effectively. Poor communications can result in small misunderstandings that can grow quickly. A simple barrier, such as a different assumption about what a word means, can lead to difficulties. For example, an external consultant was talking on the phone with a client on a Friday morning. The client was rushing to a meeting and asked the consultant to call back "next Tuesday." The consultant understood the date to be a few days after the day of their conversation. When she called the client the following Tuesday, she was unprepared for the client's rather rude reception. "I told you I didn't have the information you need when we talked on Friday. Were you listening? I said that I'll try to have it for you next week!" The client understood the word "next" as the Tuesday a week after the upcoming Tuesday. Unfortunately, the consultant reacted a bit defensively, and the relationship was rocky for several weeks. The whole situation could have been avoided if they had clarified which Tuesday the client was referring to.

Business people spend up to 80 percent of their work time communicating. Of that time, nearly half is spent in activities that involve listening. In addition, communication focuses on the two business elements that determine the success or failure of people or businesses: information and relationships. Because information and relationships are so important, understanding what successful communication is and how to achieve it is critical.

The word communicate comes from the Latin word *communicare*, which means "to share." Therefore, communication means sharing information or sharing yourself. The goal is mutual understanding—making "visible the invisible." The main objective for all communication—whether information- or relationship-based—is results. It requires an interactive, dynamic process.

Communication barriers often come up during work (and other) relationships. The best way to overcome them is to first admit that they exist, at least to yourself. Then select a strategy for overcoming them.

Hidden Agendas

A hidden agenda may drive a conversation. Your client may have a goal he or she is unwilling to share, yet tries to steer discussions and decisions toward that goal. Hidden agendas may be unintentional. For example, your boss may not tell you all you need to know to complete a project. The omission is not a result of a purposeful effort to thwart you; rather, it may stem from a distraction.

On the other hand, some people may hide their true agendas. They are not open to someone else's message. Personal agendas steal attention away from what others are saying. If the listener has other priorities, he or she will tend to give less attention to what others say.

For example, you might be scheduling interviews with people your client identified as key to your new project. One manager is available only late in the day. At the same time, you promised to provide the snacks for your daughter's soccer game for that same day. You don't want to tell the manager you're concerned about a conflict with an eight-year-old's soccer game. By hiding your agenda, you have created a potential for misunderstanding. If you acknowledge the possible conflict in your own mind, you can decide whether you need to call for backup on the snacks or try to find an alternative meeting time. You do not necessarily need to share the purpose of the scheduling conflict, but you do need to resolve it.

Resist the temptation to work from a hidden agenda. Be open to all ideas. Remember that you might find something of benefit to you, even if you don't expect it. Stay tuned in, even though it's difficult; you never know when a useful idea will pop up. Ask questions for better understanding. For instance, the manager who was looking for information on the meeting found that satisfactory results took longer than she had intended. Because the communication was not effective the first time, the employee had to spend more time obtaining the information the manager wanted.

Clear and Unambiguous Words

Language does not determine how anyone thinks: it transmits thinking. Words often mean different things to different people—for example, common jargon in one field becomes meaningless to outsiders. Meaning comes, to some extent, from each person's experiences and perceptions. Seemingly ordinary words can be ''loaded'' for some people. If ''discipline'' implies a reprimand to you, it is a loaded word. If someone else interprets it to mean strong, self-starting work habits, it does not carry the same emotional baggage for him or her. Remember that words take on meaning based on people's life and work experiences. Ask questions to make sure you and the speaker are using the same definitions for words:

- Does the speaker use any words that act as red flags for you? Are there any words you should avoid when responding?
- Avoid words you know are likely to inflame others. Also, be aware of your own reactions. Try to keep your temper when someone uses a word or phrase that is a red flag word—at least until you are certain of the speaker's intent!

Facts or Opinions

People often differ in what they consider fact. If you are working to improve the health of the people in your department, you might cite as fact articles on eating various foods and the effect of food on health. On the other hand, someone else might think the findings are opinions of well-meaning researchers—but lack the rigor necessary to call them facts. This difference is a barrier to communication. Check out what you accept as fact and what would satisfy you in presenting such a case. One person might have more technical information than another person has. Recognize the difference between facts and opinions.

Do you really understand the other person's perspective? Ask questions to make sure you really do disagree. Try restating the other person's position before offering your own ''fact'' or ''opinion.'' Effective communicators are flexible; they recognize that they and another person may never agree on what the facts are, but in most cases they can respect each other's interpretation.

Interpersonal Barriers

Interpersonal barriers may exist between the speaker and listener—different values, difficult dealings in the past, or poor ''chemistry.'' The speaker's style or personality can turn people into poor listeners. If the person talks in a boring, unemotional tone of voice, others will tune out. If the speaker is someone who comes on as a ''know-it-all,'' listeners will also tune out. If you simply don't like the speaker, you might listen for just enough information to contradict him or her and miss the point completely.

In these situations, first acknowledge differences so that the message is not lost in the middle: ''I know we sometimes have different views of things, but let's see if we can find a way to help each other.''

Clarifying and mirroring can also make you and the speaker feel like you are part of a team, overcoming your differences for the good of the organization. Use ''we'' statements to foster a feeling of solidarity. Another way to break down interpersonal differences is to substitute ''I'' messages for ''you'' messages. ''You'' messages can sound accusatory or blaming: ''You don't follow my instructions.'' ''You'' statements often prevent communication because the listener is caught in an emotional trap, feeling accused or blamed. The statement, ''I hope my instructions were clear,'' can better promote communication, demonstrating that you are open to all possibilities of why things went wrong. ''I'' statements also leave other people with their dignity intact, even though the communication that follows may reveal that they need to make some improvements in performance or behavior. No one loses when communication is open to all possibilities. Respect and friendship can be retained by openly asking for ideas and opinions.

Distractions

If you're talking on the phone to a colleague who is explaining an important procedure and several people talk loudly as they walk past your cubicle, you will usually focus on the distraction rather than on the phone conversation.

Distractions can be mental or physical. Some examples of mental distractions are fear, worry, being unprepared, daydreaming, boredom, poor self-esteem, and anger. Sometimes, fatigue or poor health limits your ability to listen well. You may fully intend to devote all your attention to the other person, but your physical state won't allow it.

Physical distractions can also include other people and outside noise. Telephones, and the perceived need to answer them, also make poor listeners of many people. Here are some questions to ask yourself when you notice you are being distracted:

- Can you go somewhere else to avoid noise? Close a door or window to muffle it?
- Can someone else answer your phone for a few minutes? Are you in a public place where others can overhear? Is privacy important for this conversation?

Unless you do something to eliminate distractions, you are not likely to be able to get the speaker's full message.

Prejudging

Sometimes you may decide in advance what someone else is going to say—either positive or negative. Think of a situation in which you are sure you can predict what the other person will say, even before he or she has said a word. In these situations, you often end up hearing what you want or expect to hear. Closing your mind to a subject or an opinion blocks effective listening. Emotionally charged issues can bring about long discussions but very little listening. Avoid prejudging.

Ask neutral questions to clarify the other person's message. How do you feel about the speaker? Are you convinced the other person is wrong before saying a word? How will this affect what you hear? Try to imagine how you would react if another person were communicating the same message. What is the speaker's knowledge about the topic? Do you know? How important is the topic to you? To the speaker? To your company? Examine your own prejudices and keep an open mind.

Processing Speed

Most people think at a rate of around five hundred words a minute. The average person speaks at about 125 words a minute. The gap between processing speed and talking speed lets our minds wander while others talk. So you have much ''room'' left over for considering other things—work you have to do, phone calls to return, what to have for dinner, a bill you forgot to pay—as more pressing priorities.

Writing speed is even slower than speaking speed. Taking notes can help you focus on what the speaker is saying. Resist letting your mind wander. Instead, use the extra time to listen better—process the information by outlining the speaker's ideas, summarizing them, or categorizing them.

Organizational Barriers

Another important factor is organizational barriers—when someone has more power, status, or authority. For example, when you talk to your boss or people who report to you, you may encounter organizational barriers. You may ignore what a new person says just because he or she is inexperienced. On the other hand, you may ''freeze'' when your boss speaks to you.

Jobs also sometimes cause people to view situations differently. For example, most salespeople view paperwork as nonproductive. Accounting people value paperwork as something they need to create financial statements. These differences can create a barrier to communication.

If you're the boss, arrange settings that show you view the person as an equal, with a point of view you want to hear. If you are the employee or peer of the speaker, use reflective statements and other active listening techniques that indicate you think that you and the speaker are on the same level.

Physical Barriers

Physical barriers include floor plans that isolate people from each other. If employees in different departments must communicate often with each other, and they are on different floors, in different buildings, or in different cities, they must overcome significant physical barriers to be sure their messages are received.

Listening also suffers when the environment works against communication. A room that is too hot or cold can distract you from the speaker's words. Trying to discuss a confidential matter in a crowded or noisy room creates challenges for both speaker and listener. Other physical barriers are wearing dark glasses or putting a desk between you and the speaker.

If the physical barrier is distance or structural (you're on one floor and a person you must work with is on another), see whether one of you can relocate, at least temporarily. Perhaps you can meet in a neutral location. If you are in your own office, sit next to the person, rather than put your desk between you. Take off your dark glasses so you and the speaker can make eye contact.

Gatekeepers

Gatekeepers, another barrier, are filters people use when taking in information. The three main gatekeepers are deletion, distortion, and generalization.

You *delete* when you leave out parts of a message the speaker is sending. For example, think of a meeting in which you were tuning in and out, letting your thoughts drift and missing much of what the speaker was saying. As a result, the words you did hear left you thinking, ''He doesn't really care about employees.'' Having heard only part of the message, you draw conclusions that may not be valid. Deletion is a problem when you hear only what you want to hear—either positive or negative—and ignore the rest.

Distortion is when you enlarge or diminish your experience. You look at it as in a mirror that distorts size and shape. For example, if you pass a co-worker in the hall and she doesn't speak, you might decide she was intentionally ignoring you. Without more information, that would be distortion.

Generalization is when you look at only certain aspects of a situation and build them into broader conclusions, ignoring the exceptions. For example, ''I can never get to sleep before 2:00 a.m.'' Chances are that you do get to sleep earlier. Generalization helps you respond to new situations based on previous experience, but it can be dangerous if you close off learning from new experiences.

Gatekeepers are also ways of responding to information, ideas, words, and even nonverbal signals. You've learned filters throughout your life—such as biases, experiences, and expectations. For example, someone may decide she doesn't like ''new jazz.'' She will use this as a filter to delete really listening to the music and generalize about all such music.

You can usually manage filters by first becoming aware of them, then acting to control or eliminate them. Think about the barriers you put up. Are any of them gatekeepers that keep you from listening to someone?

Effective listening involves good habits. Forming good habits requires awareness, practice, and sometimes the help of others. Understanding why people listen poorly will help you focus on creating better communication skills. One of the greatest barriers to communication is when people simply do not listen to each other. Communication is about give and take. You can always hold up your hand and say, ''Wait a minute. Can we backtrack so I can see if I understand you correctly?''

Consider difficulties you've had in communication. Were barriers involved? What effects did they have? How might you have resolved them? Use the form in Exhibit 18.5 to help you analyze barriers to communication.

Verbal and Nonverbal Communication

People communicate two ways—verbally and nonverbally. Both methods share information and reflect the relationship between the speaker and the listener. As long ago as 1971, Albert Mehrabian at UCLA conducted studies that indicated that, in situations that have an emotional component, body language and tone of voice can contribute as much as 90 percent of the communication of attitudes, while words contribute only 10 percent.

Exhibit 18.5 Analyze Barriers to Communication

Barrier	*Effects on My Communication*	*How I Might Resolve Them*
Hidden agenda		
Words		
Facts versus opinions		
Interpersonal differences		
Distractions		
Prejudging		
Speed differences		
Organizational		
Physical		
Deletion		
Distortion		
Generalization		

When there is a conflict between your words and your body language, most listeners will ''listen'' to the body language. And if the nonverbal message contradicts the verbal one, the nonverbal message wins nearly every time.

Verbal Messages. Verbal includes written and spoken messages. It refers to the words we use and how we speak them. For example, if a customer calls a company's office, the person answering the telephone might say, ''How can I help you?'' The words seem to indicate the person is ready to help the caller. On the other hand, depending on how the person answering the phone says the words, the customer could read the tone as:

- The call is welcome.
- The caller is interrupting.
- The person answering is attentive.
- The person answering is angry.
- The person answering considers the caller a ''pain.''

Indeed, how we use the words is often more important than the words themselves. Think about a recent conversation during which the inflection affected how you interpreted the other person's words.

Our jobs demand that, as consultants, we listen carefully to the whole message from clients, bosses, peers, subordinates, vendors, contractors, and even competitors. You must listen to your customers to ensure that your organization offers the best products or services. Active listening means the listener becomes involved in the communication process and takes on part of the responsibility to make sure he or she understands the message being sent.

The message received *is* the message. To ensure that you receive the message the speaker intended, monitor your own nonverbal communication and emotional filters. Try to put yourself in the speaker's position and glean the whole message by focusing on more than the words. Look for body language and tone as well to ascertain the whole message he or she is sending.

Nonverbal Messages. Nonverbal messages are powerful in communicating the attitudes behind the words. They can signal whether or not you like the other person, your perception of your status or power in relation to the person, how responsive or attentive you are, and how accepting you are.

We communicate nonverbally in many ways: eye contact, facial expressions, gestures, and posture. Examples of these are rolling eyes in disagreement, wringing of hands, tapping a foot or fingers, shaking or nodding the head, shrugging the shoulders, and leaning toward or away from the other person. Other nonverbal ways may be physical, such as touching—a handshake or a hand on the shoulder.

The distance between people is also a form of nonverbal communication. The term *proxemics*, introduced by anthropologist Edward T. Hall (1966), describes how distances between people affect the way they interact. Loosely stated, social or personal distance between people is reliably correlated with physical distance. When one person moves physically closer to another, he or she is indicating a closer personal relationship. Acceptable space varies between cultures, with Latin cultures comfortable being closer and Nordic cultures preferring more distance. You may find it hard to listen to someone if you are uncomfortable with his or her distance from you.

Silence, clothing, personal grooming, and even office furniture are other examples of nonverbals. Although little valid research has been done on these, everyone has a personal story about how such factors affected communication between him or her and another person. Nonverbal messages quickly convey acceptance, impatience, anxiety, agreement, disagreement, puzzlement, warm feelings, or discomfort, all without speaking one word.

Nonverbal language is not an exact science. You cannot be guaranteed that specific gestures always mean the samething to all people. For example, when people sit with their arms folded, they may be indicating a closed attitude or simply that they are cold.

Your interpretations of body language are likely to be more accurate when you observe a person over a period of time, enabling you to correlate gestures with attitudes. When you sense negative attitudes in body language, modify your message. Back off, make your message milder, or stress the benefits to the listener. Keep modifying until the attitudes seem positive again. Of course, be mindful of your nonverbal messages as well. To enhance your verbal message:

- Make eye contact with the listener. (It's OK to glance away from time to time, but be sure to make eye contact again.)
- Speak in a pleasant tone of voice at a reasonable pace.
- Use positive facial expressions and body positions—smile, nod, and look alert and interested.
- Lean toward the other person.

Active listening is a process by which you listen attentively to the speaker, give signals that you are listening and understand what he or she is saying. People rarely agree with others just by getting more information; they are more likely to agree when they feel understood.

The process of active listening forces you to overcome barriers to respond correctly and effectively. Active listening does not mean that you agree with or unquestioningly support the speaker. It means paying attention to the words used and how the speaker delivers those words.

Active listening brings in your visual and emotional channels as well as the auditory senses. It involves four key techniques:

- Attending
- Restating and paraphrasing
- Asking questions
- Reflecting

Attending. Attending is giving the speaker signals that you are listening and that you understand what the speaker is saying. You do this in the following ways:

Eye Contact—Maintaining eye contact shows the speaker that you are interested in what he or she is saying. Gazing into the distance, staring at the floor, or daydreaming shows a lack of interest and shuts down communication. The speaker then spends more time trying to get your attention and may lose track of the original message.

Oral Comments—Comments or questions that move the conversation along tell the speaker that you are listening. Uttering phrases such as "Yes," "Umhum," "OK," "I see," "I understand," or "How interesting" signals that you are paying attention (as long as your nonverbal signals match your comments). If you are busy shuffling papers, reading, or walking away, obviously the speaker will spot your insincerity.

Posture—Sitting back in a chair with arms folded and feet propped on the desk sends several signals. The person may be nonverbally saying, ''Go ahead, tell me something I don't already know!'' Or the posture may indicate boredom or a display of power. When a person leans forward, it can signal a willingness to pay attention. A person who sits on the corner of a desk and peers down at you also sends signals of power or superiority. A boss who invites you into the office for a meeting and then stands while talking generally indicates a short meeting. If he or she comes to your office, sits down, and draws the chair close to your desk, chances are the topic is important and the meeting could take a while. Adjust your posture to let the speaker know you are interested and willing to listen.

Gestures—Nodding your head in agreement or shaking it in disbelief shows the speaker you are listening. You can do the same by shrugging your shoulders or by tilting your head. Make your gestures consistent with the content of the message. If the speaker says, ''I think everyone agrees that we should go ahead with this new plan,'' and you nod your head, the speaker assumes you agree. If you disagree, you will be more effective if you frown slightly, shake your head, or tactfully state your opinion.

Restating and Paraphrasing. Active listening can include repeating or paraphrasing the content or feeling of the speaker's message as you understand it. To properly reflect the speaker's words, ideas, or feelings, you must listen. You can use the same words as the speaker: ''He's a great worker'' or paraphrase what is said: ''He's a great worker. You're impressed with his productivity.''

Asking Questions. Asking questions helps you get at the message ''between the lines.'' Short questions such as ''Really?'' or ''Is that so?'' work. Some good questions and comments that serve as door openers to encourage the speaker to keep talking are ''Tell me more about it'' ''I'd like to know what you're thinking about this'' ''Would you like to talk more?'' and ''Let's discuss it.''

Long questions that challenge or distract the speaker are not active listening. They are often signs of not listening. Use either open-ended or closed-ended questions to show you are listening. Open-ended questions are those that require a longer answer. They are usually seen as less threatening and ''good'' when you are trying to build a relationship as well as get information. For example, ''What happened at the meeting?'' is open-ended and sounds like the listener is ready to spend time with the speaker. Use the speaker's responses to form follow-up questions.

Closed-ended questions require short answers; sometimes one or two words are enough. ''Was Joe at the meeting?'' is closed-ended. These questions are more efficient in gathering information, but too many of them at one time take on the tone of an interrogation instead of a conversation.

When you make comments or ask questions that keep the other person talking, you are in listening mode. If you comment or question in a way that threatens, challenges, criticizes, or judges the other person, you are no longer listening.

Reflecting. Reflecting is tying into the emotion behind the words. It takes the most time and is the most powerful of all the active listening techniques. Use reflecting when you want to listen for relationship as well as for information.

You can reflect the feeling behind the statement. ''He's a great worker.''

''You enjoy having him on your team,'' reflects the speaker's feelings.

Reflection keeps the speaker talking. It signals that you are listening because you care about the speaker, the topic, or both. It shows you're not listening just to challenge or criticize.

Consider how reflecting only the content keeps the following conversation friendly and helpful.

SPEAKER: ''I'm really excited about my new job.''

LISTENER: ''You have a new job.''

SPEAKER: ''Yes, and I've already learned how to use the word processing and spreadsheet packages.''

LISTENER: ''You're really into this computer business!''

SPEAKER: ''And how! Not only does it save me money, it has already paid for itself by allowing me to take on a project I would never have been able to complete on time with just a typewriter.''

LISTENER: ''You're really getting your money's worth from your investment.''

SPEAKER: ''Best money I ever spent.''

The following example reflects the content and the feeling.

SPEAKER: ''I can't wait until they've completed that new house behind us.''

LISTENER: ''You're not too happy with the construction going on in your neighborhood.''

SPEAKER: ''Yeah, I wish they'd pack up their tools and go home.''

LISTENER: ''You're really upset with that project.''

SPEAKER: ''And how! Every morning at 7:00 sharp they start their incessant banging and sawing and drilling and chattering. They drive me nuts.''

LISTENER: ''Their early starts really disturb you.''

SPEAKER: ''Normally, I wouldn't even be at home so I wouldn't care. But just last Monday, I was put on the midnight shift at work. I can't get any sleep with all that noise.''

Notice how the real problem did not surface until the speaker reflected three times. Only then did the listener understand why the speaker became so upset at the new home building project. Here are a few samples of reflections of both feeling and content:

You're ____________ because ____________.

You feel ____________ about____________.

You seem____________ with____________.

Using neutral comments lets you learn more about a situation before committing yourself to a position. Active listening also creates time to cool off when emotions might be heated. While the other person speaks, you can use your ''think speed'' to decide how involved you want to get with the other person or the situation he or she is describing.

When you know you have received the whole message, you can plan a better response. Clear responses let the speaker know that the message was clear—or that he or she needs to revise it to make it clearer.

Many people overlook good ideas because they do not focus on the speaker, preferring to spend the time composing their next comment. By listening closely to the speaker, the listener often gets good ideas or can respond to work with the speaker to build better ideas.

For example, the marketing director in a telecommunications company was discussing the appointment of yet another president (the third in two years). ''I'm willing to give him a chance,'' she said. ''At least when you talk with him, you feel you've been heard. That's some improvement.''

Active listening lets you show the speaker that you care about what he or she is saying. It also puts you on the same level as the speaker, which can lead the speaker to recognize you as a co-equal, or at least in a higher status position than before.

Think of the people who are known to be ''good listeners.'' Now think about who usually gets recognition for good ideas. Are they the same people? They are in many organizations.

In the 1920s, a salesman for a Procter & Gamble was taking a train from Kansas City to St. Louis when he was engaged in conversation by a salesman for a rival company. With only a little encouragement, the man couldn't stop talking about the new ''flaked'' soap the company was introducing. The P&G representative got off at the next station and wired the company, which sped up its introduction of Ivory Snow. It quickly cornered the market while the rival Palmolive Beads languished in a distant second place. P&G told the story to new recruits for years to admonish them to protect company secrets, but it is also an excellent example of the potential benefits of active listening.

Sometimes just attending to what the other person is saying will give you valuable information. In other situations, a few reflecting questions will do the trick. Either technique should help you determine quickly whether the speaker's message is of real value to you.

Properly practiced, active listening lets you place the responsibility where it belongs: on the speaker. If the speaker has a problem, you can help by being a sounding board, while not taking on unwanted tasks. If you are the one who must take action, active listening helps you ensure that you clearly understand what is expected. Exhibit 18.6 provides a summary of the main points.

Exhibit 18.6 How to Be a Better Listener

Speaker's Responsibility

- Know the listener—tailor the message to the listener's point of view. You don't have to agree with the speaker; just try to understand what he or she is saying.
- Check your volume, speed, and clarity.
- If the listener gives no feedback, ask for it.
- Become a good listener.
- Sharpen your observation skills. Do the words match the nonverbal cues?

Listener's Responsibility

- Know your purpose: What do you want in this situation?
- Be sensitive to the speaker's language, but don't let words you consider ''loaded'' interfere with your job as a listener.
- Check out the emotional content (feeling).
- Listen actively.
- Attend to the speaker.
- Restate or paraphrase the speaker's words.
- Reflect the speaker's emotions.
- Give feedback.
- Concentrate on the message; don't drift.
- Retain the important points.
- Use verbal and nonverbal cues to show that you are listening.
- Don't interrupt or finish the other person's sentences.
- Sharpen your observation skills.
- We listen with our ears, our hearts, our minds, our eyes, and our mouths.

HANDLING RESISTANCE

As Peter Block notes in *Flawless Consulting*, if a client tells you he thinks you're way off the mark, it's not resistance. In the real world, performance improvement consultants may have difficulty identifying the precise type of resistance. They usually just need to know whether the resistance is from lack of knowledge, an inability to adapt, feelings of being threatened, or a belief that proposed solutions lack value. While being alert to a client's resistance, be aware of your own uncertainties. Is the client truly resisting, or is he or she picking up on your own fears?

Once you are convinced the client is resisting you or your information, Block advises that you first try to understand it. At its heart, resistance is an emotional reaction, and a major cause is fear of the unknown. So despite your clear, compelling presentation, the client may be unable to accept any change. An unpleasant, known situation is often preferable to the unknown.

1. Identify the resistance through verbal and nonverbal cues. You may need more than one meeting to confirm your belief that the client is resisting you.
2. Look for other reasons. Is it really resistance or a difference in styles? Respond as though there is no resistance.
3. If the client continues to resist, let the client talk. The problem may be that he or she wants to make sure you understand the problem. Use active listening skills to uncover underlying concerns.
4. Name the resistance in neutral language. Or deal with it through questions.
5. Be quiet and let the client talk.
6. Listen to your own thoughts and feelings.
7. Don't take it personally.
8. Give two good faith responses.
9. If necessary, renegotiate or escalate.

Identify the Type of Resistance

Peter Block has an excellent list of the types of resistance a consultant might encounter. In the early 1990s shortly after reading the first edition of *Flawless Consulting*, the author was so impressed that she tried identifying resistance with nearly every client she met. The truth is, resistance was less common than she assumed. The result was confusion on the part of several clients. The lesson here is that, until you are experienced in the process, stop and think before jumping to conclusions.

Table 18.1 lists the types of resistance that Block identified, when a client may not be resisting, and how to test for it. In some situations, you may need to name the resistance.

Clients often resist because they do not feel understood. Simply saying ''I understand'' may be more of an irritant than a help. You must first indicate who is resisting (the client), how he or she is showing the resistance, and what you need if you are to understand. Naming the resistance involves reflective listening and frequent use of ''you'' statements.

Table 18.1 Identifying Types of Resistance

Type of Resistance	*Might Not Be Resistance If*	*You Can Test by*
Asking for more detail	The client is asking reasonable questions. Some performance improvement practitioners jump to conclusions very quickly, and a request for detail may indicate that the client isn't comfortable with your conclusions or the approach you are taking. The client is more detail-oriented than you are. Conflicts can often occur between ''detail'' and ''big picture'' individuals when the latter doesn't give the former enough detail. (The opposite is also true. Detail types may flood big picture people with information.)	In the first or second meeting, assume the questions are true requests for information. If the client asks for increasingly detailed information, it might signal resistance. Review previous meetings in your own mind. Did the client ask for information or raise concerns you haven't addressed? Ask the client if you are missing something. ''I will look further into this area. Is there something in particular you think I'm missing?'' Evaluate whether there might be a style difference. If you think the client wants more detail than you normally provide, can you negotiate a compromise?
Flooding you with detail	Again consider style differences. You may think the client is flooding you with detail; he or she may believe it is important information to the project.	Try to find out whether the client is always so detail minded. Unless you think the ''flood'' is holding up progress, a little patience may be the best solution. If you think the details are unrelated (or barely related) to the project, say, ''You are giving me a lot of information, I

		am having a little trouble connecting it all to the project we're discussing. Might this mean a bit more in a few days (weeks/months/years)?
No time	The client really has no interest in the project. Trying to push a client into a project is a poor way to get started.	Ask direct questions, such as ''Is this project something we should postpone?'' Busy businesspeople are often unaware of the effect of their behavior. A direct question will open a discussion on the project. If the client's priorities have truly shifted, you may be better off to delay it for a while.
Impracticality	Your ideas may be outside the client's comfort zone.	Ask the client just which ideas he or she views as impractical and validate against your own understanding of the client's work culture. Consider whether your ideas will work and propose potential compromises.
Attacking	The client deals with everyone this way.	In low-risk situations, be assertive—let the client know what you'll accept or not accept. (See the discussion on assertion later in this chapter.)
Silence	While silence does not mean agreement, it is not necessarily a sign of disagreement. The silence might also be a sign of style differences. Some people think through situations in full sentences before responding; others often don't know what they think until they've said it!	Wait for the client to respond, and then evaluate the response. If the client is being unduly passive, consider the situation. Practice patience.
Theorizing	Some people just like to talk and can expand on almost any topic. With some of those who seem to go on forever, the talk may turn to theorizing.	Let the person talk for a bit, but try to bring the conversation back to the situation at hand with a comment such as ''You seem to have an expansive knowledge in

(*Continued*)

Table 18.1 (*Continued*)

Type of Resistance	*Might Not Be Resistance If*	*You Can Test by*
		this field. I know our time is limited, so I am hoping to get in a few questions on some specifics.'' Then follow up with real questions.
Methodology	You are using outdated techniques or you have recommended a methodology or other methods that the client sees as unproven.	First, get specifics and clarify the client's objections. Then ask yourself whether your methodology is outdated or too new for the client's comfort level. Can you achieve valid results with the client's choices? Will learning new technology help you in the future?
Disappearing problem	The problem was misdiagnosed and is no longer an issue. Another possibility is that new problems have arisen than are more important to the client.	Unless you are sure this is a form of resistance or you have a very open relationship with him or her, accept the client's decision. Otherwise, you might appear to be a solution looking for a problem.
Pressing for solutions	The client is genuinely frustrated. Have you missed key delivery dates? Are there unexplained delays? Sometimes the problem is not client impatience but consultant tardiness.	If you know you've met all your deadlines, ask the client if something has changed in his or her area. Perhaps the client has been handed new demands. Ask, ''You seem to want the dates moved forward. What was wrong with our original schedule?''

Handle Resistance Assertively

As with all relationships, small misunderstandings left untended can balloon into big problems. It is difficult, if not impossible, to read others' emotions and thoughts, but observing their behaviors and choosing the right behavior yourself will help you navigate even the most difficult situation. A useful model is one involving assertive, aggressive, non-assertive, and passive-aggressive behaviors.

Assertion is a behavior that lets people express their thoughts, feelings, and values about a situation openly and directly while maintaining regard for the other person. Aggression is when one stands up for his or her own rights but ignores the rights of others. Non-assertive behavior occurs when one ignores his or her own rights and allows others to infringe on them. Passive-aggressive behavior is when one complies with the wishes of others but resents them and acts to subvert the original intent.

Assertive Behavior. Assertive behavior is important in consulting relationships because consultants often do not know the other person's wants and needs. By speaking honestly and tactfully, the consultant stands up for his or her own rights while recognizing those of others. An assertive approach helps build mutual respect, strengthen relationships, and create a perception of decisiveness. The negative side of assertion is that sometimes people will view the person acting assertively as too independent or even pushy.

Aggressive Behavior. If people are not careful, what they perceive in themselves as assertiveness may be seen by others as *aggression*. Monitor your own intentions. Aggressive individuals feel they always have to win and make decisions that do not consider the rights or views of others. You may think you are being persuasive and passionate about your cause; others may think you are single-minded, inflexible, and unwilling to compromise. Some signs of aggressive behavior include interrupting others, using sarcasm, and being condescending by pointing out small mistakes.

Non-Assertive Behavior. *Non-assertive* behavior occurs when an individual does not state his or her own needs, ideas, or feelings. The person may be overly concerned about the other person's rights to the extent that he or she foregoes some or all of his or her own rights to express ideas or influence events. One is acting non-assertively if his or her behavior includes over-accommodation, being too apologetic, and using self-putdowns. Some ways these could be reflected in the consulting relationship are

- *Being overly accommodating*—There is a line between being responsive and overly accommodating. Answering a client's questions quickly is responsive. Turning one's schedule upside down is not.
- *Being too apologetic*—A brief apology for being late to a meeting along with a brief explanation is appropriate. Long explanations just make the person sound defensive.
- *Putting oneself down*—What one person perceives as self-deprecating humor, others may see as lack of confidence.

Because one's actions and words may not agree with his or her real thoughts and feelings, the result is often suppressed anger and resentment. Such emotions are bound to surface eventually, severely damaging the relationship. If left unchecked, non-assertion can lead to passive-aggressive behavior.

Passive-Aggressive Behavior. This sort of behavior occurs when, rather than express negative feelings openly, one expresses them indirectly and often negatively. People who are passive-aggressive are often seen as complying with the requests of others, sometimes even enthusiastically so. Then they may not follow through or deliver only part of what they promised. Consultants often complain about clients not meeting their commitments. The client may be extremely busy, having second thoughts, or simply avoiding dealing with a situation. The consultant's response should be to deal with this behavior as a form of resistance.

Communicate Assertively

To succeed, you need to be able to state your needs in an assertive, constructive way. In fact, there are times and situations in which assertiveness is critical for consultants.

During the contracting phase, one example of assertiveness is to explain the benefits of an approach while being prepared to discuss any limitations. The client needs to own the approach as much as or more than the consultant, and he or she cannot do that with less-than-full information. Be sure to also indicate the importance of both sides meeting their responsibilities. For example, you will want to indicate the seriousness of missed deadlines.

During the consultation, the consultant must take the lead in informing the client about progress and any problems that arise. To communicate assertively, you need to tell the other person how his or her behavior affects you and offer a solution and closing statement. For example, assume your client had promised you access to plant supervisors so you could ask them why they think absenteeism has recently soared. You might say, ''Without access to the plant supervisors, I am concerned that my report will be incomplete. I have some free time on Wednesday and Thursday. Which day do you think would be more convenient?'' The key is to focus on ultimate outcomes rather than on mistakes made by the client or others.

Deal with Others' Behavior

Dealing with assertive behavior is simply a matter of responding to the other person's statements and behaviors. The difficulty arises when the other person appears to be behaving aggressively, non-assertively, or passive-aggressively.

Aggressive Behavior. Aggressive behavior may be a sign of resistance or a deliberate style to put others on the defensive. It is fairly easy to detect. The other person's speech is faster, louder, and more demanding than usual. He or she may give the impression that your rights don't matter. The best response depends on the risks present in the situation. In a low-risk situation, when you are sure the other person is unaware of his or her behavior and its effect on you, an assertive response would be appropriate.

If, however, the other person is obviously angry and you perceive the risk to be moderate to high, become more passive. Become quiet and monitor your own emotions. Breathe through your nose and assess what action would be appropriate. Speak quietly and more slowly.

1. Calmly acknowledge the behavior.
2. Use assertive communication. Describe the effect the behavior is having on you and on the discussion.
3. Determine whether it's possible to continue constructive conversation.
4. Suggest a way to focus on the work issue.
5. Express support and reassurance.

Sometimes aggressiveness takes the form of rudeness. The person has learned that rudeness often results in compliance by others and may even think it speeds processes.

- Watch your language—four-letter words and sarcasm add fuel to the fire.
- Avoid responding in kind—rudeness sparks more rudeness.
- Get in control of yourself—breathe deeply and take a time out, if necessary.
- Manage your anger by stating your own emotions: ''I'm angry because . . . .''
- Deal with the effect of the other person's behavior on you: ''When you play the radio so loudly in your cubicle, I can't talk to customers.''
- Escalate to your boss, the other person's boss, or HR—as a last resort.

Non-Assertive Behavior. Clients can display non-assertive behavior through silence or compliance. They may want to avoid an unpleasant situation, such as being unhappy with your work or having to deliver bad news about difficulties with a project. It may be that the client is having difficulty accepting his or her own responsibility for the problem. When you are dealing with non-assertive behavior:

- *Draw the other person out with gentle questioning*—Most clients have no problem with telling a consultant he or she is on the wrong track, but they find it difficult to acknowledge their own responsibility.

- *Offer praise when appropriate*—You may find it difficult to have to reassure a client that he or she is doing a good job, but this may be the only way to encourage the other person to open up.
- *Watch the client's responses*—Non-assertive behavior is either a sign that the other person feels very uncomfortable or an indicator that he or she has no intention of cooperating with you. Peter Block refers to the latter situation as ''consulting with a stone'' (2000, p. 138). Such situations rarely end happily. You would be wise to remove yourself from the project.

Passive-Aggressive Behavior. Sometimes we think a person is passive-aggressive when there are other explanations for his or her behavior. For example, if someone didn't follow through on a request or we feel he or she betrayed us, our adrenalin surges and we feel an urge to get revenge. Resist the feeling until you can be sure it is passive-aggressiveness. Two situations often viewed as passive-aggressive are when co-workers fail to follow through or when they seem to betray you.

When Someone Doesn't Follow Through

- Did something prevent him from keeping his word? Most people take pride in honoring their commitments. If this is the first time the person has failed to follow through, assume he had some kind of emergency or heavy demands from the boss.
- Did she make the promise too quickly? Sometimes people agree to take on tasks before they realize the scope of the commitment. To avoid such situations, be sure to give the other person the whole picture.
- Is he a guy who just can't say ''no''? Non-assertive people often agree to all tasks because they don't think they have the right to say ''no.'' Confirm that he has time, interest, and commitment for the task you've requested. You may need to negotiate to help the person meet all his obligations.
- Did she really intend not to keep her word? People rarely make commitments when they have no intention of honoring them. If you are convinced that this is the case, you are probably dealing with true passive-aggressive behavior. The following actions will help:
  - *Document everything*—The author was meeting with the manager of a group she was going to lead in a planning session. Her custom at the time was to tape all meetings for note-taking purposes. The manager had some specific goals she wanted the group to achieve and outlined them to the author. When the author included the goals in her preliminary outline, the manager denied ever discussing them. The author apologized and said she'd recheck her tape to see where she'd misunderstood the conversation. The manager immediately dropped the subject.

  - *Avoid meeting alone*—If you have to meet alone, consider recording the meeting. Of course, all parties agree to the recording in advance. If the other person objects or raises questions of trust, explain that you can't always trust your own memory or handwriting. Taping ensures that you don't have to interrupt the other person while you update your notes.
  - *Surface issues in public*—Express your concern as one of understanding. Remind yourself that you are just trying to ''make the invisible visible.'' Be very careful to keep a neutral tone and use assertive, reflective statements as appropriate.
- Keep trusting others—Don't taint other relationships because of one bad one.

When You Feel Betrayed

- Stifle the impulse toward revenge. People who betray you have probably done it to others as well. Act like the rational adult you are, and resist spreading rumors.
- Try to talk to the other person—privately and calmly. Focus on how the person's behavior affected you. ''When you denied saying you would call the client, I felt (betrayed/let down/disappointed/hurt) because my memory of the conversation is so clear.'' Then give specific facts to support your statement. If the other person acknowledges your version, follow up the understanding in an e-mail with copies to others who attended the meeting.
- Try to let go of the grudge—It only lets the other person keep winning. Forget, even if you can't forgive.

MAINTAINING RELATIONSHIPS

The key to maintaining agreements or relationships is follow-up. Performance improvement practitioners often excel at attracting business, but they then skimp on the follow-up with the client. Often, they think the work should speak for itself and that ''client handholding'' is a distraction. In fact, successful practitioners know that maintaining the relationship is just as important as the substantive work.

Key activities in managing agreements include:

- *Stay in touch*—Decide on a method that works well for you and your client. You should agree on the following points:
  - *Frequency of updates*—Revisit the frequency issue from time to time. Clients often want more frequent updates early in a project and then

    require them less often. There is no need to provide updates the client feels are unnecessary.
  - *How the client wants to be updated*—Is an e-mail sufficient or are more formal methods needed?
  - *Reporting in*—Does the client want to know every time you are on the premises? Is prior approval required?
- *Follow through with follow-up*—If you promise a deliverable, deliver it.
- *Be consistent*—Project updates can be tedious and time-consuming. Consider the level of commitment you offer and plan to maintain that level throughout the engagement.
- *Use integrity*—It is possible to recover from most mistakes when consulting on performance improvement projects. On the other hand, a lack of integrity is the one type of situation clients find hard to accept. Maintain your good reputation, even if you must occasionally admit mistakes or take responsibility for less-than-perfect outcomes.
- *Look for win-win interactions*—The best results are those that please both parties.

A few rules help the consultant ensure collaboration with the client:

- *View others as they are*—It's hard enough to change yourself; you can't change others.
- *Recognize differences in personal working style*—Some people have neat desks. Others operate amid stacks. Either can do a good job.
- *Treat others as they want and need to be treated*—This is the platinum rule of successful relationships.
- *Remember that different is not the same as stupid*—Differences can lead to stronger solutions.
- *Consider results, not processes*—Some people know intuitively where they can cut a corner or two, while others are lost if they don't follow a process precisely. Of course, you needn't accept inferior work, but recognize that there is often more than one valid way to complete an assignment.
- *Own your own feelings*—Consider your intent, and recognize when it has changed.
- *Acknowledge your own wants and needs*—Remember, you have a right to succeed.
- *Identify underlying client concerns*—Most fear losing control. Your role is to help them through what is often a difficult time.

Complete the SWOT analysis in Exhibit 18.7 after reviewing the principles presented in this chapter.

Exhibit 18.7 My Personal SWOT Analysis

Strengths: List technical, interpersonal, and consulting skills. Identify other types of value you might add (languages, art, etc.) not commonly found in your field.	*Weaknesses:* Identify areas where you are weak, including technical, interpersonal, and consulting skills.
Opportunities: What opportunities do you seek? What ones are currently open to you?	*Threats:* What technological trends work against you? What is your competition doing?

References

Block, P. (2000). *Flawless consulting* (2nd ed.). San Francisco: Pfeiffer.

Marquis, H. (2008). The three keys to ITIL success. Retrieved July 28, 2008, from itsmwatch.com.

Hall, E. T. (1966). *The hidden dimension*. New York: Anchor Books.

Schein, E. H. (1999). *Process consultation revisited*. Reading, MA: Addison-Wesley.

CHAPTER NINETEEN

Managing ID in the Context of a Training Organization*

Heather C. Maitre
Susan A. Smith

OVERVIEW

An organization's human capital is among its greatest assets. In the new knowledge economy, organizations that lack a strategy for building and maintaining knowledge workers will struggle to compete on a global scale. Workforce demographics play a large part in this because of the impending lack of skilled workers, which will grow to be the greatest that the world has ever experienced—a shortage of ten million skilled workers by 2015 and thirty-five million by 2030 (Employment Policy Foundation, 2002). Retiring Baby Boomers will not be replaced with a generation equal in size, and thus workforce growth will stagnate. Organizations will require rigorous strategies to attract, retain, and build talent organically. The delivery of organizational learning and development plays a key part in the building and maintenance of talent.

Unfortunately, despite the criticality of training in organizations, its contribution to a corporation's bottom line is often not as clearly seen as the

* The authors would like to express a hearty ''thank you'' to Michele A. Graham, who provided incredible editorial support. This chapter would not have been completed without her! We would also like to thank John R. Mattox II, Ph.D., associate director of KPMG's Performance Measurement Group, and his staff, Peter Sanacore and Dr. Ralph E. Grubb, for their research support and intriguing insights.

contributions of other corporate infrastructure groups such as marketing, information technology, or finance. There is a growing body of research on the impact of training, however. Bassi and McMurrer (2002) found that a firm's current training investments are the single most important statistical predictor of total stockholder return. According to the American Society for Training and Development (ASTD) (2000), an increase of $600 in training expenditure per employee yields a 6 percent improvement in shareholder return. Today, most companies realize that their human capital is an asset worthy of investment, and this is resulting in more support from senior management than ever before (Lee, 2008). Dr. Gary Becker, Nobel Laureate from the University of Chicago and expert in human capital investment, stated, ''Any company has to recognize that not only is the human capital of their employees a major asset, it is also a depreciating asset that needs continuing investment'' (Becker, 2001).

Still, many training departments fall into the trap of being ''order takers'' of learning requests from non-learning professionals within the organization. An operation that falls into this trap ends up with an abundance of ''scrap'' learning, or learning that is provided but not applied on the job (KnowledgeAdvisors, 2005). According to KnowledgeAdvisors, creators of Human Capital Analytics software, half of the average training organization's budget is wasted (KnowledgeAdvisors, 2005, p. 3).

The good news is that this trend is shifting. Bersin and Associates reported in *The Corporate Learning Factbook* (2007) that the average direct learning expenditure per employee rose by 7 percent from the prior year, and training department headcount has increased by 6 percent (pp. 19, 23). This investment in learning is a testament to its growing importance in today's corporations.

Training organizations are moving from being reactive to the business to being strategic and forward-thinking. Evidence of this trend includes the learning and development (L&D) function having a ''seat'' at the executive table. The number of chief learning officers (CLOs) has increased from a ''handful'' in the late 1990s to hundreds (Caudron, 2003). In a 2007 CLO magazine survey of 1,438 members of the Business Intelligence Board (BIB), 39 percent were CLOs or other heads of corporate education. CLOs are generally found in organizations that generate revenues from $500 million to $5 billion, with more than 10,000 employees (L'Allier, 2005).

Bersin and Associates (2005), a highly respected training research organization, has stated: ''Despite the large amount of spending in this business area, there is a dearth of best practice books and materials available to help companies understand how to organize well. Because training is a high-touch, company-specific problem, it is difficult to find repeatable best practices'' (pp. 9, 18). This chapter serves to begin bridging this gap and provide a solid foundation for properly structuring and operationalizing a training department.

THE TRAINING ORGANIZATION

Following is a discussion of the key organizational characteristics that provide the environment necessary for producing learning offerings that drive results and a high level of credibility: organizational structure, governance, budgeting, and resources. Other important organizational considerations, such as globalization and external recognition, and how they can be configured to drive success, are also discussed.

Organizational Structure

There isn't a correct answer when wrestling with questions of organizational structure and budget, particularly for large, complex, compliance-driven or global operations. Key questions that an organization must answer include the following:

- Should my organization have a CLO or senior-level person in charge of learning?
- Should the technology and learning delivery infrastructure be funded and managed centrally or left to business units?
- Should core training programs such as on-boarding, leadership, and management be standardized across the organization or left to other groups within the organization?
- What is the best way to fund and manage business-specific training programs (for example, technical, sales) such that they meet business needs?

This section presents the key considerations, and a model to help aid in decision making based on organizational size and importance of training.

Executive Positioning. Organizational learning and development departments, like any other department, require strong leadership to have the greatest level of success. Similar to a corporation's finance and information technology functions, learning should have a seat at the executive table. This will ensure the learning organization is included in conversations about where the business will be taken so that learning and development reflect corporate strategy, engender credibility across the organization, and receive support at the highest levels. This structure starts with the presence of a senior-level executive in charge of learning (a vice president of learning or chief learning officer) who reports directly to the company's president or CEO. High-impact learning organizations with a CLO ''have much higher impact, efficiency, sharing of best practices, e-learning strategies and measurement strategies'' (Bersin & Associates, 2008, p. 122). Effective CLOs are often business-focused and have the ability to speak to, understand, and support the businesses effectively.

Centralized Versus Decentralized. Determining whether the organization should be centralized or decentralized is one of the most difficult decisions to make (Chief Learning Officer Business Intelligence Industry Report, 2008).

In a centralized learning organization, learning services for the enterprise are provided by a core learning organization and funded centrally. Centralized training organizations often provide learning infrastructure services such as learning delivery operations, learning systems technologies (such as a learning management system [LMS] and as learning content management system [LCMS]), instructional design, e-learning design and development, measurement and standards, and enterprise-wide training content such as management and leadership. Success with this type of configuration requires adequate funding and ''tentacles'' into the business; otherwise, business units tend to create small training groups in hidden pockets of the organization where training needs are not being met. Centralization is most impactful when it successfully implements technology, design, and process standards for learning across the organization.

In a decentralized model, training is usually funded by the line of business. If training is decentralized and the organization is large, there is often still a central training organization, but it is much smaller and provides minimal core services (for example, basic LMS technology). Decentralized models are often viewed as necessary in larger organizations, as many think it would be impossible to meet the needs of a large, complex organization centrally. Whether centralized or decentralized, very specialized business skills (such as technical topics or product training) are often best addressed by the business unit in any organizational scenario, with a high degree of collaboration with a central learning and development group. The key is the extent to which local training resources are connected to a central group that standardizes processes and nomenclature and ties learning together for the broader enterprise.

The magic is in determining the right balance for your organization, particularly if it is large. We are partial to a centralized model and believe it is more strategic, provided that it has very strong leadership and linkage to the top of the organization. An esteemed colleague once said, ''Unless you want fifty sessions of *The 7 Habits of Highly Effective People*, then the training organization must be centralized.'' The rationale is that, if a learning and development organization is centralized, funding and directional decisions will be more transparent, strategic, properly governed, and thus command a higher level of executive scrutiny. It will be less susceptible to the whims of any particular line of business that may have its direct interests at heart, versus those of the organization as a whole.

The arguments for centralizing or decentralizing the training organization go hand in hand with the arguments for centralizing or decentralizing the learning and development budget. We believe that a centrally funded L&D organization, controlled by the office of the president, chairman, or chief executive officer of the organization, is the most strategic and effective budgeting model. A strategically located, central entity will fund, support, and provide resources for initiatives that will directly relate to metrics that are important to business sustainability (growth, revenue, turnover, sales). It is likely to result in an organization that sees training as a strategic business investment.

Many learning organizations struggle most with technology, providing innovative learning solutions and measuring business impact. These key areas represent an even greater struggle within a decentralized training organization. The reason is that measurement and technology require centers of excellence supported by research and development and specialized expertise; funding such entities over several departments within the organization is expensive and ineffective.

Technology, in particular, often drives centralization because an LMS implementation is expensive and resource-intensive. It causes an organization to ask how to define its learning offerings in a way that makes sense in the context of the business; how these offerings should be targeted by position, level, and line of business; and how its critical work functions should be addressed through competency models. More importantly, it serves as a central source of data from which the learning organization can actually manage its business. When executed properly, the learning organization is powered by an incredibly powerful LMS machine: after all, if it can't be tracked and measured consistently, it can't be managed.

Table 19.1 provides a summary of the pros and cons of centralized and decentralized organizations. Note that the larger the organization, the more the negative issues are magnified.

Budgeting

It is not uncommon for an organization to have the following basic budgeting problem: How much should be spent in total on learning? A newly appointed CLO or VP of learning, particularly in a large, decentralized organization, can spend years trying to answer this question. Doing so is critical, however, for control and impact of organizational learning and development. Strategies for determining total spending can be employed through corporate expense reporting systems, purchasing systems, and the annual budgeting processes.

The structure, level, and rationale for budgeting the learning organization vary depending on industry, type of business, employee headcount, geographic reach, organizational structure, politics, and value placed on learning within the organization.

Table 19.1 Pros and Cons of Centralized and Decentralized Organizations

Centralized		*Decentralized*	
Pro	*Con*	*Pro*	*Con*
Transparency in L&D investments—governance easier	Line of business needs may be stifled without appropriate L&D ''tentacles'' into the business	Better supports and understands line of business needs	Redundancy in learning offerings
Total L&D spending is known	Small training groups and initiatives may emerge in hidden pockets of the organization	Line of business control/satisfaction	Missed opportunity for economies of scale/discounts with vendors used by multiple departments
L&D investments are more strategic (when controlled by the top of the organization)	Central L&D can be seen as unresponsive		Often results in nonexistent or multiple LMS systems, ''makeshift'' LMS systems, or multiple websites that contain learning information and offerings; this makes it difficult to brand learning, present learning cohesively, or represent learning holistically across all professional development offerings.

(*Continued*)

Table 19.1 (*Continued*)

Centralized		*Decentralized*	
Pro	*Con*	*Pro*	*Con*
Economies of scale with learning management systems (LMS), learning delivery, and content quality (given solid core standards/process)			Lack of consistency in learning delivery and content across the organization
Robust core resources available for the rest of the organization to leverage			Lack of robust resources for learning platform technology, instructional analysis and design, etc.
Core programs can be delivered consistently and cost-effectively across the organization			Higher degree of administrative overhead for coordinating decentralized units
Enterprise-wide compliance with regulatory standards is easier			Measurement is not possible if learning data is inconsistently represented (learning offerings) and not centralized

Determining and advocating for the appropriate level of funding may seem like a daunting feat. Fortunately, due to an abundance of benchmarking, general funding guidelines can be determined by sources such as the ASTD Benchmarking Forum (BMF) and Workplace Learning and Performance (WLP) Scorecard®, and Bersin's What Works® Series. They include L&D headcount level based on the number of employees the learning organization must serve and L&D spend as a percent of the organization's revenue, profit, and payroll. For example, Bersin and Associates (2008) recommends 1.5 to 2 percent of payroll for base investment, with additional investment added for strategic programs (since their research has shown that 80 percent of the value of L&D organizations comes from 20 percent of their programs).

Another budgeting decision that must be made is whether to employ a chargeback system for training, wherein departments pay the central training organization for training consumed by their employees. It may or may not be connected to centralized funding—that is, a centrally funded L&D organization may also employ chargebacks in some or all cases. A chargeback system ensures that business units only pay for what they use, controls ''no shows'' or taking training for granted (also culture dependent), and makes overall training costs and program allocations by business line more transparent. The learning and development organization is considered a ''vendor,'' which is positive when business units are inclined to prefer the internal organization's offerings, and it keeps L&D ''honest'' to deliver a good product for the value.

A chargeback system can be negative when external programs are chosen over internal ones. It can also seem illogical to the recipient organization when training on specific content is mandatory for all or large portions of the organization.

THE LEARNING ORGANIZATION DECISION MODEL

Figure 19.1 provides recommendations for four types of learning organizations: Standardized and Lean (for large organizations in which training is not strongly considered strategic/required for the business); Collaborative and Strategic (large organizations that consider training to be business-critical); Nimble and Strategic (small organizations that consider training to be business-critical); and Simple and Effective (for very small organizations in which training is not considered strategic/required). The quadrants can serve as a best practices guide in organizational decision making. Note: We would argue that training is always strategic. However, this is not reality for many organizations; hence, the recommendations below are based on acceptance of that reality. We also believe that ''Standardized and Lean'' training organizations can move to ''Collaborative and Strategic'' if core services are

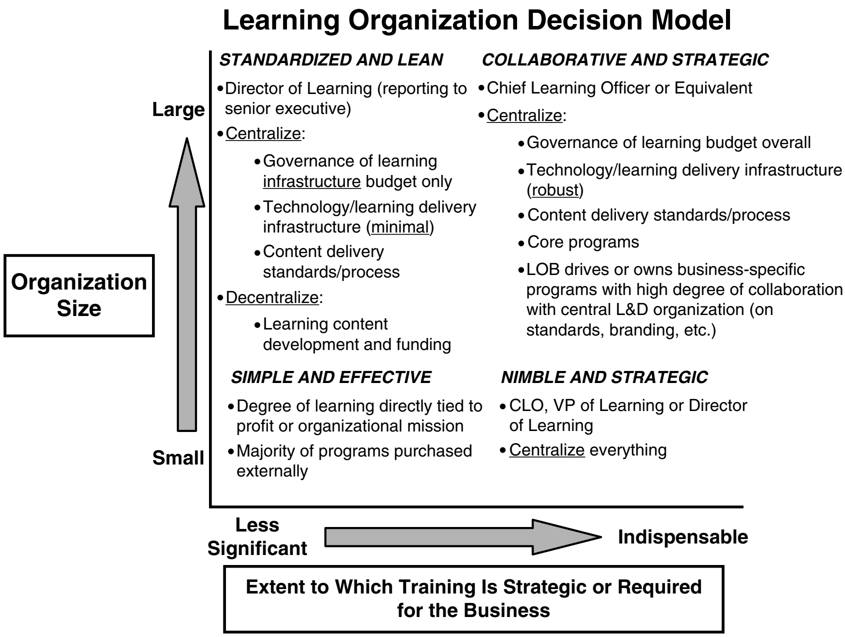


Figure 19.1 Learning Organization Decision Model.

provided consistently and effectively. The power of having this foundation paves the way for organizational maturity with learning.

Governance

There is usually no shortage of training requests in any given organization. There are, however, almost always limited learning and development resources. The ultimate balance a learning and development organization should strike is one in which the training produced has marked business impact and the process used is as cost-effective and efficient as possible. Governance is a critical component for striking this balance. Without organizational governance of the learning function, stray requests will come in from every pocket of the organization, and they will generally not be informed by what is best for the organization overall. In addition, these types of "stray" requests usually come from non-learning and development personnel. This opens the door for requests for solutions to problems that are not necessarily learning issues. Performance is not just a function of learning—it is a function of many things, such as environment, managerial support, and motivation, to name a few. To avoid an abundance of "scrap learning" (a term coined by Knowledge Advisors) in an organization, learning requests must be validated to ensure that the learning solution will meet stated goals, and that it is worthy of the investment requested.

Effective governance requires the following:

- A steering committee with the appropriate level and type of representation (membership should include leadership representation from each business unit, HR, and functional departmental heads such as marketing, finance, and IT);
- Requests that are accompanied by a business case, criteria for program success;
- An effective prioritization process; and
- An expectation that there will be follow-up to ensure that programs meet their stated objectives.

Approving every single training request by committee vote is clearly impractical. There should be a threshold for requests that must be reviewed before L&D resources are allocated. The threshold should be in proportion to L&D resourcing overall.

Overall Structure

Learning and development organizations usually have the following areas of staffing or departments: planning and strategy; business management; analysis, design, and development; measurement and analytics; technology; and delivery. The organizational model in Figure 19.2 shows each of these areas and how they might be organized.

Figure 19.2 Model of a Learning Organization.

MANAGING TRAINING DELIVERY

High performing learning organizations, whether centralized or decentralized, are learner focused. They understand that the learner experience is all-encompassing, starting with the communication of a learning event, delivery of the course, timely posting of learning credits, and ending with evaluation and post-course follow-up. The learner experience should also exceed expectations regardless of the delivery method—instructor-led or e-learning.

Training Delivery Activities

The objective of the following sections is to highlight the key activities associated with the delivery of training. This may not be an all-inclusive list; however, it frames the key areas that one must consider for the delivery of live training. Delivery activities are discussed in the context of whether they must be done before, during, or after learning events. An overarching theme is how the skillful execution of key tasks can enhance the learning experience, thus increasing overall learner satisfaction.

In this section, three basic methods of delivery are discussed: instructor-led classroom training (classroom), instructor-led virtual training (virtual), and web-based training (WBT).

Pre-Delivery Phase Activities. For a training program to be successful, both from a learner and business perspective, several decisions, processes, and procedures must be in place, such as the following:

Method and timing of delivery. Prior to the development and delivery of the training course, the team responsible for the design and development must determine the method of delivery. As mentioned, this can include classroom, virtual, or WBT. Effective delivery of training requires an organized, detailed approach to the activities that must take place before the learning offering begins.

The first delivery decisions the organization must address are the date(s) and time(s) for virtual and/or classroom events. If the event is delivered virtually, the team will need to determine who will facilitate the session(s) and where the sessions will be held. Such options include central ''broadcast'' rooms, local offices, or a training facility. If the training is to be delivered in a classroom, the team must secure the space/venue to conduct the event. Venue considerations are discussed later in this chapter.

With the learning program delivery method and timing determined, focus can be placed on the key activities that need to occur before the actual delivery of the training.

Communications. The compilation, authoring, and dissemination of program information to learners are the key activities for program communications. The information that needs to be communicated during this phase includes any prerequisites and requirements that must be met in order to attend/consume the training. Logistical details also need to be provided to the learners and the instructors attending classroom training.

Communication of training offerings can be delivered in multiple ways, including standardized e-mails, pre-established print vehicles, and personalized/targeted messages on the learners' pages of the LMS. These communications should also include clear, simple instructions on how to register for the course. Next come logistical details related to the offering—location, lodging, start and end times, closest airport, and ground transportation, to name a few. This information can be sent to those registered by e-mail, or posted on a website or in the LMS. See Exhibit 19.1 for a sample of a consolidated classroom course logistics document.

If your training organization conducts a large volume of instructor-led classroom programs that require learner travel, we highly recommend providing an emergency phone number that learners can be instructed to call in the event of inclement weather or an emergency situation in the host city. A recorded message on this number can be used to communicate to the learners whether the course is canceled, starting late, or running as scheduled, along with any other important details.

Course registration. The registration activities include selecting the method for registration, setting up the registration process, and managing registration once begun. Registration options depend on what technologies are available, but usually include event registration systems, home-grown registration websites, or an LMS. Alternatively, registrations can manually be handled with e-mail responses logged into spreadsheets. If you are conducting training that will require hotel and travel for the learner, then a hotel and travel reservation process would also need to be addressed. Setting up the registration process also involves creating the offerings in an LMS or website or crafting an e-mail with voting buttons. The process should be easy for the learners to understand and execute. Registration processes, especially those found in an LMS or on a website, that involve repetitive, unnecessary clicks or steps can frustrate learners to the point of tainting their overall opinion about the learning organization (and its supporting technologies). Unfortunately, this issue is notoriously true for organizations that use an LMS.

Once registration is opened, their management becomes important. This includes monitoring enrollments against capacity, handling waitlists, and taking cancellations. A way to positively affect the learner experience is to provide participants with a way to ask questions or have concerns addressed about the registration process, the content of the course, or other topics. A service/call center or specific e-mail mailbox manned by a course coordinator/manager are two of many ways to meet this need. It has been our experience that monitoring

Exhibit 19.1 Sample Logistics Document for Instructor-Led Course

[Enter Course Name Here]

The following is important information to know about your class/program before you go:

Meeting Number	**[Enter Meeting Number]**		
Training Location	**[Replace below with course information]** Hotel Name Street Address City, State Zip Code **Phone:** **Fax:**	**Sleeping Accommodations**	**[Replace below with course information]** Hotel Name Street Address City, State Zip Code **Phone:** **Fax:**
Course Start Date and Time	[______] at [______]	**Course End Date and Time**	[______] at [______]
Arrival Date	**[Enter Arrival Date]**	**Departure Date**	[Enter Departure Date] Please do not book your flights before **[Insert three hours after training]**
Learning Credit Guidance	Learning credit will be awarded for this course. All actions that are required by the learner to earn credits should be entered here.		
Advanced Prep	[______]		
Prerequisite	[______]		

Items Required for Training	You **must** bring your laptop, power cord, computer lock, SecurID, issued headset, and a form of picture ID***.
	***Important: For security purposes, you must present a photo ID, such as a driver's license or passport, in order to receive your name badge for the program.
Attire	Business Casual
Recommended Airport	**Airport Code**
Ground Transportation	**[If participants are their own, please provide direction on where to pick up transportation and approximate transportation cost from airport to hotel]**
Hotel/Travel Reservations (U.S. Participants only)	Instructions on making both hotel and travel arrangements. Also include a contact number if there are issues. Insert Meeting Registration link or instructions and phone number.
	If you need to modify your hotel reservation, **[insert instructions and link].**
Expenses	What the learner will be expected to pay for is entered here. Also included is how they will be billed by the hotel and any information on payment (i.e., company credit card). Expense policy may also be reiterated.
Accessibility Needs	If you have a special physical or communication need that may impact your participation in this training, please send an e-mail to **[insert address]** at least two (2) weeks prior to the start date of the program.
Emergency Hotline	In the event of inclement weather or other incidents which may cause the course to be postponed or cancelled, contact the Emergency Hotline at **[enter phone number]** for up-to-date information and other messages.
Policies	Cancellation Penalty Policy; Rooming Policies, etc.

these services is critical. Customer service, as measured by the learner, must be pleasant, timely, and accurate.

If hotel accommodations are necessary, it is important to reconcile the course registration list with the hotel reservation list. Sending out reminders to learners to make sure they register for the course and reserve a hotel room avoids many unpleasant conversations onsite.

Materials. Course materials must be created ahead of time, duplicated, and shipped; supplies and materials must be ordered; and shipments must be confirmed as received at the venue.

Instructors. Selecting the instructors, conducting train-the-trainer classes and rehearsals, registering the instructors for the program, and ensuring the instructors have transportation and hotel reservations are also responsibilities of the delivery team.

Technical setup. For virtual and web-based training, the course and materials must be uploaded into the virtual training system, the LMS, and/or websites ahead of time.

Delivery Phase Activities. As the training delivery date nears, the learning delivery team shifts its focus from the pre-delivery activities to delivery activities.

Classroom programs. The following activities are associated with classroom delivery:

> *Training materials management and classroom(s) setup*—Onsite, the staff must complete the sometimes tedious tasks related to material management and classroom setup, including unpacking and sorting materials and supplies associated with the course. This ensures that all materials and supplies arrived onsite and that there are enough for all who registered. Next is setting up the classrooms with the instructor and learner materials, rosters, and any other supplies required. Do not underestimate the time it takes to perform these activities. Setting up a single classroom for forty learners can take twenty to thirty minutes.
>
> *Instructor care*—Instructor care may be the responsibility of the manager of the course or the delivery team. Making sure the instructors have administrative and course content support is critical, including checking in with the instructors periodically during the day, conducting daily debriefing sessions, and ensuring the learners are arriving in classrooms.
>
> *Technology setup and administration*—Make sure the computer projectors are all functioning and that the proper software is on instructors' laptops. Hand out audience response system keypads to the learners. The learner experience can be dramatically impacted (either positively or negatively) by the use of technology. The delivery team should have been well trained prior

to arriving onsite. We strongly recommend sending IT support to the site for larger programs or those programs using new or complex technologies. You should always have a contingency plan in the event that technologies malfunction.

Onsite registration—Onsite registration can be hectic. Checking in the learners may sound simple, but when there are hundreds of attendees and the registrars are performing multiple transactions with each learner—checking them in, handing them their name badges, and directing them to their classroom—it can become very chaotic. For a smooth registration process and a pleasant learner experience, try the following:

- Arrange with the hotel to hand your welcome letters to learners when they check into the hotel. The letter should, as a minimum, welcome them to the area, provide the location and times of course registration and where to contact the learning organization onsite.
- Organize your team and augment your team with temporary staff used by the facility for their other large events so that you will have a smooth-running registration process. These people will know the basics of a registration process and be familiar with the venue.
- Craft the process to avoid long waiting lines. Have multiple registrars, each handling a different part of the alphabet, and limit the number of transactions for each learner. Have a separate desk for specialty requests (for example, servicing walk-ins who did not pre-register, switching of elective classes). If you need to check registrants' identification, place placards near the registration location reminding folks to have identification out as they approach the counter.
- Include ''nice-to-haves'' at the registration desk, which might include maps of the facility, personalized agendas, local restaurant guides/maps, or other helpful materials. Provide emergency information cards that contain the numbers of local hospitals/walk-in clinics and onsite company staff contact information (especially if the learners are more junior/younger).

Evaluation and test administration—The distribution, administration, and collection of course evaluations are essential to a successful program. If you have paper evaluations and/or tests, the distribution can be part of the materials setup in each of the classrooms. The administration of the evaluation or the test is usually done by the instructor. The delivery team's responsibility may be to remind the instructors when and how to administer these items. Paper evaluations and tests should be collected as soon as they are completed. If you are using electronic evaluations/tests, the distribution may involve sending an e-mail with a link to the instrument to the attendees. This would be done by the delivery team.

Attendance tracking—These activities include providing the instructor with a roster and instructions on how attendance should be taken or tracked and collection of the roster. Attendance can be tracked in many different ways, which will be discussed later.

Facility liaison—Someone on the team should act as a liaison with your hotel/conference center meeting planner shortly before the course begins to review the details such as food, classroom breakdowns and resets, or technology being used. Other activities that ensure a smooth course delivery are to check that meals are ready when scheduled, that the temperature in the classrooms is satisfactory, and that hotel/conference staff is prepared for any scheduled special events. All extra-curricular activities such as entertainment or sporting events should be well-executed with timely and comfortable transportation, efficient ticket distribution, and quality food service.

Safety and care for all onsite staff/instructors/learners—The safety of their professionals is a paramount concern for all corporations. To enable onsite support staff to respond effectively to any and all situations, we recommend training them in basic emergency procedures. A template for an emergency guidelines document can be found in Exhibit 19.2. Having an HR manager onsite is a best practice, particularly for highly regulated and compliance-driven organizations or for high-risk programs (such as intern programs, which typically have young participants). Onsite HR could handle a personal emergency or professional issue such as a code of conduct infraction.

Virtual programs. The following are the standard activities managed by the learning organization during virtual programs:

Session support/moderating—These activities include arranging for the session to be opened prior to the start time of the session; staying at the session to provide instructor and learner support; and/or moderating/facilitating the session and closing out the session in the virtual classroom system. During a virtual training, the support person or moderator may also need to monitor the text chat feature in the system that allows the learners to ask questions or bring up an issue they may be experiencing.

Troubleshooting technical issues—Many organizations elect to have their IT department perform troubleshooting. Others depend on the training delivery staff to provide this service. This requires knowledge and experience with the technologies, which very often include the virtual classroom system, the LMS, and the associated processes with running a session.

The only activity for the learning organization to perform once a WBT program has been made available is monitoring the number of people accessing and completing the training. During the delivery of virtual and web-based training, the engagement of the learner in the content is paramount. The

Exhibit 19.2 Onsite Emergency Guidelines Template

Date [this document should be reviewed and revised on a routine basis]

Purpose: This document is intended as a guide for action to be taken in the event of an onsite emergency during a learning event. It is intended as a guide only. While there is no way to predict all the types of emergency situations and no way of prescribing all actions to be taken in all emergencies, this document provides procedures for the more common situations.

Types of Emergencies

- Weather, including hurricanes, tornados, major snow storms, floods
- Health, including illness, accidents, death, food poisoning
- Fire/explosions
- Acts of terrorism
- Other, including participant family emergencies, theft, power failure

Basic Emergency Preparedness

- Remain calm.
- The paramount concern is the safety and well-being of all professionals.
- Know your facility's emergency contact number (or dial ''0'' from the nearest house phone for hotel/facility security) and know the emergency plan.
- Know the guidelines in this document and have them with you at all times while on site.
- Use common sense in your interpretation of these guidelines.

Topics to Consider for These Guidelines

- Chain of command for onsite staff of your company
- Knowing the facility's emergency contact and emergency plan
- Communication escalation procedures
    - Company security
    - Chief learning officer/head of learning and development
    - Head of human resources
    - Other HR contacts
- Public relations for emergency situations
- Accidents/illness: urgent, not urgent, hospitalization required
- Armed robbery
- Participant/staff death

(*Continued*)

Exhibit 19.2 (*Continued*)

- Fire/explosions
- Participant's family emergency
- Power failure
- Terrorist acts
- Theft
- Weather emergencies affecting program arrivals/departures, occurring during the program
- Emergency hotline number that has been set up to inform participants of training schedule changes in the event of emergency situations (weather, travel, etc.). This number will be included in every confirmation letter sent to every registrant of all company-sponsored programs. This number is to be used to inform the participants of a schedule change prior to the start of the program.

''surround'' of the training has fewer interactions with the learner, so it is critical to have the technology operating properly. To ensure a positive learner experience: provide easy access to the content, minimize any technical issues, and ensure access to technical assistance if needed.

Post-Delivery Activities. The following activities are typically conducted after the classroom or virtual training has been conducted:

Recording completions. Once training is complete, the learning organization must record completions, which involves the reconciliation of attendance rosters and the recording/input of completion status and training credits into the tracking mechanism your organization is using. This should be performed as soon after the completion of the training as possible. Learners will appreciate having their records updated in a timely and accurate manner, especially if they are in a highly regulated profession. Self-service printing of all documentation required for certifications, professional licenses, and regulatory boards is an added service you can provide your professionals.

Document retention and materials storage. The retention of key training documents such as attendance rosters, instructor qualifications, and instructional materials is recommended in order to provide a history of the course and assist in responding to inquiries about the course. It also provides documents to use should the course be conducted again. Document retention is usually a requirement for compliance with professional regulatory standards. If you are running multiple sessions of a course during your training year, we strongly recommend that the retention of documents/materials be part of your normal session close-out process. This will minimize redundant work and the risk of losing/misplacing documents.

Measurement instruments (evaluations and tests) processing and reporting. The ultimate goals are to provide measurement reports for review and to provide the delivery staff with test scores that may be required. These activities might include scanning/collating and summarizing the evaluation/test data for paper-based instruments, closing down the evaluation/test in the technology application, and running the reports for the assessments used in each class/session.

Summary of Delivery Activities. It is critical to note that all of the above activities need to be executed in a timely, accurate, consistent, and efficient manner. We recommend the creation and use of standard operating procedures (Exhibit 19.3), service-level agreements (Exhibit 19.4), and process flow diagrams (Exhibit 19.5) to achieve quality delivery of training. These documents are excellent process training tools and reference materials, and they are helpful during service negotiations with internal clients. See Exhibits 19.3, 19.4, and 19.5 for templates and/or examples of these tools.

The remainder of this section discusses, in greater detail, the major delivery activities: communications, including training policies, venues, and instructors; materials management; tracking and reporting; and conducting these activities in compliance-driven environments.

TRAINING DELIVERY COMMUNICATIONS

An integral part of all learning organizations is the communication function. While this function is not limited to the delivery of training, it is frequently associated with the training infrastructure. For that reason, we have chosen to include some basic communications tips in this chapter.

Four questions should be answered before crafting a communication. These are related to content, audience, vehicle, and the author.

What needs to be communicated? Content. Include information on any new technologies that enable training, such as an LMS, virtual classroom systems, and audience response systems. If any upgrades to these systems are available or training is offered, include that information in any communication. Also include any new training processes or policies such as a new process map or reference to specific organizational policies (dress code, use of organization equipment, general conduct, or others). If your organization is subject to any regulatory/compliance-related requirements/policies, then this would also be included in the messaging. The content might also be any additional training offerings and/or special learning and development events the learner should be aware of.

Exhibit 19.3 Standard Operating Procedure Template

OBJECTIVE:	The objective of this SOP is to define the process for . . . . This SOP includes		
SCOPE:	Name of Departments Involved with This Process		
REFERENCE DOCUMENTS:			
PROCEDURE:			
<u>**Responsibility**</u>		<u>**Action**</u>	
	1.		
	2.		
	3.		
PROCEDURE:			
<u>**Responsibility**</u>		<u>**Action**</u>	
	4.	*How to (for specific steps):* 1. 2.	

APPROVALS

Prepared by:		
	Author	Date
Approved by:		
	Approver 1	Date
Approved by:		
	Approver 2	Date

Attachments

Exhibit 19.4 Example of a Service-Level Agreement

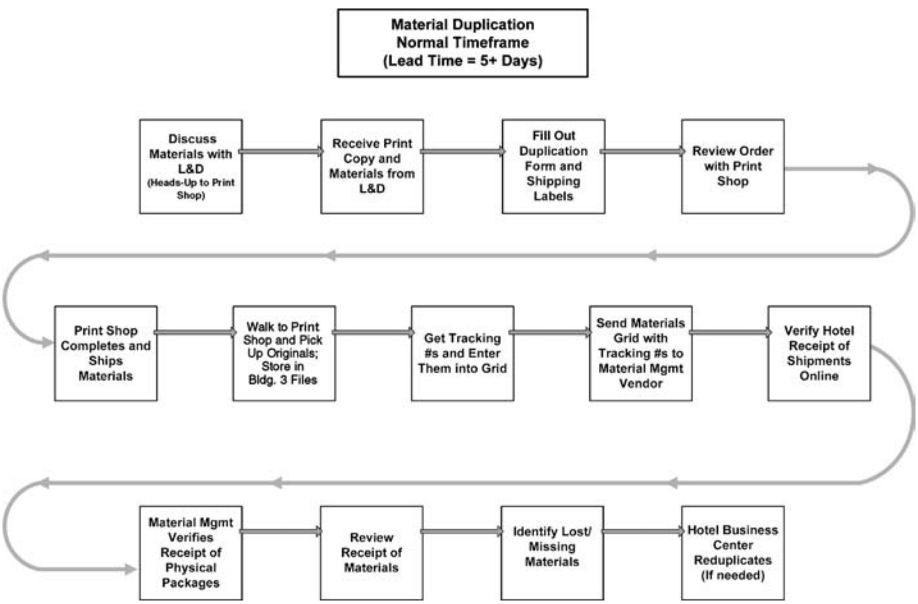
DUPLICATION OF PAPER MATERIALS

Purpose: To ensure the accurate and timely printing and shipping of training materials.
L&D responsibilities: To provide delivery team (DT) with ''camera-ready/print condition'' originals six or more days out and with electronic materials three to five days out.

If materials are received by DT in ''print condition''	**The duplication will be handled by**	**Timely onsite delivery?**
Ten or more work days out	DT and Internal Print Shop	Yes
Six to ten work days out	DT and Internal Print Shop (with Vendor A, who provides twenty-four-hour service as back-up)	Yes
If materials are received by DT electronically with a ''read me'' file of print format instructions	**The duplication will be handled by**	**Timely onsite delivery?**
Three to five work days out	DT and Vendor A	Yes
Two or fewer work days out	DT/L&D, Vendor A and, if required, the Hotel Business Center	*Yes (this will depend on the volume and capacity of the resources)

Note: Camera-Ready/Print condition means printed out and arranged in the same format that will be requested of the print shop or print vendor.
*Have Meeting Manager alert the hotel as early as possible if it becomes necessary to use this option.

Exhibit 19.5 Example of a Process Flow



Who needs to be informed of this content? The Audience. Once your organization determines the content of the communication, it is important to consider who the recipients will be. These may include the entire company, leadership only, L&D, a specific department within L&D, and/or external audiences such as those for achievements and award applications.

Who is communicating this information? The Appropriate Author. Determining who will communicate the information is important and directly relates to the content. The criticality of the message will generally indicate whether the author should be a specific person in the L&D department, the L&D organization/leadership, or company leadership.

How will the content be communicated? Vehicle. In today's technological world, learning organizations have a variety of vehicles to deliver communication. Organizations can choose the more traditional vehicles, such as desk drop or internal hard-copy distribution, or move to electronic vehicles, such as e-newsletters, LMS/website, or e-mail distribution.

We would like to share some lessons learned related to electronic communication vehicles. Specifically, when establishing a learning and development e-newsletter or a training-related e-mail protocol, conduct an official launch (or re-launch) of the vehicle to brand the newsletter or e-mail as an official communication from training. Make the first edition compelling by incorporating leadership endorsements, industry information, and/or targeted training announcements. If your organization has a virtual learning system, you may want to launch your newsletter by conducting a session(s) devoted to discussing the purposes and features of the newsletter. Establish a routine publication schedule to get learners in the habit of looking for it on certain dates. The format of the e-mail or newsletter should be consistent each time it is posted or sent. Don't forget to make it easy to read and navigate.

Communications that lend themselves to hard-copy format are those in which the printed content contains detailed instructions that assist the learner in performing a task or information that the organization would like learners to refer to on a frequent basis. Examples might be quick reference cards on how to log into and navigate a new LMS, a document stating a new policy, or information about mandatory training that must be attended over the next few months with a calendar of the sessions. If your resources allow, make these communications eye-catching and distribute them to learners by placing them at their desks or in their work areas. Keep in mind that having this information available electronically will be required for mobile work forces.

If your organization has an LMS, use all functionality related to the personalization of the system. For example, communications regarding training offerings can be placed on the landing page of the LMS. These pages are capable of displaying communications tailored to the individual learner. If you choose the messages presented on this page wisely, your LMS can be a powerful communication vehicle.

TRAINING POLICIES

Training policies are an inevitable part of providing learning and development activities in any company. We recommend that L&D functions consider establishing and institutionalizing the policies listed below.

State who in the company is entitled to what training. If you are a part of a training department, then company leadership already has a vision and a strategy for training that includes general content that training should cover and who in the enterprise should receive this training. If your company is in the process of establishing a training function, you will have to make these decisions in order to formulate a policy. The amount of training an employee is mandated to attend each year should be stated. Senior leadership should be consulted when developing this policy and for approval/endorsement of the final version.

There should also be policies related to the financial aspects of training. Depending on your L&D budget, you may need to establish policies regarding tuition; learner lodging, meal and travel charges for attendance at training events; and/or training cancellation penalty policies, among others. Cancellation penalties may be used to offset the administrative costs.

Human resources training policies should also be crafted and communicated. We strongly recommend that you have policies covering rooming while attending training; a code of conduct for both learners and instructors outlining expected behavior and actions; and a policy on who should be consulted for ad-hoc HR-related matters. Some examples of ad-hoc matters include support for disabled attendees, any concessions made for nursing mothers, and waivers for recommended training given special circumstances presented by the learner. Trying to provide policies that cover each of these ad-hoc circumstances is not practical. Providing the person or group who can assist in addressing these matters in an effective manner may be a more effective ''policy.''

Here are some tips that may help you with learning policies:

- If the policy pertains to human resource, legal, or financial issues, consult your in-house subject-matter professionals when formulating the policy to ensure that the policy is aligned with other internal policies, practices, and procedures.
- If you are affected by regulatory board standards, consider communicating the standards related to training in the form of a company policy or as part of your policies. This provides your learners with one location for all policies.
- The consequences of not complying with a training policy should be determined before issuing the policy and included as part of the communication.
- To ensure that all professionals are aware of the major training policies, communicate them during company orientation programs and house them in easily accessible websites/repositories.

Two cautionary pieces of advice: (1) if you have a policy, you need to have processes for monitoring adherence to the policy and for ensuring the consequences of non-compliance are enforced and (2) there will always be unusual situations that fall outside of the established policies. Apply prudence, consistency, and common sense when addressing these occurrences.

Examples of training-related policies can be found in Exhibits 19.6 and 19.7.

VENUES

The training environment must be conducive to learning while meeting the many needs of the administrators, participants, and instructors. This section is devoted to training venues.

Exhibit 19.6 Business Casual Attire Policy

The intent of this policy is to help maintain a certain level of professional business attire standards among our employees. In support of that policy, the Learning and Development department would like to supplement this policy for our professionals representing the firm at all learning and development-related events.

How We Identify Ourselves
New this year, we are providing you with a symbol that quickly identifies you as a learning and development team member. Going forward, everyone who is working onsite for a learning and development event will have a yellow banner on his or her name badge.

When Working with Clients

Working with clients means direct contact with them, for example, at the learning event, or when you're attending evening events, for example, dinner or company-sponsored events. Please use the current dress code policy as a guide, including:

Slacks: Dress slacks, pantsuits, khakis, and corduroys are acceptable alternatives to traditional business slacks. Inappropriate choices include jeans of any fabric or color, sweatpants, bib overalls, cargo pants, Capri, leather, spandex, or other form-fitting pants and/or shorts.

Shirts: Casual shirts with collars, polo shirts, blouses, tunics, golf shirts, sweaters, sport jackets, and turtlenecks are within guidelines. Those not within guidelines include: midriff or spandex tops, t-shirts, sweatshirts, tank tops, halter tops, tops with bare shoulders (unless worn under a jacket), or denim shirts.

Footwear: Loafers, leather pumps, flats, deck shoes, and dress sandals are acceptable footwear. Work or hiking boots, athletic or tennis shoes, leather or canvas sneakers, sport sandals, thongs, and slippers are not acceptable choices.

Dresses and Skirts: Casual dresses and skirts, and moderate-length split skirts or culottes are acceptable. Minis, spaghetti straps, open back, low-cut, or revealing dresses are not acceptable.

During Free Time
Free time is considered when you are onsite but are not expected to work with clients or vendors. A more relaxed appearance is acceptable; however, please consider your appearance since you are still representing the company at a company-sponsored event.

During Setup
We understand the rigors of organizing learning and development events onsite and realize a business dress policy is inappropriate during that time. However, we would like to emphasize the desire for a neat appearance, even during this time. We encourage you to wear clean slacks (jeans are okay), and appropriate shirts during this time. Limit the amount of logos on shirts (e.g., t-shirts, etc.). You may also consider the type of shoes you're wearing to help ensure comfort and protect your feet from items that may drop or fall on them during the setup process. Clean tennis shoes or sneakers are acceptable footwear during setup.

Tip
Remember that you should always dress appropriately for the day regardless of our casual day policy. Be prepared to respond to client needs, and dress accordingly.

Questions
If you have any questions or comments about this policy, please speak with your performance manager or department leader.

Types of Venues

While it may be obvious that you need a classroom to conduct ''classroom'' or instructor-led training, what may not be as obvious are the number of options that are available. Your company's conference center or corporate university is always an option. If you have a training center, this section may not be as relevant to your needs but still offers some interesting tips to consider. Besides a corporate conference center, other options include using your training vendor's

Exhibit 19.7 Example of a Learning Policy

ACCESSIBILITY TO LEARNING EVENTS

In keeping with the company's EEO policies, the following accessibility statement must be placed on all L&D websites and included in all training announcements, invitations, and registration confirmations:

"If you have a special physical or communication need that may impact your participation in this training, please send an e-mail to Learning and Development at least two weeks prior to the start date of the program."

We expect these requests will be from professionals with handicaps such as vision and/or hearing loss or impairment, etc.; most of these professionals have standard accommodations that they are comfortable with and will continue to utilize. Should you receive a request, please forward it to [enter name], who will contact the requestor, handle the special accommodations for the participant, and keep a log of all arrangements made as a reference document.

space; an assortment of your own company's conference rooms in one or more locations; renting a local conference facility; and using a conference center or hotel conference facilities.

First consider your instructor-led training venue needs. Map out your classroom events and the specific needs for each event for some extended period of time, at least six months, as this will assist you in the negotiation process. We recognize that this defies the "laws" of good instructional design in that you are predetermining the delivery method prior to the final design of the training. But we assume that your organization will conduct a fairly consistent volume of classroom training in a consistent fashion using roughly the same number of classrooms each year. Table 19.2 offers some considerations when choosing a venue.

Keep in mind that many training organizations conduct training in a variety of types of venues. They may use their local office conference rooms for training that is shorter in duration and is mandatory for all professionals to attend. Conference centers or hotels may be used for large national/international programs when leveraging leadership presence and networking opportunities are critical. Other companies may arrange all of their training offerings based on the capacity and configuration of their own training facility.

Determining the needs for each instructor-led training offering will assist you in selecting the appropriate venue. Course content, instructional design, and the availability of instructors may dictate where your training is held. Some courses may lend themselves to a large-capacity facility, while others are better delivered in smaller settings at the local level. Nail down the details to be sure the site you select can host your event. Check on the dates of the program, the number and

Table 19.2 Considerations for Venue Options

Venue Option	*Considerations*
Internal Conference Room Space	Does training get first/high priority? Are there resources to perform logistical tasks? Do you need catering services for meals? Do you need lodging for out-of-town attendees?
Training Vendor's Space	May only be available for their programs. Do you need lodging for out-of-town attendees?
Rented Space from Local Corporation/Company	Each facility will have its own services. Do you need lodging for out-of-town attendees?
Conference Centers	Do you also need hotel space nearby? Generally have better business centers that are agile and equipped to handle large or complex last-minute materials print jobs. Long-term contracts versus ad-hoc contracts for negotiating pricing and services. Optimal lead time for securing space for your requested dates can be twelve months.
Hotel Conference Centers	Training events have different requirements than those needed for meetings. Make sure the hotel has experience with training. The larger hotel chains now have conference center core competencies. Check to ensure that their business center can meet your unanticipated needs. Long-term contracts versus ad-hoc contracts for negotiating pricing and services. Optimal lead time for securing space for your requested dates can be twelve months.

sizes of the classrooms, and other space requirements such as staff and instructor offices and materials storage space. If you need hotel rooms, you will also need to supply the venue with the number of single and double occupancy rooms needed.

Additional Venue Planning Considerations

Numerous other details are critical to the venue planning for your training event. They include:

- *Meal space*—Keep in mind that you may need two spaces that can seat your entire population—one for general sessions and one for meals. If using the same room for both activities, make sure to leave ample time for resetting the room between the activities. Remember this point for any space that will be serving two or more purposes.

- *Special events/receptions space*—These events might be a simple reception or a complex outdoor team-building event. Take the view of the host/instructor/facilitator of the event and then take the view of the learner/participant of the event. This will ensure you have covered the details, such as microphones for the instructors/hosts at the reception and the bug spray for the learners at the outdoor team-building session.
- *Audio-visual/technology needs*—These items include flip charts, white-boards, computers, projectors, screens, electrical connections, Internet connectivity, printers, copiers, fax machines, and phones. Remember that some of this equipment may be for the staff/instructor offices.
- *Food and beverage requirements*—There are many options to consider. Most conference centers and hotels offer all-meals packages that include all meals and snacks. We recommend using buffet service over plated meals, as it provides a wide variety of food choices and is more efficient for breakfast and lunch when time is an issue. Be sure to make advance arrangements for any special dietary choices, from a reception to boxed lunches for early dismissals on the last day of class.
- *Production services*—Leadership presentations or large general sessions may require staging, screens, lighting, audio systems, technicians, or other equipment. Most large hotel chains and conference centers can provide these services.

Venue Negotiations

It is obvious that the more space you require and the more often you require the space, the more clout you will have when negotiating with a venue. However, a nurtured long-term relationship and positive spend history with a hotel chain or facility will always be the most valued attributes in a negotiation process. Table 19.3 lists some of the customary concessions that can be provided when negotiating.

Major negotiated contracts may include such clauses as no charge or a reduced charge for a cancelled event if your organization schedules another same-size event within a stated period of time (this is usually six months). Knowing your ongoing and future venue needs and negotiating in advance with the sites will enable your organization to budget well for instructor-led programs.

Classroom Configurations

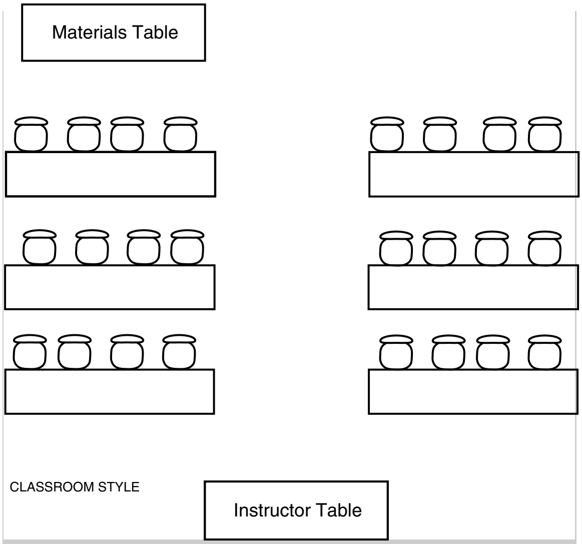
While determining a venue, it is important to determine whether the facility can support the type of classroom interactions you require. Simple ballrooms are not always appropriate. Table 19.4 lists some options for configuring classroom programs.

Table 19.3 Meeting Space Concessions

Type	*Concession*
Meeting Space	Waived meeting space rental fees given the guest room occupancy and food and beverage spending. One free meeting room per a certain number of guest room nights (given your buying power, this ratio can go as low as 1/40).
Materials	Complimentary handling of incoming shipments; typical charges are levied by the weight of the shipments, so negotiating a flat fee for all shipments may be a cost-saving alternative if ‘‘complimentary’’ is not offered.
Technology	Complimentary fax/printer/copier machine for staff office; complimentary AV equipment for classrooms such as flip charts, power strips, etc.; waived high-speed access charges in guest rooms.

Table 19.4 Classroom Configurations

Configuration	*Use*
Classroom	Instructor-focused and better for systems application training when viewing the front screen is essential. It is not conducive to team exercises.
Conference and U-Shaped (Figure 19.3)	Appropriate for small groups of fewer than twelve and conducive to discussions. While being instructor-focused, conference seating can also have a more formal nature.
Team Style and Crescent Rounds (Figure 19.4)	Optimal style for participation and team exercises when room capacity is under forty learners. More conducive to interaction and is learner-focused. Crescent round seating may be used in large general/plenary sessions, as it allows for interactivity, learner comfort, and easy viewing of the front of the room.
Theatre	Least effective for learning; little personal space and no tables for taking notes or referencing documents. May be appropriate for short leadership/key speaker presentations.



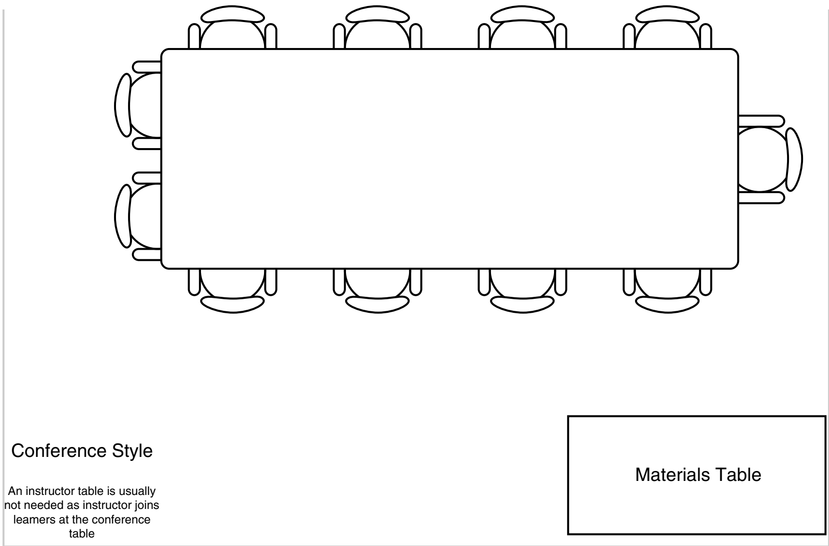


Figure 19.3 Conference Style Seating.

Regardless of which seating configuration you choose, remember that you need space in the room for AV equipment and additional tables for instructors, observers, and materials. Here are the formulas used by a large professional services company when accessing the participant capacities of rooms of each of the aforementioned configurations:

- *Conference/U-Shaped*—Square footage of the room divided by 18 = the maximum number of participants this room can accommodate

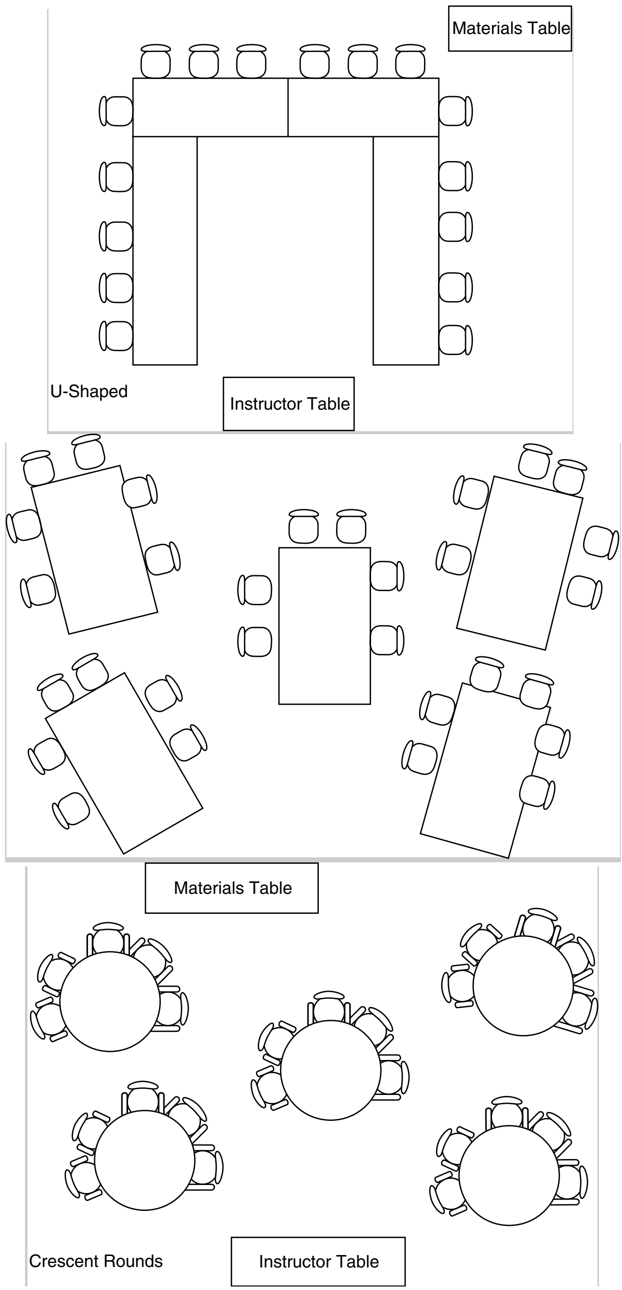


Figure 19.4 Team Style and Crescent Round Seating.

- *Team/Crescent Rounds*—Square footage of the room divided by 179 × 6 persons per table = the maximum number of participants this room can accommodate
- *Theatre Style*—Square footage of the room divided by 11 = the maximum number of participants this room can accommodate
- *Classroom* (2 participants per 6-foot table)—Square footage of the room divided by 26 = the maximum number of participants this room can accommodate
- *Major Production*—Square footage of the room divided by 32 = the maximum number of participants this room can accommodate

These calculations assume the following: each classroom has instructor and materials tables; an unobstructed room (no pillars, built-in furniture); and that the width of the room is more than half of the length of the room (that is, the room is more square than oblong). According to Jean Barbazette (2008), ''Be careful of odd-shaped or longer rectangles. (It's like teaching in a bowling alley!)'' (p. 325).

Post-Venue Selection

It is a best practice to have a venue secured (with a signed contract) before opening registration. It avoids all the learner inconveniences and administrative rework that a potential change in location may create.

Registration for training and registration for a hotel sleeping room may require different learner-specific information and different systems. Depending on your registration management systems (your LMS or other system), you will need to capture other details such as smoking/non-smoking sleeping rooms, special requests (dietary or handicap needs), arrival/departure dates not aligned with class dates, and nicknames for name badges, among others. The ease of registration for the learner must remain of paramount concern.

When using any external space, keep in mind the contractual agreements that need to be met prior to the actual event. A major one is your meeting room and sleeping room blocks—the number of rooms being blocked for your event. Usually these need to be confirmed two or three weeks in advance of the course start date. Attrition fees can be expensive, so your delivery team needs a way of managing registration and cancellations for the event that results in reliable numbers. Closing registration a few weeks out from the program and having a penalty for cancellation can be useful.

Impact of Facilities on the Event

Several facility challenges can impact the participants' overall reaction to the training program. Some of these are cleanliness of and excessive noise near guest rooms; uncomfortable chairs, temperature control issues, or excessive exterior noise in meeting rooms; not enough food or long lines at buffets; and poor transportation to and from the hotel. We strongly recommend a post-event facility/hotel evaluation be distributed to and completed by onsite staff and instructors. A comprehensive sample evaluation can be found in *Managing the Training Function for Bottom-Line Results* (Barbazette, 2008).

Global Venue Considerations

More and more companies are conducting training programs for employees from multiple global locations. How does the host learning and development team choose the right location and venue for an international audience?

Given the vast number of variables influencing global training and venues, we do not claim expertise on this topic. But based on our limited experience, we can offer some points to consider when selecting your venue for a global event.

The first thing to consider is the budget. Organizing and administering global training will be much easier if there is a global budget. Having all training development and delivery costs attributed to one budget will mitigate local budget constraints and eliminate charge-backs for tuition and lodging.

Another consideration is the number and locations of the participants who will be invited. If 50 percent of the learners are from one country, it may be prudent to conduct the training in that country. However, if that country is not centrally located for the other invitees, than travel time and costs may warrant a different location.

Once you have selected the city for your event, take the necessary steps to ensure your selected facility has a sustained track record of customer service excellence.

Venue Security. Most outside conference venues will have very detailed, effective processes and procedures to ensure the safety and security of their guests. The safety of the local neighborhood should also be considered. The security of your equipment and intellectual capital is also a concern. Additional security personnel may be required for your equipment. Your learners may need computer locks.

Are you OK with your competitors being present in the same hotel while your employees are there, or do you need a site where only your employees are present? This is a concern for some companies, and they will take extra security precautions to protect their intellectual capital.

''Green'' Venue/Transportation. Use facilities that have their own mature/effective (maybe even publicly recognized) green initiatives, such as eco-friendly laundry processes, using water coolers rather than bottled water, sending leftover food to local food distribution centers, and providing recycling containers.

Also consider using buses to transport participants in groups to/from airport and facility rather than individual cars/taxis. Provide information on alternative public transportation like the ''train to the plane'' in cities like Chicago and New York.

If you do not have a corporate travel service or meeting management department and learners will need to travel to training, you will need to have policies and processes governing their use of transportation.

We recommend outsourcing transportation arrangements. Many vendors specialize in these services. Airline flights, transportation to/from the learner's home to the airport, to/from the airport to the hotel; use of car services, buses, trains, taxis—all can be facilitated for the learner.

INSTRUCTORS AND INSTRUCTOR CARE

The selection, training, and care of instructors are critical to the delivery of training.

Instructor Selection

Will your instructors be internal employees, external vendors/contractors, or a combination of both? If your resources are internal, will they be full-time, part-time, or as-their-work-commitments-allow instructors? How will you ensure you have enough internal resources for instructing when you need them? The answers are essential to your program delivery success.

There are only a few basic needs for the transfer of skills and knowledge, in other words, for learning. There have to be a teacher, a student, and a learning need. While we cannot tell you who in your organization should be trained or what the content of the training should be, we can share some valuable information related to instructors and the delivery of training.

The selection of instructors/facilitators for your learning events is usually managed by the developers of the content. The delivery of training is made easier when the instructors are internal full-time employees of the L&D department in that you do not need to train them in instructor skills and they already should know the delivery processes and procedures affecting instructors. They are also paid to be instructors, which minimizes other work assignments conflicting with their schedules. External instructors are generally hired for their knowledge of a specific body of content and only need instructions on their administrative responsibilities related to the delivery of the course. Contracted relationships with external instructors minimize last-minute cancellations. The instructors who present the greatest challenge to the training delivery team are internal professionals with full-time responsibilities who are requested or volunteer to be instructors given their knowledge and experience in a specific or specialized area. While it is not deliberate, these instructors opt out of their training assignments more often than the aforementioned groups due to work load/client needs. To mitigate this risk, Jean Barbazette suggests a ''letter of agreement between the internal trainer candidate and that employee's supervisor'' (2008, p. 245). Basically, the delivery team needs to be agile and prepared to handle last-minute instructor replacements, including registration and logistics.

Our experience has shown that the delivery of training is smoother when there are two instructors in the classroom, whether physical or virtual. Based on our experience, co-facilitation is a best practice. While some instructors may not appreciate being paired, most enjoy it, as it generally enables a more robust learning experience for the participants. Some instructor pairing strategies are to pair:

- A content expert with a great facilitator,
- A less experienced instructor with a more experienced instructor, or
- An internal instructor with an external instructor/college professor.

Instructor Care

Once you have selected your instructors, we recommend that you go the extra mile to ensure they are trained and treated as valued contributors to successful learning. This is of particular importance when your instructors are internal professionals who have other full-time responsibilities. Consider the following instructor care tips:

Enable their instructor skills and content knowledge. Conduct general instructor training to allow subject-matter professionals to develop their skills and provide specific train-the-trainer sessions on course content. When training instructors for a specific program, include a piece on all the technologies that will be used: overhead projectors, retractable screens, audience response systems, simulations, printers, virtual technology, and so forth. Enable instructors to troubleshoot technology issues. Both internal and external instructors should be trained; don't just assume they know how to work with the technology.

Do not assume that a talented classroom instructor will be a good virtual instructor. The virtual world requires a specialized skill set. We strongly recommend rehearsals for all virtual training instructors.

Create ''rules of engagement'' for your instructors. These rules should include their attendance at train-the-trainer sessions, when they need to arrive/depart for the session, their administrative responsibilities while conducting a class, a ''code of conduct,'' and any financial/invoicing requirements. Minimize the number and complexity of instructor-related administrative tasks such as taking attendance; making announcements; and the distribution, administration, and collection of evaluations and tests. An obvious but often forgotten best practice is to contract with instructors to show up in their classrooms at least twenty minutes before start time to get organized and greet learners. This is critical to ensure a timely start to the session, the correct number of materials in the classroom, and a better learner experience all the way around.

When other training staff are not onsite, the instructor must be self-sufficient. He or she will need to take attendance, provide evaluations and tests to the learners, and return unused materials, evaluations and tests, and office supplies to designated employees. A pre-packaged instructor kit can be the solution; it may include the items mentioned above along with administrative instructions, pre-printed return envelopes, and any other required communication.

Provide onsite instructor assistance. Morning and afternoon classroom check-ins by the course manager can provide the instructor with a process for a debriefing or a review of upcoming content highlights or areas for emphasis. Another way to assist instructors is to cultivate and provide ''master instructors'' onsite, content experts who can assist during the delivery with key content, team exercise/simulation tips, and debriefing points. If master instructors are not an option, having an onsite person to assist with any content/facilitation questions is also effective.

Consider providing amenities for instructors. This is a way to treat the instructors as valued contributors while enabling them to keep current with other job responsibilities. Consider having the delivery team handle instructors' hotel and transportation reservations. Accommodate instructors in suites rather than standard hotel rooms. Provide an onsite instructor lounge—a place for instructors to relax and have a snack in the evening. Arrange for an onsite instructor office: a quiet room with a phone, an Internet connection, and a printer.

Establish an instructor recognition process. This should be formal and company-wide and include things such as offering ''comp time'' for all/part of the time spent instructing; considering higher performance ratings/promotions/bonuses for instructor excellence; and/or issuing and communicating awards for instructor achievements.

MATERIALS MANAGEMENT

Training materials are essential to a meaningful learning experience. This section will cover everything from the types of training materials to ''going green.'' For the purpose of this discussion, all documents and supplies that facilitate the learning experience will be considered materials.

Types of Materials

Table 19.5 lists common types of instructional and administrative materials.

Table 19.5 Common Types of Materials

Instructional Materials	*Administrative Materials*
Leader/Instructor/Facilitator Guide—Usually the participant materials with notes/facilitator tips, exercise setups, and debriefing points	Personalized agendas—An agenda for each learner that details the offerings for larger programs with an assortment of electives
Participant Materials—The key instructional points in print that support each learning objective	Week-at-a-glance—Used more for the administrative/instructor groups as a quick way to communicate the program offerings/ complexity, but can certainly be used as a type of table of contents for a program being offered over more than one day
Advanced Preparation—Any learning activity done prior to taking the WBT or attending the live session	Participant lists for participant networking—The participant list with their contact information to encourage a community of practice/networking after the event
Handouts—Exercises/Solutions/Case Studies that are kept separate from the participant materials for ease of use during the training and/or to ensure that participants don't see it prior to its use in the classroom	Name badges—Clip-on or lanyard badges with the participant's name/nickname and other information that may be important to the course, such as their functions/ departments or for participant networking such as client name, office location, etc.
Wall charts—To emphasize a key point, path, process of the training content, or to have as a constant reference during the program; also may be used during team assignments	Staff and instructor office supplies—Anything from extra copier paper to paper clips that keep the program running smoothly
Rosters/Sign-In Sheets—A list of registered learners for the program (enough room for each learner to indicate his or her attendance for each timeframe of the program in order to be compliant with any regulatory standard for tracking attendance)	Prizes, give-aways, candy—Anything purchased in advance of the program that will be used/distributed during the course
Evaluations/assessments—The end-of-class participant and instructor feedback	Registration packets—May include things like name badges, personalized agendas,

(*Continued*)

Table 19.5 Continued

Instructional Materials	*Administrative Materials*
surveys (Level 1 measurement instruments) and/or tests or exams administered after or at the end of class to assess the learners' learning (Level 2 measurement instruments)	facility maps, local restaurants, emergency cards, and so on

Electronic Materials

In our current environmentally conscious climate, offering materials electronically may provide a competitive advantage and can positively impact your training budget by eliminating duplication, shipping, and storage costs. It can also improve your speed to market by eliminating the time for duplication and shipping. Another advantage to electronic materials is that it addresses the volatility issue of training materials. Edits can be made quickly and the new version can be accessible within hours and with minimal cost.

Electronic materials may be presented in the following ways:

- Smaller documents can be made accessible as downloadable files from learning management systems or content management systems. This method works well for reference materials, advance preparation, and logistics documents.
- Larger documents, requiring more electronic storage space, can be made available for download or reference from any learner-accessible web-enabled system that has sufficient space. These larger documents might be learner materials or instructor/facilitator guides.
- Distribute CD-ROMs or flash drives. Distribution of CDs or flash drives can be done during onsite registration. This method ensures that each participant has the materials. Assuming the participants have ordered the CD/drive in advance and remembered to bring it with them has not proven to be effective. Of course, producing CDs/drives in time for the course is contingent upon receiving the content with the necessary lead time. Vendors are available who produce/copy materials in either way. To make this method ''greener,'' the content can be copied on CDs made of recycled materials and that are also recyclable. Be alert to the different ''grades'' of recycled CDs. Security should be considered: How will you prevent non-company people from gaining access to the contents? Encryption or passwords are some possible approaches.

Caution: All of the ''green'' methods discussed above assume that all learners have computers. Having computers in the classroom can also be a major distraction given the learners' access to the Internet and e-mail. Make sure you budget for electrical and Internet connectivity when conducting training in hotels or conference centers. Wireless cards and/or hard wiring can be undependable. We recommend testing all connectivity in advance of the training, having technical support and replacement computers onsite, and having a contingency plan in the event of connectivity failure.

Hard-Copy Materials

Materials can be copied in any number of ways, from having your delivery staff stand at a copy machine, using an in-house print shop, or having a printing/fulfillment vendor. Which way to go depends on the number of materials to be copied, the number of participants, and other issues posited below. First, does your L&D team have the required resources? Large-volume print jobs should be outsourced to your in-house print shop or to a printing/fulfillment vendor. Given the volatility of training materials, we recommend printing session by session rather than printing for all sessions up-front. While this requires more flexibility and adds time to the process, you can avoid the costs of reprinting, replacing a few outdated pages in large binders, and storing large volumes of printed matter.

A standard process and format for the submission of materials are essential. The developers should submit the documents in camera-ready format to save time and avoid misinterpretation of instructions. Camera-ready means providing each module/file in the way it should be duplicated: double-sided or one-sided; color or black and white; one slide or multiple slides to a page. It also means assembling the files in the required sequence with the dividers/tabs noted. This ensures the accuracy of the duplication work and avoids iterative conversations between the development and delivery staffs.

If you always have plenty of time to copy, proof, and ship your materials, then you only need to work with your in-house print shop or your vendor on your duplication schedule. If you have to accommodate a few last-minute/rush jobs, your print shop or vendor may be able to deliver these. However, if you have more rush requirements than ''normal'' requirements, you will need to find a vendor specializing in this service. Some vendors can guarantee twenty-four-hour turnaround. In some instances, you may need to make copies onsite, for example, if materials were not received in time to be duplicated in advance, a box of materials was not received onsite, and/or the wrong number of copies was made. The choices are using the hotel/conference center business center, having a high-speed printer as part of your staff office equipment, or using a local print vendor. Keep in mind that the hotel/conference center option can be costly. Another best practice is to use your ''twenty-four-hour'' turnaround vendor if the materials are not needed immediately.

Best practices for the duplication process include:

- Using a print request form ensures that you and the printer are using the same language to describe the same output.
- Binding of materials can take many forms. If ring binders are used, we strongly recommend shipping the binders empty. Shipping filled binders may cause the rings to break, creating an unanticipated clean-up task and a shortage of binders. Other popular binding alternatives include spiral and tape.
- Shrink-wrapping is efficient for assembling pre-counted materials for breakout rooms, supplies for team tables/classrooms, and so forth. For example, if you need to have twenty-five copies of an exercise solution for each of ten classrooms, then shrink-wrap in sets of twenty-five. It saves time onsite for counting and assembling materials.
- ''Kitting'' learner and instructor materials is also very efficient, especially for large training events using multiple classrooms. Learner kits could contain participant materials, a pen, a notepad, a tent card, and an evaluation form. Instructor kits could contain leader's guides, attendance rosters, and instructions regarding administrative responsibilities.

Storage

Most learning departments need storage for training materials. Your storage needs will be determined by the size and volume of your learning offerings and the number of materials produced. Storage could range from a small room with shelving located in your department to off-site storage at a fulfillment vendor. At any point in time, you will need to know what items and how many of each are being stored. The administrative tasks associated with the proper and accurate management of your inventory should not be minimized. Unpacking, packing, counting, recording, reordering, and shipping are just a few tasks. Many efficient vendors excel at these services for reasonable prices. Many ''fulfillment'' vendors also print and reorder low stock items in your inventory, allowing your team to focus on more value-added responsibilities.

Shipping

If all of your materials duplication, materials storage, and training are done in the same location, then shipping is not one of your training delivery needs.

But if you do need to ship materials, we recommend using one of the big, dependable shipping vendors for a number of reasons. They provide volume discounts, tracking of shipments on the web, proximity to all major cities/hotels/conference centers, and insurance on your shipments. Meeting with your shipping vendor to discuss your specific needs can avail you of services that you may not be aware of otherwise, such as a representative to be onsite at your large events if necessary or customized services such as invoicing or special-size boxes.

Additionally, we recommend that you set up and use an internal shipping label template. This assists the team with the appropriate labeling of boxes. It also ensures that you know what's in each box when tracking, logging in, and sorting materials onsite.

If your organization works internationally, we recommend the use of a fulfillment vendor with international locations and international shipping experience. This will prove invaluable with services such as printing and storing materials locally and navigating through all the international customs regulations.

Best Practices for Materials Management

While there are many ways in which an organization can manage materials, we recommend the following:

- Arrange to have a box room at your training venue where all materials can be stored and sorted. Arrange to get the key to this room in advance, as those setting up the training will arrive before the participants. The room should have double doors to accommodate pallets and distribution hand-trucks and be close to the classrooms being used. If you are running the program's classes in two facilities near each other, then have a box room at each location.
- Onsite materials management is a key to effectively managing classroom training for three hundred or more learners. This service is usually provided by a vendor who becomes familiar with all materials beforehand by reviewing a materials grid prior to arriving onsite; arrive onsite one or two days early to open all boxes, log all materials, make note of what's there and what isn't there; make copies of anything that didn't arrive; sort the materials; and set up all classrooms. These vendors may also be contracted to assist with other administrative tasks while they are onsite, such as delivering and collecting rosters, evaluations, and tests; manning the information desk; distributing announcements; and other tasks. At the end of the program, a materials management vendor will collect unused materials/staff office supplies and ship them back or arrange for disposal or shredding.
- To ensure that your onsite staff/instructors have basic office supplies, consider having ''site boxes'' provided by a national vendor with locations in most major cities. Find a vendor who can supply these items locally to

avoid storage, pick/pack, and shipping and restock charges. An alternative is to have a fulfillment vendor stock your site box items for you and then ship them to training locations upon your request. See Exhibit 19.8 for an example of a site box order form when using a fulfillment vendor.

- Contract with a ''shredding'' vendor for leftover/extra proprietary materials that must be destroyed. There are national vendors that will pick up from most national locations and shred the materials.

''Going Green'' with Materials

Obviously, the easiest way to eliminate hard copies is to go electronic. This instantly cuts your printing, storage, shipping, and distribution costs and decreases your carbon footprint. It assumes a ''paperless'' culture where everyone has a computer and a working Internet connection to download content and that you have the content in advance to allow time for downloading. Environmental impact can also be reduced from using less drastic measures.

One option to consider is the use of recycled materials. Paper made of recycled products is readily available. Paper can also be manufactured using more environmentally friendly processes. There are binders made of 100 percent recycled materials. Most common classroom materials, including pencils, pens, flip-chart pads, and CD ROMs, are now also being made of post-consumer goods.

We also suggest printing only what is needed in the classroom, such as exercise materials and key content. Print PowerPoint slides in outline format three to a page. Reduce the size of the paper from to 8½ by 11 to 8½ by 7 without sacrificing readability. These practices can not only meet your going green initiatives, but also save the organization unnecessary expense.

TRACKING AND REPORTING

For the purposes of this section, tracking and reporting are being differentiated from ''measurement.'' *Measurement* pertains to assessing the efficacy of learning activities, and an entire chapter (Chapter Fifteen in this volume of the *Handbook of Improving Performance in the Workplace*) and all of *Volume Three: Measurement and Evaluation* are dedicated to this subject. Tracking and reporting are defined as the activities necessary to provide basic information on usage, accuracy, timeliness, and compliance. Learning organizations track and report in order to:

- Have an accurate picture of what they are offering;
- Respond to regulatory board requirements;
- Justify the time, money, and resources spent on training;
- Provide training history/transcripts to the learners;

Exhibit 19.8 Request for Materials Return from Site

Course/Session:			Job Number:
Program Title:			Project Number:
New	**Revised**	**Cancelled**	**Return to Inventory**
Site Box must be delivered by Date:			
Billing Information			
Course Manager Name:	E-mail Address:		
Shipping Information			
Requester Name:	Requester E-mail Address:		
Attention: (Please Note If a Guest of a Hotel)	Site Name:		
Street Address:			
City, State, Zip Code:			
Site Box Materials	**Vendor Item Code**	**Quantity**	**Revised Quantity**
KIT: Tent card, pencil, pad (1=1)	Pencil, Pad, & Tent Kit		
Pencils (1=1)	Pencil – blue		
Note Pads (1=1)	Notepad – Logo		
Tent Cards (1=1)	Tent Cards		
Transparencies (1=1)	Transparencies		
Overhead Markers (1=1 box of 4 markers)	Markers – 4 Overhead		

(*Continued*)

Exhibit 19.8 (*Continued*)

Flip Chart Markers (1=1 box of 4 markers)	Markers – 4 Flipchart
Highlighters (1=1)	Highlighter
Index Cards (1=1 pack of 100)	Index Cards
Packing Tape Roll w/ Dispensing Portion	Tape Gun w/ Tape
Masking Tape (1=1)	Masking Tape
Tape Dispenser (1=1)	Tape Dispenser
Tape Refills (1=1)	Tape Roll
Three-Hole Punch (1=1)	Three-Hole Punch
Glue Stick (1=1)	Glue Stick
Course Evaluation Form	Form Number
First Aid Kit (1=1)	First Aid Kit
Speakers – Audit (1=1 set)	Speakers – Audit
Stapler (1=1)	Stapler
Staple Remover (1=1)	Staple Remover
Staples (1=1 box of 5000)	Staples
Scissors (1=1)	Scissors
Medium Binder Clips (1=box of 12)	Binder Clip – Medium
Envelope – Letter Size (1=1)	Envelope – #10

Roster and Evaluation White Envelopes	Envelope 10 × 13 Plain
3 × 5 Post-it-Note (1=1 pad)	Post-It Note-3 × 5
Large Paper Clips (1= box of 100)	Paper Clips – Large
Plain Pens (1=1 box of 12) Blue	Pen – Blue
Rubber Bands (1=1 bag)	Rubber Bands
Utility Knife (1=1)	Utility Knife
White Out (1=1)	Liquid Paper

E-mail Site box Request form to: (enter e-mail address)
Shipping Address for Returns to Inventory: (enter address)

- Ensure the employees are consuming required training and/or continuously developing; and
- Demonstrate the development of employees to potential candidates and other concerned externals.

Your organization's reasons for tracking and reporting will be dictated by any number of variables, from your organization's business/industry to the style of your top-level management. Some tracking is essential to all learning and development departments: the number and delivery type of course offerings, how many learning hours/credits each offering is worth, and how many learners attended/consumed each offering. Tracking of learner demographics is also helpful for ''slicing and dicing'' data by different populations.

If you are establishing an L&D function for your company and you do not have regulatory requirements prescribing your ''units/credits'' awarded for learning events, we strongly recommend that you make this one of your first tracking decisions. A generally accepted unit of training is a training hour. When tracking hours at classroom courses, make sure you do not include breaks and meal times.

Tracking is one of the main advantages in having an LMS or web-enabled training site. These systems provide online access to most data with a few simple clicks or by running a standard report. Manually tracking the essential training data requires a lot of organization to monitor processes and multiple resources. Attendance tracking is a particularly important process, especially in highly regulated/compliance-driven environments. Most regulatory bodies that oversee such industries as pharmaceuticals, engineering, medicine, and auditing have standards for the continuing education that professionals must acquire to maintain their licenses. This requires specific evidence of attendance for each learner. Table 19.6 lists some approaches to attendance tracking for classroom training.

Tracking attendance in virtual sessions is easier, given that these systems generally provide log-in/log-out times for each learner as well as how many interactions the learner responded to. This tracking will suffice for most regulatory requirements.

Tracking consumption of WBTs has also become easier with an LMS and the use of technical standards. Tracking can be done automatically and provides the following types of information: number of learners who launched the content but did not complete the course; the time it took for each learner to complete the course; final test scores; and a list of the learners who successfully completed the training.

When you summarize basic tracking information on your training and combine it with financial data (actual training expenses), you will be able to provide the more common training delivery reporting such as:

Table 19.6 Approaches to Tracking Attendance

Method	*Description*	*Considerations*
Sign-in rosters	Rosters can be generated reports from an LMS or a web-enabled training management site containing the names of those who registered for the program and space for them to check, sign, or initial their attendance. In the absence of an LMS, rosters can be a manually produced Excel spreadsheet with the same information. Multiple rosters can be used for attendance tracking in large plenary/ general sessions by providing copies on each table or row.	Roster should contain the names of the participants rather than having the attendees write their names. Deciphering handwriting can be an administrative challenge and time waster. If your attendance tracking requires the marking of full and partial attendance, then the roster will need to provide space for the instructors to make these notations and for the instructor's signature. Copies of multiple rosters will need to be collated onto one master roster after the program.
Collect business cards	Provide collection bins at each of the room's entries.	You will need to remind learners to bring their cards. The delivery team will need to check the cards against the registration list after the session.
Instructors mark attendance directly in the LMS.	Instructors manually enter participant attendance in the LMS.	Recommended for smaller events. Requires a computer, Internet/systems connectivity, and instructor training.
Leverage results collected in the audience response systems (ARS).*	ARS systems are applications that enable learner interactivity in the classroom through hand-held keypads. These devices can track an individual's responses to questions asked by the	Using ARS responses becomes challenging when partial attendance needs to be tracked or when specific types of study need to be associated to specific individuals.

(*Continued*)

Table 19.6 (*Continued*)

Method	*Description*	*Considerations*
	instructor. Reports from these systems on learner responses can be used as evidence of attendance.	
Use a bar code scanning system.*	Systems are available that will record an individual learner's bar code as he or she passes it through a scanner. The reporting from the device can serve as evidence of attendance.	Keep in mind that bar codes will need to be assigned and made available to each learner in advance of the program.

*An LMS can generally accept uploaded attendance files from these technologies.

- Volume of training by delivery method and content skill type (enabling, technical, leadership, etc.);
- Uptake of training;
- Costs per course per delivery type, per learning credit, and per participant day;
- Costs for development and delivery of training;
- Percentage of those in the company who attended training; and
- Average number of learning hours/credits earned by different number of learning units by every cut of your population that is possible.

COMPLIANCE-DRIVEN ENVIRONMENTS

Learning and development functions in certain industries (such as financial services, accounting firms, pharmaceuticals) face training compliance standards that require rigorous and detailed education, communication, tracking, and reporting processes.

All learning managers should have a working knowledge of the standards/ rules as stated by the pertinent regulatory board or governing body. We strongly recommend assigning a senior L&D professional the responsibility of being the in-house expert. This professional can provide practical interpretation of the standards and liaise with regulatory board and other ''like'' companies. This provides obvious benefits to the organization as a whole by ensuring the alignment of learning processes with compliance requirements. However, the person who fills this role can also be your organizational voice to the regulatory body, providing valuable perspectives on the implications and interpretation of the standards, which can impact future regulations.

Educating professionals on their individual training compliance responsibilities is another critical success factor. Their responsibilities should be explained to learners as clearly and frequently as possible. Make them part of the new hire orientation program and every training event that is offered. Providing compliance guidance on easily accessible websites can assist in the education of your professionals. Periodic reporting and/or LMS learner interfaces that provide the learner with the delta between compliance requirements and a learner's current status will facilitate the education process.

Embed the rules/standards into the key affected processes. For example, if the standard for training attendance tracking states that each learner must physically sign a roster, then make sure your processes include the production of a paper roster that is circulated in the classroom, instructions to the trainers to circulate the roster, and a method for collecting each roster. Provide quality assurance steps in the standard operating procedures for all learning processes and conduct periodic internal audits of the processes. This is always recommended, but is critically important for compliance-related activities. Institutionalizing the required education/communication, tracking, reporting, and document retention mechanisms is the ultimate goal. Seamless integration of these mechanisms will avoid redundant efforts and last-minute fire drills.

Compliance-driven organizations must determine the consequences for any non-compliant learners. Some typical consequences include negative performance ratings and stalled promotions; learner counseling; and in some cases termination of employment. These consequences need to be communicated frequently. There also should be pre-emptive reporting, messaging, and an escalation process that supports timely compliance by learners.

It is important to consider the impact that compliance requirements and the related education/communications have on learners. If the organization's training compliance ''messaging'' is not balanced with the importance of professional development, the climate can turn into one in which the learners become numb to the true benefits of training and simply consume the training to meet the requirement and stay off the non-compliance list. To counterbalance this sort of atmosphere, learning and development teams in regulated industries must provide the most engaging and efficacious learning activities their resources can support.

LEARNING DELIVERY THROUGH TECHNOLOGY

The use of technology-based learning methods has been on the rise for the last decade (ASTD, 2007). In 2006–2007, nearly one-third (30 percent) of all learning provided was technology based and just over one-fourth of learning was consumed through technology-based methods of delivery (ASTD, 2007; Bersin & Associates, 2007). The advent of learning management systems (LMS) and the ability to ''plug'' content into them have made scalable enterprise consumption of e-learning possible. Technology has thus become a major part of the learning function, particularly in large organizations in government, banking/finance, industrial/manufacturing, high-tech, healthcare/pharmaceuticals, and insurance industries (Brandon Hall Research, 2008).

While it may seem new, learning via technology-based means has been around as long as or longer than telephones and television (Masie, 2006). Many large corporations began training over the telephone as soon as individual telephones became common in the corporate setting. Learning through technology has become pervasive, however, with technological advancements such as the home PC and the World Wide Web; learning standards such as SCORM; and, more recently, virtual worlds and social networking technologies. To determine what is next for technology and learning, all one needs to do is track general technology advancements and consumer innovations, and draw parallels and analogies as to how these technologies will play out in the learning world. In Figure 19.5, one can clearly see the innovations that led to the ''dot-com'' era (web browsers, home PCs, and consumer broadband, for example) and the learning innovations that followed (virtual classroom technology and web-based learning management systems).

Clearly, as organizations merge, globalize, and become increasingly complex, and as the pace of business change continues to be extraordinary, technology as an integral part of the learning organization will continue to grow. Organizations that invest in learning management systems have much higher efficacy and efficiency levels than those that do not (Bersin & Associates, 2006). However, implementing enterprise learning systems is challenging, complicated, and requires ''boutique'' technological knowledge, process expertise, and knowledge of the learning industry. One reason for these challenges is that learning systems are relatively new. After all, the first major learning management system was only released in 1997; corporate enterprise resource planning (ERP) and human resource information systems (HRIS) have been around far longer and are thus more stable and easier to support. A second challenge is the newness and complexity of the standards that govern learning content interoperability: they were first released in the 1990s, and only formally adopted by the IEEE in 2002. Interpretation of the standards has not been consistent or clear at this time.

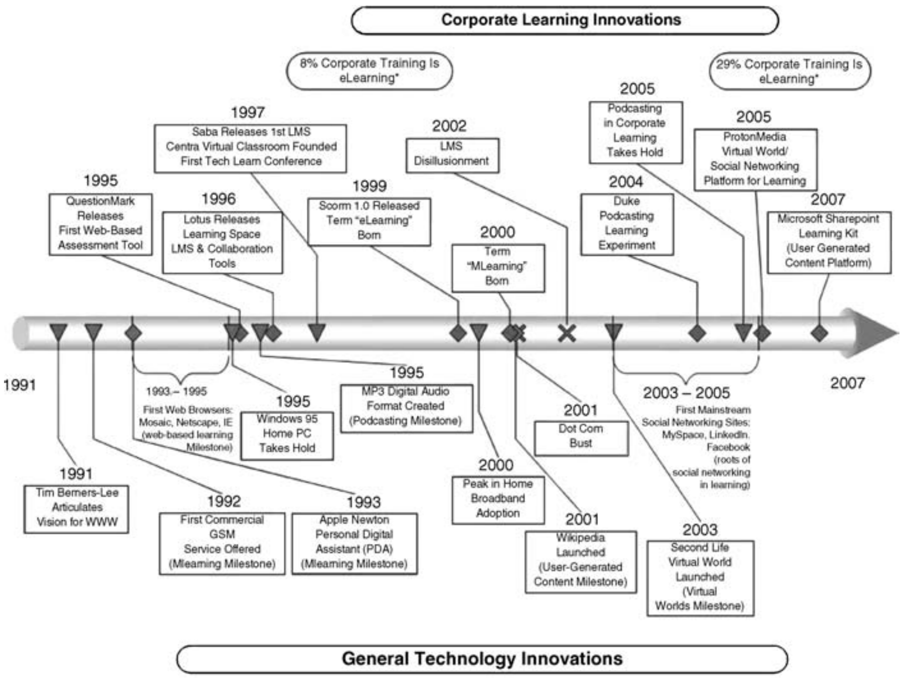


Figure 19.5 History of Technology-Based Learning.

Source: American Society for Training and Development

ENTERPRISE LEARNING TECHNOLOGIES AND SYSTEMS COMPONENTS

The technology infrastructure required for a standard enterprise learning platform can include some or all of the following components: learning management system (LMS), virtual classroom (VC), assessment/e-testing system, database, reporting system, learning content management system (LCMS) (less common), and learning analytics (less common).

Learning organizations that have sophisticated enterprise learning systems struggle to integrate the individual components, which consist of many disparate systems (see Figure 19.6). The power of these systems in the future will hinge on the ability of an organization to leverage the capabilities of its learning management systems for performance management, talent management, and knowledge management. The LMS is seen as the most important learning technology investment priority for senior learning leaders, according to a 2007 IDC survey of the *CLO* magazine Business Intelligence Board, nearly 50 percent of companies will invest in upgrading their LMSs in the coming fiscal

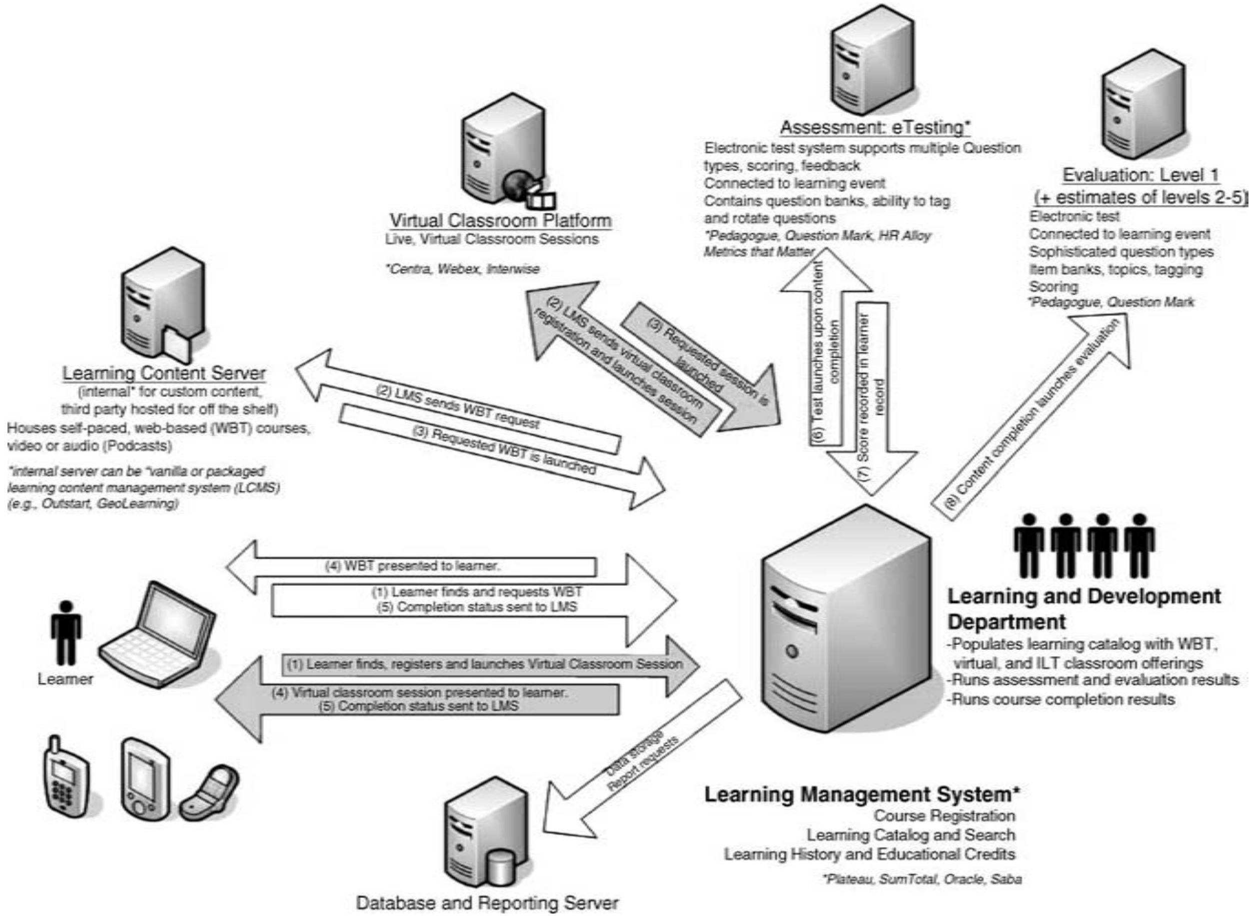


Figure 19.6 Enterprise Learning Systems Architecture.

year, consolidating other LMS systems or integrating them with talent management solutions (Lee, 2008). Other reasons for this undertaking include a high degree of dissatisfaction with LMS systems and the need to consolidate multiple LMS systems into one enterprise-wide system. It is costly, but provides significant efficiencies. To manage these platforms, learning organizations typically allocate 5 percent of their budgets to learning technology operations and maintenance (Bersin & Associates, 2007).

Learning Management Systems (LMS)

The traditional function of an LMS is to power learning and to be the central location in which all of a learning organization's offerings and completion records are housed. Its primary functions include learning event/offerings, registration, status, and completion data. Some LMS systems also offer learning analytics, content management, assessment, and reporting. An LMS generally serves as the repository for all data and information related to learning. It is not a human resources information system (HRIS): the HRIS is usually the repository for all people data (''rank, file, serial number''). The LMS and HRIS are usually integrated so that data found in the HRIS populates the system that needs the information to operate the LMS. SAP and PeopleSoft are the HRIS systems that most commonly provide people data to the LMS. Many HRIS systems do have LMS functionality; however, at the time of this writing, they are typically not ''best of breed.'' HRIS vendors continue to try to break into the LMS market.

The LMS started as the central learning systems platform. As of 2009, many LMS vendors are reinventing themselves as human capital management systems, attempting to morph into the system of record for not only learning, but also for the definition of organizational skills and competencies and the learning diagnostic/prescriptive engine for competencies and learning. Some other systems are ''best of breed'' for competency management. Determining which system to use depends on the size and complexity of the organization, the sophistication level of the functionality desired (if low, then a one-size-fits-all system can often work), and the size of the maintenance and integration effort, should a best-of-breed system be selected.

The LMS is not always the front-end interface. Today, some LMS systems are used as the learning system's back-end database application, and web services are used to leverage key LMS functions in other corporate applications (for example, the learning registration ''web part'' may be placed into the web page of another system, such as the performance management system).

According to Brandon Hall Research (2008), over 65 percent of organizations have an LMS, and they are most commonly found in mid- to large-sized organizations with sales of over $100 million. Most organizations are on their second or third LMS, and many have more than one. LMS systems are notoriously difficult to implement and maintain effectively due to the immaturity of LMS systems (they are relatively young compared to HRIS systems) and the lack of skilled personnel to manage them. Successful operation of the learning management system requires that one group, preferably a very technical team inside the learning department, ''own'' the LMS and all other enterprise learning systems, ensuring that they are properly integrated with one another and with the other systems of the organization. This group must set the standard operating procedures that govern data input to ensure system data integrity or work closely with the responsible team. This is extremely critical and should be a well-resourced component of a centralized, ''shared services'' learning systems or operations group.

Learning Content Server

The learning content server houses self-paced, web-based learning content. Some organizations have a server that has folders organized to house content by subject, business unit, or some other organizing principle. The LMS communicates with the learning content server when a learning content request is made by a learner. This type of server can be considered a ''poor man's'' learning content management system (LCMS).

Other organizations use a more sophisticated, off-the-shelf learning content management system (LCMS) application. An LCMS also serves as a central authoring system that has learning content authoring workflow for designers and subject-matter experts and facilitates a high degree of content reuse. An organization should have a high level of custom content volume to merit the investment in yet another system that must be maintained and integrated with other learning systems applications. The use of LCMS systems is surprisingly low in corporations, and generally employed by those that have a high volume of custom content authoring.

Virtual Classroom Platform

A virtual classroom (VC) platform provides a live environment in which to host a virtual class. VC platforms can usually host up to one thousand participants at once, although most organizations keep class sizes small (below fifty). VC platforms usually provide a list of the students who are in attendance and icons next to a student's name, which the student clicks to communicate with the instructor or the rest of the class (to answer ''Yes'' or ''No,'' for example). Teaching tools in the VC include a whiteboard with instructor and student markup functionality, ability for the instructor and students to share an application and mark up a screen, ability to share in a web-page navigation view, chat, polling, in-class testing/evaluation, virtual breakout rooms, and the delivery of multimedia such as audio and video.

The VC platform is usually integrated with the LMS so that its classes appear in the LMS catalog and learner completions are registered on the learner's transcript in the LMS. This level of learning systems integration can contribute significantly to efficiency in a high-volume operation.

Creating effective virtual classroom sessions is a challenge for several reasons: (1) without the direct physical contact of the traditional classroom, instructors can't read body language and students are more likely to multitask; (2) there is a tendency for VC sessions to be ''shovelware,'' or PowerPoint dumps; (3) VC is often the solution an organization employs when it needs to educate its employees live and in a hurry; (4) the speed with which employees are reachable through this medium tends to cause learning design and development timelines to be significantly abbreviated; (5) there is a lack of technology-based instructional design skill in the learning marketplace; and (6) there is a lack of technology-based instructor facilitation skills. If the instructor is hard to follow in the classroom, putting him in the virtual classroom is deadly for student engagement. KPMG LLP, a ''big four'' accounting firm, executed a study that found that poor instructor facilitation negatively impacts the overall course satisfaction rating by .39 points on a 5-point scale; instructors who were trained performed better, on average, receiving an instructor rating that was .25 points higher than untrained instructors (Hanssen, 2007). Refer to Chapters Twelve and Thirteen for more detail on how to effectively use virtual classroom and virtual worlds as learning solutions.

A less obvious but still important factor for the success of a corporate virtual classroom platform is that, if your organization does not have an alternative collaboration software solution for non-training events (meetings, large-scale communications), then it will not take long for members of the organization to discover how to use the training department's virtual classroom platform for other important corporate uses (particularly sales, client-facing, or other externally focused applications). A solution to this is for the training department to work with the group responsible for live meetings, helping them to add a ''virtual'' meeting service to their offerings. This can be done using the same platform (Centra, WebEx, Interwise), but different graphics and branding (it can even be synergistic in terms of platform maintenance and upgrading). At KPMG LLP, Meeting Services Online® (MSO) went live eighteen months after the virtual classroom platform because training was getting a reputation for something important but not learning-focused. Today, KLEARN LIVE®!, a virtual collaboration session that is branded for learning, has vastly different interactions and exercises and is designed with learning objectives in mind. A ''virtual collaboration decision tree'' was created for end-users to send them to the right department for support, depending on what their objectives were. KPMG LLP employees know what to expect when they attend an MSO versus a KLEARN LIVE! session. Interestingly, MSO volume surpasses KLEARN LIVE!

volume today: thank goodness, or the training department's resources would be diverted to running meetings and supporting corporate communications instead of designing interactive learning sessions designed to improve job performance (and vice versa).

A new type of virtual classroom environment is emerging: the virtual world. A virtual world is a more realistic, graphical representation of a place (usually with people) that has a ''third'' dimension: instead of a web page with ''flat'' graphics representing locations and people, a virtual world has ''three-dimensional'' graphics that cause the user to have a more realistic and lifelike experience (picture the image of a simple triangle versus an image that shows a triangle plus a ''shadow'' that represents its backside). At the time of this writing, Second Life was the most well-known virtual world, and many organizations have set up virtual spaces in Second Life for purposes of training. ProtonMedia was one of the first virtual world products designed for corporate settings. Its primary benefit is that it is a closed environment and not open to other members of the Internet. IBM's CEO, Sam Palmisano, has said that the 3D web is the next evolution of the Internet because it provides a dimension that enables the Internet to have immersive, realistic qualities. Second Life and Proton Media are good examples of how the 2D virtual classroom may evolve into a more immersive learning platform in next generation, 3D web.

Assessment/Testing. An electronic assessment system provides the ability to administer a test to participants immediately after an event. Typical testing system functionality includes:

- Presentation of immediate feedback;
- Results to the learner;
- Repeat tests administered to participants at set timelines after the learning event;
- The ability to limit the number of attempts to complete an exam;
- Test scheduling for specified groups;
- Random presentation of questions and choices to allow for multiple tests/ retests;
- Multi-author capability;
- Secure delivery for high-stakes exams;
- The provision of results to performance managers when participants are in need of remediation;
- The creation of large, topical item banks that can then be assembled to make individual tests;
- Multiple question types (multiple choice, drag and drop, etc.);

- Computer adaptive testing; and
- Item meta tagging, and more.

A primary benefit of electronic testing is the consistent availability of rigorous psychometric test statistics, which would be tedious to produce manually. These statistics, such as P-value and item discrimination, are critical for creating assessments with high evidence of reliability and validity. So for the organization that wants to get serious about testing, an electronic assessment tool is a requirement. Other reports typically produced by these systems include a summary of all questions for a class in aggregate format, with more tactical analysis at the question level; pre- and post-test data computation of knowledge gain scores; and triggered notifications such as when a specific learning event (class) falls below a pre-set threshold.

The testing platform can be integrated with the LMS so that tests are associated with learning offerings and result in assessment scores for the learner in the LMS. This level of learning systems integration can contribute significantly to efficiency in a high-volume operation.

End-of-Program Evaluation. The majority of companies track learner satisfaction (Level 1) (Kirkpatrick, 1998). This is typically handled either through paper, a general web-based survey system, the in-house electronic assessment system, or a specialized Level 1 learning evaluation system.

Table 19.7 depicts the benefits and drawbacks of each approach. It is noteworthy that, in the early 2000s, there were no systems that elegantly handled both end-of-course learning evaluation and assessment/eTesting, yet both require similar functionality (that is, ability to present a question and allow an answer to be entered). The primary difference between what these types of systems should provide is reporting.

Database and Reporting Servers. The LMS relies on a database to store data for the application (Oracle, MS SQL). The reporting application always pulls its data from the database and either (1) also sits on the database server; (2) is part of the LMS; or (3) sits on a different, specialized reporting server. Many organizations that have become serious about measurement house learning data in a data warehouse that also includes HR data (such as turnover and promotions) and enterprise data (sales, productivity) so that complex analytics can be run to determine the impact of training on the bottom line.

Reporting is usually the bane of the learning organization's existence, unless the organization has excelled in building a reporting strategy, tools, and platform. Typically, LMS reporting tools are not enough. If LMS reporting tools are used, they often need to be supplemented and generally provide data, but not robust information (that is, learning analytics).

Table 19.7 Learning System Evaluation Types

Type	*Benefits*	*Drawbacks*
Paper	Higher response rate due to in-class completion.	Time delay in receiving results. Very difficult to standardize questions. High distribution costs (people operations, paper, printing, report distribution). Lack of accurate, real-time demographics of respondents.
General Web-Based Survey System (SurveyMonkey, Web Surveyor)	Lower distribution costs. Easier to standardize questions. Real-time results.	Additional system for learning group to manage. Lack of learning specialization in questions. Lack of learning analytics (predictive Level 3 or 4 statistics).
Assessment/eTesting System (Pedagogue, Question Mark)	Lower distribution costs. Real-time results. Cost reduction in having one system handle both assessment and evaluation.	Lack reports and statistics relevant to learning evaluation (they are focused on testing statistics and analysis). Generally tough to standardize questions, as it is not designed to administer the same survey over and over again.
Specialized Level 1 Learning Evaluation System (Metrics That Matter)	Lower distribution costs, especially if it can be integrated with LMS and automatically triggered upon learning event completion. Real-time results. Learning specialization in questions. Learning analytics (predictive Level 3 and 4 statistics).Powerful internal benchmarking on key, learning-specific dimensions. External benchmarking possible in hosted system.	Additional system for learning group to manage (unless it handles testing).

Usually the combination of (1) housing LMS data, assessment data, and evaluation data together in an HR data warehouse and (2) the use of a sophisticated reporting application enables robust learning management reporting and analytics.

One must use caution when selecting database/reporting solutions. They have varying ranges of scalability. A system that can handle up to fifty reporting users may not scale to one that requires reporting access for one hundred or one thousand users. When architecting solutions, business analysts ensure scalability by determining parameters such as the number of database records that must be retrieved to produce a given report, the number of users who require the available reports, and the number of total reporting hits that a server will be expected to handle during peak times.

Wikis, Blogs, Podcasts, and Other Forms of Informal Learning

The popularity of user-generated content exploded in the past decade, with the advent of user-friendly web authoring technologies, Wikipedia in 2001, and the ease with which digital audio can be created and distributed.

Wiki is the Hawaiian word for ''fast.'' A wiki in the web world is an application that easily enables users to create, link, store, and update content collectively in one place. It is considered a ''Web 2.0 technology'' because of its user-generated nature. Wiki technology often underlies some of the most powerful, collaborative sites on the web. One of the greatest wiki learning phenomena is Wikipedia; as of 2008, it contains ten million articles in 253 languages and is one of the most popular reference websites on the Internet. The key to the phenomenon is that it continues to be created collaboratively by users around the world; one would be hard-pressed to build such a vast, comprehensive encyclopedia with any amount of corporate resourcing and initiative. A criticism of wikis is the lack of authentication and verification of information. As wiki technology increases in sophistication, these limitations are being mitigated.

Podcasting gets its name from the iPod digital audio device from Apple Computer. A podcast is simply an audio clip. While many listen to these audio clips on iPods, podcasts can be heard on personal computers, BlackBerrys, Treos, cell phones, or anything that plays digital audio. Duke University's famous 2004 experiment, in which it provided 20G iPods to 280 freshman in nineteen courses for the purpose of replaying lectures, popularized podcasting for learning. Many organizations employ podcasts for learning today. Creating and posting a podcast is simple, can be done from any laptop, and can be learned from various sources on the Internet.

A blog is a website maintained by an individual that contains a chronological list of entries. Generally, it is an ongoing journal of commentary by the individual. ''Blogging'' refers to the practice of regularly posting commentary on one's blog-formatted website.

There has been much talk about the use of wikis and blogs for learning, but little in terms of standards, best practices, and research on learning and performance impact. KPMG LLP is experimenting with the use of wikis as an instructor-led tool for building and housing course materials (as a green initiative), based on input from its top subject-matter professionals.

LMS systems have acknowledged these new forms of informal, user-generated learning content. Many now provide a course site that contains a course wiki, student blogs, and the ability to host podcasts.

Anther type of informal learning is ''referenceware.'' Two examples include book summaries or full-text searchable books online. Newer LMS systems enable the inclusion of this type of content. This functionality, however, tends to overlap with that of knowledge management systems; thus, many organizations for whom KM is also a priority are determining strategies to ensure that formal and informal learning content is synergistic versus duplicative and confusing to find for the end-user.

The availability of informal learning and the popularity of social networking sites, combined with the sheer volume of new information that becomes available on a daily basis, are sure to change the learning landscape over the next decade. While there are a lot of unanswered questions about how user-generated content can be relied upon and measured for business impact, the old method of highly structured/designed content is fast becoming outdated due to its highly impractical nature in the information age.

References

American Society for Training and Development. (ASTD). (2000). *Profiting from learning: Do firms' investments in education and training pay off?* Alexandria, VA: ASTD.

American Society for Training and Development. (ASTD). (2007). *State of the industry*. Alexandria, VA: ASTD.

Barbazette, J. (2008). *Managing the training function for bottom-line results*. San Francisco, CA: Pfeiffer.

Bassi, L. J., & McMurrer, D. P. (2002). *Investing in companies that invest in people*. Retrieved June 16, 2008, from http://www.bassi-investments.com/downloads/Article_hr.com.pdf.

Bersin & Associates. (2005). *The high-impact learning organization*. Oakland, CA: Bersin & Associates.

Bersin & Associates. (2006, October). *The high-impact learning organization*. Oakland, CA: Bersin & Associates.

Bersin & Associates. (2007, February). *The corporate learning factbook 2007*. Oakland, CA: Bersin & Associates.

Bersin & Associates. (2008). *The high-impact learning organization. What works in the management, governance and operations of modern corporate training*. Oakland, CA: Bersin & Associates.

Brandon Hall Research. (2008). LMS Knowledgebase 2008: Introduction. Retrieved June 5, 2008, from http://www.brandonhall.com.

Caudron, S. (2003, February). CLO accountability. *T + D*, *57*(2), 36–46.

Chief Learning Officer Business Intelligence Industry Report. (2008). Retrieved June 5, 2008, from www.clomedia.com/business-intelligence-report/.

Employment Policy Foundation. (2002). Challenges facing the American workplace. *American Workplace Report* 2002. Washington, DC: Author.

Hanssen, C. E. (2007). TSS redesign 2007: Evaluation analysis and approach. Internal study performed for KPMG LLP Learning & Development.

Kirkpatrick, D. (1998). *Evaluating training programs: The four levels* (2nd ed.). San Francisco: Berrett-Koehler.

KnowledgeAdvisors. (2005). *Human capital contribution model*. Chicago, IL: KnowledgeAdvisors.

L'Allier, J. (2005, March). The CLO's role: Preparing for future challenges. *CLO* magazine. Retrieved June 5, 2008, from http://www.clomedia.com/content/templates/clo_article.asp?articleid = 869&zoneid = 59.

LiNE Zine (2001, Spring). *Talking human capital with professor Gary S. Becker*. Retrieved May 11, 2009, from http://www.linezine.com/7.1/interviews/gbbmthc.htm.

Lee, S. (2008, February). A look ahead: The training industry in 2008. *CLO* magazine. Retrieved June 5, 2008, from http://www.clomedia.com/includes/printcontent.php?aid = 2083.

Masie, E. (2006, September). Lies about e-learning. In L. Israelite (Ed.), *Lies about learning* (pp. 127–139). Alexandria, VA: American Society for Training and Development.

ABOUT THE EDITORS

Kenneth H. Silber, Ph.D., CPT, is president of Silber Performance Consulting and an Associate Professor Emeritus in instructional technology at Northern Illinois University. He is also currently an Adjunct Professor at Capella University, in the ID and HPT programs. For forty-five years, he has been working in corporate, academic, and consulting settings to produce results for clients. He designed and implemented HPT interventions (ID-based and non-training) that have saved millions of dollars. He has taught HPT/ID skills to over two thousand students and employees. Silber has edited ISPI's *From Training to Performance* series, co-authored two chapters for the third edition of ISPI's *Handbook of Human Performance Technology*, co-authored *Writing Training Materials That Work*, published sixty articles, and made over one hundred presentations, all of which are available on his website, www.silberperformance.com. He has earned his e-mail handle, wiseoldken@aol.com.

Wellesley R. (Rob) Foshay, Ph.D., CPT, is director-research for the Education Technology Group of Texas Instruments. Prior to joining TI, he served as a consultant and as the chief instructional architect of the PLATO learning system, developing and researching learning architectures for the system. Prior to joining PLATO, he directed instructional design and quality assurance for Applied Learning, Inc. His background includes a faculty position at the University of Illinois–Champaign, as well as experience as a high school teacher and school media coordinator. His doctorate is in instructional design from

Indiana University. He has contributed over seventy major articles to research journals and book chapters on a wide variety of topics in instructional design, technology, and education and speaks frequently before educators and trainers world-wide. He is co-author of the award-winning textbook, *Writing Training That Works: How to Teach Anyone to Do Anything: A Practical Guide for Trainers Based on Current Cognitive Psychology and Instructional Design Theory and Research*. He currently serves on the editorial boards of three research journals. He has served on the boards of directors of the International Society for Performance Improvement (ISPI) and the International Board of Standards for Training, Performance, and Instruction (IBSTPI), as well as on an ANSI/ASQ committee that developed standards for training and education under the ISO 9000 system. He was cited by ISPI with honorary life membership and a distinguished service award and is a Certified Performance Technologist. Many of his publications are available at www.foshay.org.

ABOUT THE CONTRIBUTORS

Jeanne Hites Anderson, Ed.D., CPT, is a professor at the Center for Information Media at St. Cloud State University. She has been working and consulting in the instructional design field for over thirty years, both in businesses and academia. She is a professor of information media at St. Cloud State University in St. Cloud, Minnesota, where she teaches instructional design and research courses. She has conducted many analyses using the methods described in her chapter in this book to design live and distributed training and education using a variety of media. She can be reached at Jeanne.Anderson@StCloudState.edu.

Dee H. Andrews is a member of the Senior Executive Service and is a senior scientist for the Warfighter Readiness Research Division, 711 HPW/RHA, Air Force Research Laboratory in Mesa, Arizona. He is the Air Force's principal scientific authority for training research. His responsibilities include sustaining technological superiority for training by planning and conducting theoretical and experimental studies. Dr. Andrews graduated from Brigham Young University with a degree in psychology. He has a master's degree in instructional design and development and a doctorate in instructional systems from Florida State University.

Dr. Andrews' research interests include training in distributed environments, instructor-operator station design, performance measurement, command and control, cost effectiveness, and decay and retention of higher-order cognitive skills. He has written more than seventy publications. Dr. Andrews is a certified

Level III acquisition professional in systems planning, research, development, and engineering. The International Society for Performance and Improvement and the American Society for Training and Development have certified him as a Human Performance Technologist. He has been elected to Fellow status in the Human Factors and Ergonomics Society, the American Psychological Association, the Royal Aeronautical Society of the United Kingdom, and the Air Force Research Laboratory.

Herbert H. Bell is a technical advisor for the Warfighter Readiness Research Division, 711 HPW/RHA, Air Force Research Laboratory in Mesa, Arizona. His responsibilities include the coordination of research programs to establish the scientific and technical foundation that enables development of personalized, continuous training and performance aiding for individuals and teams throughout their Air Force careers.

Dr. Bell was one of the leaders in the development of technologies and training methods for Distributed Mission Training/Distributed Mission Operations. He has worked on a variety of military, commercial, and academic programs involving both system design and training and has coordinated a number of research activities within the Air Force, the Department of Defense, and NATO. Dr. Bell received his Ph.D. in experimental psychology from Vanderbilt University and is a member of Human Factors and Ergonomics Society, IEEE, and the Psychonomics Society. He has received the NATO Research Technology Organization Scientific Achievement Award as a member of the Mission Training via Distributed Simulation Joint Task Group, the AFRL/HE Scientific and Technical Management Award, the Human Systems Center Commander's Trophy, and the Federal Laboratory Consortium Award of Merit for Excellence in Technology Transfer.

Conrad G. Bills is a systems and architecture lead for training system acquisition, Lockheed Martin Maritime Simulation and Training Systems, Akron, Ohio, integrating systems engineering with instructional systems development (ISD). Prior responsibilities include serving on multiple training system programs covering all phases of the ISD process. He was responsible for the F-16 MTC pilot training system concept, and his understanding of training systems was incorporated into the training concept for the Joint Strike Fighter (JSF) Program. He joined Lockheed Martin in 1993 upon his retirement from the U.S. Air Force. Before his retirement, he directed the project updating the Air Force's ISD process. He also served as training systems analyst for Air Force Aircrew Training System acquisitions. He holds a Ph.D. from Kent State University and master's and bachelor's degrees from Brigham Young University. Dr. Bills has over thirty years of experience in scientific management, applied psychology, and instructional system development integrated with systems engineering. He is a member of the Air Force Association (AFA) and the International Society for Performance Improvement (ISPI) and is a Certified Performance Technologist (CPT).

Kerry J. Burner, Ph.D., works part-time on the faculty in the Richard W. Riley College of Education and Leadership at Walden University. She holds a Ph.D. in instructional systems from Florida State University, along with certificates in human performance technology, program evaluation, human resource management, and online instructional development. Currently working in the educational leadership Ed.D. program at Walden University, Dr. Burner recently rolled off the FSU-Teach program as the launch project director. Her diverse interests are not only reflected in her professional work: her publications and research projects range from investigating reflexivity in online learning, evaluating the role of PC-tablets in a middle school, and more recently studying cognitive apprenticeships, particularly in communities of practice. She can be reached at kerry.burner@waldenu.edu.

Richard E. Clark is a professor of educational psychology and technology in the Rossier School of Education at the University of Southern California, where he is co-director of the Center for Cognitive Technology. He is the author of over two hundred published articles and book chapters, as well as three recent books. He has received the Thomas F. Gilbert distinguished professional achievement award and a Presidential Citation for Intellectual Leadership from ISPI, and in 2007 he received the Outstanding Civilian Service Award from the U.S. Army for his work in distance learning.

Ruth Colvin Clark has worked for more than twenty-five years with instructional professionals who design, develop, and select effective workforce instructional programs for classroom or computer delivery. She has authored or co-authored six books, including the best-selling *e-Learning and the Science of Instruction*, co-authored with Richard E. Mayer. Dr. Clark and her associates provide seminars, short programs, and consulting to organizations interested in applying evidence-based practice to workforce learning environments.

William C. Coscarelli is a Professor Emeritus in Southern Illinois University at Carbondale's Department of Curriculum. He has been elected president of the International Society for Performance Improvement and the Association for Educational Communications and Technology's Division for Instructional Development. He was the founding editor of *Performance Improvement Quarterly*. He is also the co-author of *Criterion-Referenced Test Development for Corporate Training: Technical and Legal Issues*. He is also an author of *The Decision-Making Style Inventory* and *The Guided Design Guidebook*. He was recently selected as a co-author for the *International Encyclopedia of Education*'s review of the literature for criterion-referenced test development.

Deborah Singer Dobson, M.Ed., Dobson Solutions, is a human resources/organizational development executive and the co-author of three management books. She is currently AVP/director of leadership organizational development for SAIC. Previously, she was senior vice president/management development, training, and staffing, US Foodservice/Ahold USA, and vice president/HR for GATX Terminals Corporation, Chicago. She helped found and co-managed ERIS Enterprises, a management development consulting firm in Columbia, Maryland.

Michael Singer Dobson, PMP, Dobson Solutions, is the author of twenty books, including six books on project management. He speaks and trains internationally on management topics. He was vice president, Discovery Software International; vice president/marketing, Games Workshop (U.S.); director of marketing and new product development, TSR, Inc.; is president, Michael Dobson & Associates, a Washington, D.C., career counseling firm; and is part of the professional team that built the Smithsonian Institution's National Air and Space Museum. He is on the board of the Samaritan Medal Foundation, the only Samaritan Medal awarded by the authentic Israelite Samaritan people.

Sean Early graduated with a B.A. in political science/international relations from the University of California, Santa Barbara, where he helped lead the Global Peace and Security Program. He received his M.A. in communication disorders at California State University, Los Angeles, and his Ph.D. in educational psychology at the University of Southern California. Dr. Early's research interests center on instructional design, immersive media-based learning environments, and performance measurement.

Peter J. Fadde, Ph.D., is an assistant professor, instructional technology and instructional design, Southern Illinois University (http://web.coehs.siu.edu/units/ci/faculty/pfadde/). In his pre-academic career, Dr. Fadde served for thirteen years as video coordinator for a major college athletic program, where he developed video and computer applications for game analysis by football coaches and decision training by players. Dr. Fadde has since explored the adaptation of sports-inspired simulation training approaches in a wide range of performance domains, including semi-truck driving.

Kevin Forsberg, Ph.D., CSEP, is co-founder and chief executive officer of The Center for Systems Management. Dr. Forsberg has fifty-one years of experience in systems engineering, program and proposal management, and consulting to government and international businesses. Dr. Forsberg co-authored two books on project management and has developed and delivered many project management and systems engineering courses. His experience spans research, development, and production implementation. Dr. Forsberg is the recipient of distinguished achievement awards, including the CIA Agency Seal Medallion, the INCOSE Pioneer award, and the NASA Public Service Medal. He is co-founder of a nationally recognized university certificate in project management program. In addition, he has published over sixty journal articles and presented papers at PMI symposiums, INCOSE, ProjectWorld, and AIAA. He is an INCOSE fellow and a lifetime ASME fellow.

David H. Jonassen is a Distinguished Professor of Education at the University of Missouri, where he teaches in the areas of learning technologies and educational psychology. Since earning his doctorate in educational media and experimental educational psychology from Temple University, Dr. Jonassen has taught at the Pennsylvania State University, University of Colorado, the University of Twente in the Netherlands, the University of North Carolina at Greensboro, and Syracuse University. He has published thirty books and numerous articles, papers, and reports on text design, task analysis, instructional design, computer-based learning, hypermedia, constructivist learning, cognitive tools, and problem solving. His current research focuses on the cognitive processes engaged by problem solving and models and methods for supporting those processes during learning.

Ted Leemann, The Center for Systems Management, has over twenty-five years of management experience in the public and private sectors with particular emphasis on project management and strategic planning. He has developed and presented training internationally to thousands of students in the telecommunications, automated information services industries, and governments. Speaking engagements include the Software Technology Conference, ISPI, ASTD, The Maryland Techway Conference, and the American Institute of Architects Annual Symposium. He has also published a number of articles in key trade publications, *SBusiness Strategy, PM News Letter*, and has been interviewed for United Airlines Sky Radio. He earned a bachelor's in business and economics and a master's in logistics management at the Florida Institute of Technology.

Heather C. Maitre is director of technology-based learning and performance measurement at KPMG, which she joined in 2001 to pioneer e-learning. Her responsibilities were later expanded to include a learning measurement function startup. KPMG has since been externally recognized in the top five in *Training* magazine's top 125 rankings, received four eLearning awards and two learning measurement awards. Prior to working for KPMG, she was an information technology director at Columbia University, where she managed early e-learning experiments and was responsible for the complete technology infrastructure. She graduated with a bachelor's degree in business from Pace University and holds a master's from Columbia University in industrial/organizational psychology.

Richard E. Mayer is an internationally recognized expert in the application of learning psychology to design of instruction in multimedia learning environments. He has published dozens of research studies and is the author of *Multimedia Learning* and the editor of the *Cambridge Handbook of Multimedia Learning*.

Michael Molenda, Ph.D., is an associate Professor Emeritus, instructional systems technology, Indiana University. He is co-author of the first five editions of *Instructional Media and the New Technologies of Instruction*, an award-winning textbook on incorporating media systematically into instruction. He has written numerous journal articles, encyclopedia entries, and book chapters on instructional design/development and co-taught the first course on instructional development at Indiana University in 1974. He has lectured and consulted in Spain, Indonesia, Korea, China, and several countries in Latin America and the Middle East. In retirement he continues to do research and writing on educational technology, particularly on historical aspects.

Kathrine Moulton is a Ph.D. student in the Rossier School of Education at the University of Southern California. She earned her B.A. in psychology, an M.A. in school counseling, and second M.A. in clinical psychology. She has professional experience as a clinical therapist and school counselor. Her interests include the effects of cognitive and socio-cultural factors on motivation, the use of technology in urban schools, and the role of peer influence on the academic outcomes of urban youth.

Gary R. Morrison is a professor and graduate program director in the instructional design and technology program at Old Dominion University. He is author of two books and over sixty book chapters and articles on instructional design and educational technology. He is the editor of the *Journal of Computing in Higher Education* and is on the editorial boards of *Computers in Human Behavior, Quarterly Review of Distance Education*, and Libraries Unlimited's *Instructional Technology Serie*s. He has worked as instructional designer for three Fortune 500 companies and the University of Mid-America. He is the past-president of AECT's Design and Development, Research and Theory, and Distance Learning Divisions.

Jane Ranshaw is president of Jane Ranshaw & Associates, Inc. (www.ranshawconsulting.com), has taught thousands of businesspeople to communicate better through her workshops on business writing, internal consulting, and managing emotions and in seminars for the American Management Association. She is a member of the adjunct faculty, DePaul University. She is president of the University of Chicago's Women's Business Group and was, for several years, co-editor of *Training Today*, the award-winning magazine of the Chicago chapter, American Society for Training and Development. She has chaired the Business Issues track for ISPI conferences. Ranshaw completed her MBA at the University of Chicago and her B.S. degree in business at Indiana University.

Steven M. Ross is Faudree Professor and executive director of the Center for Research in Educational Policy at the University of Memphis. Dr. Ross is the author of six textbooks and over 125 journal articles in the areas of educational technology and instructional design, at-risk learners, educational reform, computer-based instruction, and individualized instruction. He is the editor of the research section of the *Educational Technology Research and Development* journal and a member of the editorial board for two other professional journals. In 1993, he was the first faculty recipient of the University of Memphis Eminent Faculty Award for teaching, research, and service and recently held the Lillian and Morrie Moss Chair of Excellence in Urban Education.

Sharon A. Shrock, Ph.D., is professor of instructional development and technology at Southern Illinois University at Carbondale, where she teaches graduate courses in evaluation and advanced research methods and coordinates graduate programs in instructional design and technology. She has consulted with a number of global corporations in developing criterion-referenced assessment systems and co-authored the award-winning book, *Criterion-Referenced Test Development: Technical and Legal Guidelines for Corporate Training*. She has been a senior advisor to the International Society for Performance Improvement (ISPI). She received both her B.S. degree in education and economics and her Ph.D. in instructional systems technology from Indiana University.

Susan A. Smith is director, learning operations, at KPMG LLP, Montvale, New Jersey. She started with KPMG in 1994 and was responsible for developing and delivering consulting training initiatives and key global services programs. In 2000, Susan was asked to establish a firm-wide learning operations team. Her team is instrumental in the successful delivery of over 2,500 training offerings each year. These efforts have contributed to KPMG's top five ranking on *Training* magazine's Top 125 list for the last two years. Her past experience includes financial management at American Express and AMRO Bank and non-profit project management at the Fund for the City of New York.

Kurt D. Squire is an associate professor, educational communications and technology, research scientist, ADL Colab director, games + learning + society initiative at the University of Wisconsin-Madison. He is a former Montessori and primary teacher with an interest in design-based research on digital media for social change. He is the author of over fifty scholarly articles and publications and the recipient of major grants from the National Science Foundation, the MacArthur Foundation, and the Department of Education. His current projects are investigating the potential of science games to support students in scientific citizenship and mobile media to increase civic engagement.

Since 1983, *Deborah L. Stone*, CPT, has been the CEO of DLS Group, Inc. The co-recipient of over twenty professional awards, including Microsoft's Award of Excellence, her interests include capturing expertise and improving performance by leveraging blended technologies. A frequent conference presenter, she has written numerous articles and book chapters, including ''Performance Support Systems'' in ISPI's *Handbook of Human Performance Technology* and ''Blended Learning Strategies'' in *The Trainer's Portable Mentor*. She served on ISPI's board of directors and ASTD's planning advisory board (ICE). She was awarded a fellowship/grant to Brown University and completed graduate work in instructional technology at San Francisco State University.

Steven W. Villachica, Ph.D., CPT, is an associate professor of instructional and performance technology at Boise State University. He is also a member of DLS Group, Inc.'s Strategic Advisory Board. His interests include investigating employer expectations of entry-level instructional designers and identifying and leveraging expertise across organizations. He has worked in consulting, corporate, and academic settings for thirty years. A frequent conference presenter, he has co-authored chapters on performance support systems that appear in two editions of ISPI's *Handbook of Human Performance Technology*. A two-time winner of ISPI's Outstanding Systematic Approach award, he completed his doctorate in educational technology at the University of Northern Colorado.

Wallace H. Wulfeck II, Ph.D., is a senior research psychologist at the Space and Naval Warfare Systems Center in San Diego, California, where he serves as co-principal investigator and project scientist on the Interactive Multisensor Analysis Training project. Throughout 2004 to 2007, he served as deputy science and technology officer for the Capable Manpower Future Naval Capability program at the Office of Naval Research. Dr. Wulfeck has a continuing record of training research, technology development, and transition to operational use. During the 1990s, he directed the Instructional Simulations Division and Training Research Computing Facility at the Navy Personnel Research and

Development Center. Dr. Wulfeck has also served as science advisor to the Chief of Naval Personnel. He is the author of over one hundred journal articles, book chapters, technical reports, and conference presentations. Dr. Wulfeck received his Ph.D. in learning from the University of Pennsylvania, following undergraduate and graduate degrees in math and math education from the University of California, Santa Barbara. He is a member of the American Psychological Society, the American Educational Research Association, and the Cognitive Science Society. He has received the Navy Meritorious Civilian Service Award.

Kenneth Yates is a senior research associate at the Center for Cognitive Technology in the Rossier School of Education at the University of Southern California, where he develops training and evaluation programs and conducts research in, and applications of, cognitive task analysis methods to improve human performance, instructional design, and educational technology. Recently, he developed training and evaluation programs for the U.S. Army, conducted cognitive task analysis and designed training for counterinsurgency operations in Iraq and Afghanistan, and designed training for a U.S. Army course in bilateral negotiations for the USC Institute for Creative Technologies.

NAME INDEX

T

SUBJECT INDEX

Page references followed by *fig* indicate an illustrated figure; followed by *t* indicate a table; followed by *e* indicate an exhibit.

Q

R

Pfeiffer Publications Guide

This guide is designed to familiarize you with the various types of Pfeiffer publications. The formats section describes the various types of products that we publish; the methodologies section describes the many different ways that content might be provided within a product. We also provide a list of the topic areas in which we publish.

FORMATS

In addition to its extensive book-publishing program, Pfeiffer offers content in an array of formats, from fieldbooks for the practitioner to complete, ready-to-use training packages that support group learning.

FIELDBOOK Designed to provide information and guidance to practitioners in the midst of action. Most fieldbooks are companions to another, sometimes earlier, work, from which its ideas are derived; the fieldbook makes practical what was theoretical in the original text. Fieldbooks can certainly be read from cover to cover. More likely, though, you'll find yourself bouncing around following a particular theme, or dipping in as the mood, and the situation, dictate.

HANDBOOK A contributed volume of work on a single topic, comprising an eclectic mix of ideas, case studies, and best practices sourced by practitioners and experts in the field.

An editor or team of editors usually is appointed to seek out contributors and to evaluate content for relevance to the topic. Think of a handbook not as a ready-to-eat meal, but as a cookbook of ingredients that enables you to create the most fitting experience for the occasion.

RESOURCE Materials designed to support group learning. They come in many forms: a complete, ready-to-use exercise (such as a game); a comprehensive resource on one topic (such as conflict management) containing a variety of methods and approaches; or a collection of like-minded activities (such as icebreakers) on multiple subjects and situations.

TRAINING PACKAGE An entire, ready-to-use learning program that focuses on a particular topic or skill. All packages comprise a guide for the facilitator/trainer and a workbook for the participants. Some packages are supported with additional media—such as video—or learning aids, instruments, or other devices to help participants understand concepts or practice and develop skills.

- *Facilitator/trainer's guide* Contains an introduction to the program, advice on how to organize and facilitate the learning event, and step-by-step instructor notes. The guide also contains copies of presentation materials—handouts, presentations, and overhead designs, for example—used in the program.

- *Participant's workbook* Contains exercises and reading materials that support the learning goal and serves as a valuable reference and support guide for participants in the weeks and months that follow the learning event. Typically, each participant will require his or her own workbook.

ELECTRONIC CD-ROMs and web-based products transform static Pfeiffer content into dynamic, interactive experiences. Designed to take advantage of the searchability, automation, and ease-of-use that technology provides, our e-products bring convenience and immediate accessibility to your workspace.

METHODOLOGIES

CASE STUDY A presentation, in narrative form, of an actual event that has occurred inside an organization. Case studies are not prescriptive, nor are they used to prove a point; they are designed to develop critical analysis and decision-making skills. A case study has a specific time frame, specifies a sequence of events, is narrative in structure, and contains a plot structure—an issue (what should be/have been done?). Use case studies when the goal is to enable participants to apply previously learned theories to the circumstances in the case, decide what is pertinent, identify the real issues, decide what should have been done, and develop a plan of action.

ENERGIZER A short activity that develops readiness for the next session or learning event. Energizers are most commonly used after a break or lunch to stimulate or refocus the group. Many involve some form of physical activity, so they are a useful way to counter post-lunch lethargy. Other uses include transitioning from one topic to another, where "mental" distancing is important.

EXPERIENTIAL LEARNING ACTIVITY (ELA) A facilitator-led intervention that moves participants through the learning cycle from experience to application (also known as a Structured Experience). ELAs are carefully thought-out designs in which there is a definite learning purpose and intended outcome. Each step—everything that participants do during the activity—facilitates the accomplishment of the stated goal. Each ELA includes complete instructions for facilitating the intervention and a clear statement of goals, suggested group size and timing, materials required, an explanation of the process, and, where appropriate, possible variations to the activity. (For more detail on Experiential Learning Activities, see the Introduction to the *Reference Guide to Handbooks and Annuals*, 1999 edition, Pfeiffer, San Francisco.)

GAME A group activity that has the purpose of fostering team spirit and togetherness in addition to the achievement of a pre-stated goal. Usually contrived—undertaking a desert expedition, for example—this type of learning method offers an engaging means for participants to demonstrate and practice business and interpersonal skills. Games are effective for team building and personal development mainly because the goal is subordinate to the process—the means through which participants reach decisions, collaborate, communicate, and generate trust and understanding. Games often engage teams in "friendly" competition.

ICEBREAKER A (usually) short activity designed to help participants overcome initial anxiety in a training session and/or to acquaint the participants with one another. An icebreaker can be a fun activity or can be tied to specific topics or training goals. While a useful tool in itself, the icebreaker comes into its own in situations where tension or resistance exists within a group.

INSTRUMENT A device used to assess, appraise, evaluate, describe, classify, and summarize various aspects of human behavior. The term used to describe an instrument depends primarily on its format and purpose. These terms include survey, questionnaire, inventory, diagnostic, survey, and poll. Some uses of instruments include providing instrumental feedback to group members, studying here-and-now processes or functioning within a group, manipulating group composition, and evaluating outcomes of training and other interventions.

Instruments are popular in the training and HR field because, in general, more growth can occur if an individual is provided with a method for focusing specifically on his or her own behavior. Instruments also are used to obtain information that will serve as a basis for change and to assist in workforce planning efforts.

Paper-and-pencil tests still dominate the instrument landscape with a typical package comprising a facilitator's guide, which offers advice on administering the instrument and interpreting the collected data, and an initial set of instruments. Additional instruments are available separately. Pfeiffer, though, is investing heavily in e-instruments. Electronic instrumentation provides effortless distribution and, for larger groups particularly, offers advantages over paper-and-pencil tests in the time it takes to analyze data and provide feedback.

LECTURETTE A short talk that provides an explanation of a principle, model, or process that is pertinent to the participants' current learning needs. A lecturette is intended to establish a common language bond between the trainer and the participants by providing a mutual frame of reference. Use a lecturette as an introduction to a group activity or event, as an interjection during an event, or as a handout.

MODEL A graphic depiction of a system or process and the relationship among its elements. Models provide a frame of reference and something more tangible, and more easily remembered, than a verbal explanation. They also give participants something to "go on," enabling them to track their own progress as they experience the dynamics, processes, and relationships being depicted in the model.

ROLE PLAY A technique in which people assume a role in a situation/scenario: a customer service rep in an angry-customer exchange, for example. The way in which the role is approached is then discussed and feedback is offered. The role play is often repeated using a different approach and/or incorporating changes made based on feedback received. In other words, role playing is a spontaneous interaction involving realistic behavior under artificial (and safe) conditions.

SIMULATION A methodology for understanding the interrelationships among components of a system or process. Simulations differ from games in that they test or use a model that depicts or mirrors some aspect of reality in form, if not necessarily in content. Learning occurs by studying the effects of change on one or more factors of the model. Simulations are commonly used to test hypotheses about what happens in a system—often referred to as "what if?" analysis—or to examine best-case/worst-case scenarios.

THEORY A presentation of an idea from a conjectural perspective. Theories are useful because they encourage us to examine behavior and phenomena through a different lens.

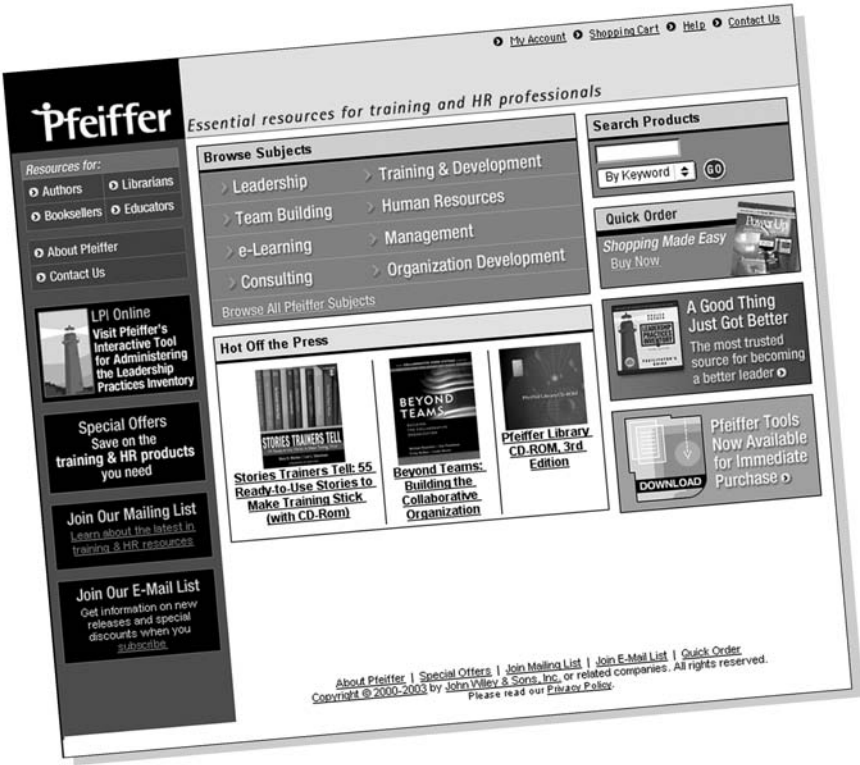
TOPICS

The twin goals of providing effective and practical solutions for workforce training and organization development and meeting the educational needs of training and human resource professionals shape Pfeiffer's publishing program. Core topics include the following:

Leadership & Management
Communication & Presentation
Coaching & Mentoring
Training & Development
E-Learning
Teams & Collaboration
OD & Strategic Planning
Human Resources
Consulting

What will you find on pfeiffer.com?

- The best in workplace performance solutions for training and HR professionals
- Downloadable training tools, exercises, and content
- Web-exclusive offers
- Training tips, articles, and news
- Seamless on-line ordering
- Author guidelines, information on becoming a Pfeiffer partner, and much more



Discover more at www.pfeiffer.com

HANDBOOK of

IMPROVING PERFORMANCE IN THE WORKPLACE

Volume 1: Instructional Design and Training Delivery

With the contributions from leading national and international scholars and practitioners, this volume provides a "state-of-the-art" look at ID, addressing the major changes that have occurred in nearly every aspect of ID in the past decade and provides both theory and "how-to" information for ID and performance improvement practitioners who must stay current in their field.

This volume goes beyond other ID references in its approach: it is useful to students and practitioners at all levels; it is grounded in the most current research and theory; and it provides up-to-the-minute coverage of topics not found in any other ID book. It addresses timely topics such as cognitive task analysis, instructional strategies based on cognitive research, data collection methods, games, higher-order problem-solving and expertise, psychomotor learning, project management, partnering with clients, and managing a training function. It also provides a new way of looking at what ID is, and the most comprehensive history of ID ever published.

Sponsored by International Society for Performance Improvement (ISPI), the *Handbook of Improving Performance in the Workplace*, three-volume reference, covers three core areas of interest including Instructional Design and Training Delivery, Selecting and Implementing Performance Interventions, and Measurement and Evaluation.

THE EDITORS

Kenneth H. Silber is president of Silber Performance Consulting, associate professor emeritus from Northern Illinois University and adjunct professor at Capella University. He has been contributing to the performance consulting and instructional design fields since their inception forty years ago. He has co-authored several books, including the *Instructional Design Competencies (IBSTPI)*, and *Writing Training Materials That Work* (Pfeiffer).

Wellesley (Rob) Foshay is director of research for the Education Technology Group of Texas Instruments. Through The Foshay Group, he speaks and publishes frequently and serves on the editorial board of four research journals in educational technology. He is a founding member of IBSTPI, and co-authored their *Instructional Design Competencies*. He served on the Board of Directors of ISPI, contributed to the creation of its certification program, and was honored with the Society's Member for Life and Distinguished Service awards. He has contributed over 70 major articles to research journal and books on a wide variety of topics in training, technology and education and is the co-author of *Writing Training Materials That Work* (Pfeiffer).

Join Us at
Pfeiffer.com

Register at **www.pfeiffer.com/email** for more information on our publications, authors, and to receive special offers.

Pfeiffer®
An Imprint of
WILEY

www.pfeiffer.com

TRAINING AND DEVELOPMENT

ISBN 978-0-470-19068-5
90000
9 780470 190685