

TENTH EDITION

PRACTICAL RESEARCH

Planning and Design

Paul D. Leedy
Jeanne Ellis Ormrod

Practical Research

PLANNING AND DESIGN

TENTH EDITION

Paul D. Leedy

Late of American University

Jeanne Ellis Ormrod

University of Northern Colorado (Emerita)

PEARSON

Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montreal Toronto
Delhi Mexico City São Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

Vice President and Editorial Director: Jeffery W. Johnston
Vice President and Publisher: Kevin Davis
Editorial Assistant: Lauren Carlson
Vice President, Director of Marketing: Margaret Waples
Senior Marketing Manager: Joanna Sabella
Senior Managing Editor: Pamela Bennett
Senior Project Manager: Mary M. Irvin
Senior Operations Supervisor: Matt Ottenweller
Senior Art Director: Diane Lorenzo
Cover Designer: Brian Huber
Cover Art: Leon Zernitsky
Project Coordination: Lynda Griffiths, Jouve India Pvt Ltd
Composition: Jouve India Pvt Ltd
Printer/Binder: Courier Kendallville
Cover Printer: Lehigh-Phoenix Color Corp.
Text Font: Garamond 3 LT Std

Credits and acknowledgments for materials borrowed from other sources and reproduced, with permission, in this textbook appear on the appropriate page within text.

Every effort has been made to provide accurate and current Internet information in this book. However, the Internet and information posted on it are constantly changing, so it is inevitable that some of the Internet addresses listed in this textbook will change.

Copyright © 2013, 2010, 2005, 2001, 1997 by Pearson Education, Inc. All rights reserved. Printed in the United States of America. This publication is protected by Copyright and permission should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. To obtain permission(s) to use material from this work, please submit a written request to Pearson Education, Inc., Permissions Department, One Lake Street, Upper Saddle River, New Jersey 07458 or you may fax your request to 201-236-3290.

Library of Congress Cataloging-in-Publication Data

Leedy, Paul D.

Practical research : planning and design / Paul D. Leedy, Jeanne Ellis Ormrod.—10th ed.
p. cm.

Includes bibliographical references and index.

ISBN-13: 978-0-13-269324-0

ISBN-10: 0-13-269324-0

1. Research—Methodology. I. Ormrod, Jeanne Ellis. II. Title.
Q180.55.M4L43 2013
001.4—dc23

2011037270

10 9 8 7 6

PEARSON

ISBN 10: 0-13-269324-0
ISBN 13: 978-0-13-269324-0

Preface

New to the Tenth Edition

It is always a joy to write a new edition of *Practical Research*, in part because there are always so many exciting new things to say! This time around, the book has been revised and updated in numerous ways. Some changes are subtle ones—wording changes to tighten the prose or enhance clarity; updates of Internet addresses (URLs), software packages, and suggested readings; minor tweaks to tables and figures to enhance their effectiveness; and so on. Two more noticeable changes are (1) a new first chapter that combines and condenses the first two chapters of the previous edition, which makes room for (2) a new chapter on mixed-methods research. Another significant change is the addition of a new Conceptual Analysis Exercise feature in which readers can apply chapter content to new situations and scenarios. And given the increasing technological sophistication of our readers, a few topics included in previous editions have been deleted, such as how to get an e-mail account and send e-mail messages.

More specific changes in the tenth edition are as follows:

- **Chapter 1 (formerly Chapters 1 and 2):** An integration of the ninth edition's first two chapters into a single chapter called "The Nature and Tools of Research," with some of the "old" second chapter's contents being moved to later chapters; a new figure on common pitfalls in reasoning (Figure 1.3).
- **Chapter 2 (formerly Chapter 3):** A discussion of independent and dependent variables (which, in the ninth edition, appeared much later in the book) along with a new discussion of mediating and moderating variables; a new Conceptual Analysis Exercise on identifying different kinds of variables.
- **Chapter 3 (formerly Chapter 4):** Significantly revised and updated discussions of library resources (including significant revisions to Tables 3.1 and 3.3).
- **Chapter 4 (formerly Chapter 5):** A reorganization of chapter content to be more logical; two new Conceptual Analysis Exercises, "Identifying Scales of Measurement" and "Identifying Problems with Validity and Reliability in Measurement"; a significantly expanded discussion of research ethics; a new Checklist, "Determining Whether Your Proposed Study Is Ethically Defensible."
- **Chapter 5 (formerly Chapter 6):** A revised discussion of APA style, consistent with the sixth edition of APA's *Publication Manual* (2010); a significantly expanded discussion of common weaknesses in research proposals (Figure 5.3).
- **Chapter 6 (formerly Chapter 7):** An expanded discussion of sampling in qualitative research; a new Conceptual Analysis Exercise, "Choosing a Qualitative Research Design."
- **Chapter 7 (formerly Chapter 8):** Inclusion of comparative-historical research as a methodology in the social sciences; a significant reduction in the length of the chapter (in large part in response to most reviewers' feedback that they didn't assign this chapter).
- **Chapter 8 (formerly Chapter 9):** A discussion of mediating and moderating variables within the context of longitudinal designs.

- **Chapter 9 (formerly Chapter 10):** A new figure presenting Campbell and Stanley's (1963) classic list of threats to internal validity (Figure 9.1); an expanded discussion of conducting experimental studies on the Internet; a new Checklist, "Looking for Confounding Variables."
- **Chapter 10 (new chapter):** An overview of various mixed-methods designs; attention to issues specific to mixed-methods research; a new Checklist, "Pinning Down the Logistics and Feasibility of a Mixed-Methods Study."
- **Chapter 11:** A significant expansion of Table 11.5 ("Description of Inferential Statistical Procedures and Their Purposes"); a description of NCEs as an increasingly common form of standard score (NCEs are described only in a footnote, as their use seems to be restricted largely to K–12 education).
- **Chapter 12:** Expanded discussions of writing abstracts and using tables and figures to present research findings; an explanation of digital object identifiers (DOIs) as a recent innovation in referencing online resources; a new section on ethics, titled "Maintaining Your Academic Integrity" (note that ethics are covered in greater depth in Chapter 4); a significantly expanded discussion of conference presentations, which now includes poster sessions as well as paper sessions; a new Guidelines feature, "Presenting an Effective Paper or Poster."
- **Appendixes A and B:** Updated discussions of Microsoft Excel (Appendix A) and SPSS (Appendix B) to be consistent with recent changes in these software packages.

Another important change is a significantly enhanced online accompaniment to the book known as MyResearchLab. When students purchase the tenth edition, they also get an access code with which they can set up a personal account for use on this invaluable supplementary learning tool.

The Purpose of This Book

Practical Research: Planning and Design is a broad-spectrum, cross-disciplinary book suitable for a wide variety of courses in basic research methodology. Many basic concepts and strategies in research transcend the boundaries of specific academic areas, and such concepts and strategies are at the heart of this book. To some degree, certainly, research methods do vary from one subject area to another: A biologist might gather data by looking through a microscope, a historian by examining written documents from an earlier time period, and a psychologist by administering tests or systematically observing people's behavior. Otherwise, the basic approach to research is the same. Regardless of the discipline, the researcher identifies a question in need of an answer, collects data potentially relevant to that answer, analyzes and interprets the data, and draws conclusions that the data seem to warrant.

Students in the social sciences, the natural sciences, education, medicine, business administration, landscape architecture, real estate, and other academic disciplines have used this text as a guide to the successful completion of their research projects. *Practical Research* guides students from problem selection to completed research report with many concrete examples and practical, how-to suggestions. Students come to understand that research needs planning and design, and they discover how their own research projects can be executed effectively and professionally. Essentially, this is a do-it-yourself, understand-it-yourself manual. From that standpoint, it can be a guide for students who are left largely to their own resources in carrying out their research projects. The book, supplemented by occasional counseling by an academic advisor, can guide the student to the completion of a successful research project.

Learning about the Research Process Is an Essential Component of Academic Training

All too often, students mistakenly believe that conducting research involves nothing more than amassing a large number of facts and incorporating them into a lengthy, footnoted paper. They reach the threshold of a master's thesis or doctoral dissertation only to learn that simply assembling previously known information is insufficient and unacceptable. Instead, they must do

something radically different: They must answer a question that has never been answered before and, in the process, must discover something that no one else has ever discovered. Something has gone tragically wrong in the education of students who have, for so many years of their schooling, entirely misunderstood the true nature of research.

Research has one end: the discovery of truth. Its purpose is to learn what has never been known before; to ask a significant question for which no conclusive answer has previously been found; and, by collecting and interpreting relevant data, to find an answer to that question.

Learning about and doing research are of value far beyond that of merely satisfying a program requirement. Research methods and their application to real-world problems are skills that will serve you for the rest of your life. The world is full of problems that beg for solutions; consequently it is full of research activity! The media continually bring us news of previously unknown biological and physical phenomena, life-saving medical interventions, and ground-breaking technological innovations—all the outcomes of research. Research is not an academic banality; it is a vital and dynamic force that is indispensable to the health and well-being of Planet Earth and its human and nonhuman inhabitants.

More immediate, however, is the need to apply research methodology to those lesser daily problems that nonetheless demand a thoughtful resolution. Those who have learned how to analyze problems systematically and dispassionately will live with greater confidence and success than those who have shortsightedly dismissed research as nothing more than a necessary hurdle on the way to a degree. Given the advantages that a researcher's viewpoint provides, considering an academic research requirement as annoying and irrelevant to one's education is simply untenable.

Many students have found that *Practical Research* has helped them both understand the nature of the research process and complete their research projects. Its simplification of research concepts and its readability make it especially suitable for those undergraduate and graduate students who are introduced, perhaps for the first time, to genuine research methodology.

We hope we have convinced you that research methodology is not a temporary hurdle on the way to a degree but, instead, an unparalleled opportunity to learn how you might better tackle any problem for which you do not have a ready solution. In a few years you will undoubtedly look back on your course in research methodology as one of the most rewarding and practical courses in your entire educational experience. We have found it so in our own experience, and many of our students have told us likewise.

Acknowledgments

No man is an iland, entire of it selfe; every man is a peece of the Continent, a part of the maine. . .

So wrote John Donne, the great dean of St. Paul's Cathedral in the 17th century. And so write we authors at the beginning of the 21st.

Those who have had a part in the making of this book, known and unknown, friends and colleagues, gentle critics and able editors—all—are far too many to salute individually. Those of you who have written in journals and textbooks about research methods and strategies, the generations of graduate and undergraduate students whom we have taught and who have also taught *us*, the kindly letters and e-mails that so many of you have written to tell us how this book has helped you in your own research endeavors—to all of you, we extend our acknowledgment and appreciation wherever you may be. You have had the greater part in bringing this book through its previous nine editions. We are especially grateful to the reviewers of the ninth edition, who offered many invaluable suggestions for strengthening the tenth edition so that it can better assist novice researchers in the 21st century: Robert Balicki, Cleary University; Ronald S. Beebe, University of Houston–Downtown; John W. Hangsleben, Fontbonne University; Chris Lüer, Ball State University; Susan Moscou, Mercy College; Angela Owusu-Ansah, Samford University; Doris Prater, University of Houston; Helen Stiff-Williams, Regent University; and Edyth Wheeler, Towson University.

We are also indebted to the students whose research proposals, doctoral dissertations, and master's theses have enabled us to illustrate some of the research and writing strategies described in the book. In particular, we extend gratitude to Rosenna Bakari, Kay Corbett, Dinah Jackson, Ginny Kinnick, Laura Lara-Brady, Peter Leavenworth, Matthew McKenzie, Kimberly Mitchell, Richard Ormrod, Luis Ramirez, Janie Shaklee, Nancy Thrailkill, and Debby Zambo. Pete Leavenworth and Matt McKenzie gave us their time as well as their research reports, and their recommendations for the chapter on historical research were superb.

Equally important is to say “Thank you, thank you, thank you” to folks at Pearson Education and Jouve India who were key players in bringing this book to fruition. Early in the process, Shannon Steed sifted through the many reviewer comments and offered helpful suggestions about how to improve the book in numerous ways, both small and more substantial. Mary Irvin devotedly and dependably oversaw production at the Pearson end, and Lynda Griffiths doggedly addressed both the big-picture issues and the nitty-gritty details during copy editing and final paging. Lynda's patience and good humor throughout many unforeseen circumstances during the spring and summer of 2011 were very much appreciated.

Finally, we must thank Paul Smith and Kevin Davis, former and current editors, respectively, for the book during the writing and production of this tenth edition. Both Kevin and Paul shared our vision for the book and struck the ever-so-important balance between providing guidance to help us improve the book while also trusting our instincts about how best to explain and illustrate the complex, multifaceted nature of research planning and design.

No author is an island, entire of itself. Every author has had many hands guiding his or her pen and many minds adding richness and depth to his or her own thoughts. All of you have been exceedingly helpful, all of you have been "a peece of the Continent, a part of the maine." For that, we offer our humble and hearty thanks.

Jeanne Ellis Ormrod

Brief Contents

PART I	The Fundamentals	
	CHAPTER 1 The Nature and Tools of Research	1
PART II	Focusing Your Research Efforts	
	CHAPTER 2 The Problem: The Heart of the Research Process	27
	CHAPTER 3 Review of the Related Literature	51
	CHAPTER 4 Planning Your Research Project	74
	CHAPTER 5 Writing the Research Proposal	119
PART III	Qualitative Research Methodologies	
	CHAPTER 6 Qualitative Research	139
	CHAPTER 7 Historical Research	170
PART IV	Quantitative Research Methodologies	
	CHAPTER 8 Descriptive Research	184
	CHAPTER 9 Experimental, Quasi-Experimental, and Ex Post Facto Designs	226
	CHAPTER 10 Mixed-Methods Research	258
	CHAPTER 11 Strategies for Analyzing Quantitative Data	270
PART V	Preparing the Research Report	
	CHAPTER 12 Writing the Final Research Report	310
	Appendix A Using a Spreadsheet: Microsoft Excel	334
	Appendix B Using SPSS	342
	References	348
	Index	353

Contents

PART I The Fundamentals

Chapter 1

The Nature and Tools of Research 1

What Research Is Not 1

What Research Is 2

Tools of Research 7

The Library and Its Resources 8



Computer Technology 8

Measurement 9

Statistics 10

Language 11

PRACTICAL APPLICATION: Communicating Effectively
through Writing 13



GUIDELINES: Writing to Communicate 13

GUIDELINES: Using a Word Processor 15

The Human Mind 16

Critical Thinking 17

Deductive Logic 17

Inductive Reasoning 18

The Scientific Method 19

Theory Building 20

Collaboration with Other Minds 21



CONCEPTUAL ANALYSIS EXERCISE: Using the Internet to Facilitate
Communication and Collaboration with Others 22

Reflections on Significant Research 22

Exploring Research in Your Field 24

PRACTICAL APPLICATION: Identifying Important Tools in Your
Discipline 24

CHECKLIST: Interviewing an Expert Researcher 25

For Further Reading 25

Answers to the Conceptual Analysis Exercise “Using the Internet to Facilitate
Communication and Collaboration with Others” 26

PART II Focusing Your Research Efforts

Chapter 2

The Problem: The Heart of the Research Process 27

Finding Research Projects 27

PRACTICAL APPLICATION: Identifying and Describing the Research Problem 29

GUIDELINES: Finding a Legitimate Problem 29

GUIDELINES: Stating the Research Problem 31

CHECKLIST: Evaluating the Research Problem 35

Dividing the Research Problem into Subproblems 36

Subproblems versus Pseudo-Subproblems 36

Characteristics of Subproblems 36

Identifying Subproblems 37



Taking a Paper-and-Pencil Approach 37

Using Brainstorming Software 38

Every Problem Needs Further Delineation 39

Stating the Hypotheses and/or Research Questions 39

Distinguishing Null Hypotheses from Research Hypotheses 40

Identifying the Variables under Investigation 40

CONCEPTUAL ANALYSIS EXERCISE: Identifying Independent, Dependent, Mediating, and Moderating Variables 42

Delimiting the Research 43

Defining Terms 44

Stating the Assumptions 44

Importance of the Study 44

Ordering the Topics in a Research Proposal 45

PRACTICAL APPLICATION: Writing the First Sections of a Proposal 45

CHECKLIST: Evaluating Your Proposed Research Project 46

PRACTICAL APPLICATION: Reappraising a Proposed Research Problem 47

GUIDELINES: Fine-Tuning Your Research Problem 47

For Further Reading 48

Answers to the Conceptual Analysis Exercise “Identifying Independent, Dependent, Mediating, and Moderating Variables” 49

Chapter 3

Review of the Related Literature 51

Understanding the Role of the Literature Review 51

Strategies for Locating Related Literature 52



Using the Library Catalog 52

Using Indexes and Abstracts 55



Using Online Databases 55

Consulting with Reference Librarians 58



Surfing the Internet 58



Locating Relevant Government Publications 59

Using Citations and Reference Lists of Those Who Have Gone Before You 59

PRACTICAL APPLICATION: Conducting a Literature Search 60

GUIDELINES: Using Your Library Time Efficiently 61

PRACTICAL APPLICATION: Evaluating the Research of Others 64

CHECKLIST: Questions to Consider When Evaluating Research 64



Organizing the Information You Collect 65

Creating a Computer Database for Your Related Literature 65

Knowing When to Quit 66

Organizing and Synthesizing the Literature into a Cohesive

Review 66

PRACTICAL APPLICATION: Writing the Literature Review 67

GUIDELINES: Writing a Clear and Cohesive Literature Review 67

A Sample Research Proposal 70

For Further Reading 73

Chapter 4

Planning Your Research Project 74

Planning a General Approach 74

The Basic Format of All Research 74

Research Planning versus Research Methodology 76

General Criteria for a Research Project 76

The Nature and Role of Data in Research 77

What Are Data? 77

Data Are Transient and Ever Changing 77

Primary Data versus Secondary Data 78

Criteria for the Admissibility of Data 79

Planning for Data Collection 80

Identifying Good Measurement Instruments 81

Defining Measurement 81

Measuring Insubstantial Phenomena: An Example 82

Types of Measurement Scales 84

Nominal Scales 85

Ordinal Scales 85

Interval Scales 86

Ratio Scales 86

CONCEPTUAL ANALYSIS EXERCISE: Identifying Scales of Measurement 88

Validity and Reliability in Measurement 89

Validity of Measurement Instruments 89

Reliability of Measurement Instruments 91

Enhancing the Reliability and Validity of a Measurement Instrument 92

CONCEPTUAL ANALYSIS EXERCISE: Identifying Problems with Validity and Reliability in Measurement 93

Linking Data and Research Methodology	94
<i>Comparing Quantitative and Qualitative Approaches</i>	95
Combining Quantitative and Qualitative Designs	98
PRACTICAL APPLICATION: Choosing a General Research Approach	98
GUIDELINES: Deciding Whether to Use a Quantitative or Qualitative Approach	98
Considering the Validity of Your Method	101
<i>Internal Validity</i>	101
<i>External Validity</i>	103
<i>Validity in Qualitative Research</i>	104
Ethical Issues in Research	104
<i>Protection from Harm</i>	105
<i>Voluntary and Informed Participation</i>	105
<i>Right to Privacy</i>	107
<i>Honesty with Professional Colleagues</i>	108
<i>Internal Review Boards</i>	108
<i>Professional Codes of Ethics</i>	109
PRACTICAL APPLICATION: Planning an Ethical Research Study	109
CHECKLIST: Determining Whether Your Proposed Study Is Ethically Defensible	109
Critically Scrutinizing Your Overall Plan	110
PRACTICAL APPLICATION: Judging the Feasibility of a Research Project	110
CHECKLIST: Determining Whether a Proposed Project Is Realistic and Practical	110
<i>When You Can't Anticipate Everything in Advance: The Value of a Pilot Study</i>	112
PRACTICAL APPLICATION: Establishing Research Criteria and Justifying Your Research Methodology	112
PRACTICAL APPLICATION: Developing a Plan of Attack	113
<i>Using Project Management Software and Electronic Planners</i>	114
<i>Keeping an Optimistic and Task-Oriented Outlook</i>	114
For Further Reading	116
Answers to the Conceptual Analysis Exercise "Identifying Scales of Measurement"	117
Answers to the Conceptual Analysis Exercise "Identifying Problems with Validity and Reliability in Measurement"	118



Chapter 5

Writing the Research Proposal.	119
Characteristics of a Proposal	120
<i>A Proposal Is a Straightforward Document</i>	120
<i>A Proposal Is Not a Literary Production</i>	121
<i>A Proposal Is Clearly Organized</i>	121
Organizing and Writing a Research Proposal	121
<i>Formatting Headings and Subheadings</i>	122
PRACTICAL APPLICATION: Writing Your Proposal	123

GUIDELINES: Writing the First Draft	124
GUIDELINES: Revising Your Proposal	128
PRACTICAL APPLICATION: Strengthening Your Proposal	132
CHECKLIST: Features Detracting from Proposal Effectiveness	132
Final Thoughts about Proposal Writing	134
A Sample Research Proposal	134
For Further Reading	138

PART III Qualitative Research Methodologies

Chapter 6

Qualitative Research 139

The Nature of the Research Problem and Planning in Qualitative Research 140

When to Choose a Qualitative Approach 140

Qualitative Research Designs 141

Case Study 141

Ethnography 142

Phenomenological Study 145

Grounded Theory Study 146

Content Analysis 148

CONCEPTUAL ANALYSIS EXERCISE: Choosing a Qualitative Research Design 150

Collecting Data in Qualitative Research 151

Sampling 152

Observations 152

Interviews 153

PRACTICAL APPLICATION: Conducting Interviews in a Qualitative Study 154

GUIDELINES: Conducting a Productive Interview 154

Using Technology to Facilitate Data Collection and Transcription 157

Organizing and Analyzing Qualitative Data 158

Using Computer Databases to Facilitate Data Organization and Interpretation 159

PRACTICAL APPLICATION: Planning a Qualitative Study 160

CHECKLIST: Pinning Down the Details of a Qualitative Study 160

Criteria for Evaluating Qualitative Research 162

PRACTICAL APPLICATION: Evaluating Qualitative Research Studies 163

CHECKLIST: Evaluating a Qualitative Study 163

A Sample Dissertation 164

For Further Reading 168

Answers to the Conceptual Analysis Exercise "Choosing a Qualitative Research Design" 169



Chapter 7

Historical Research 170

 Data Sources in Historical Research 170

 *Collecting Historical Records* 174

 *Online Databases for Historical Events* 174

 **PRACTICAL APPLICATION:** Handling Historical Data Systematically 175

Evaluating and Interpreting Historical Data 176

External Evidence 176

Internal Evidence 176

Psychological or Conceptual Historical Research 177

Searching for Roots 178

PRACTICAL APPLICATION: Historical Research Writing 178

GUIDELINES: Writing the Historical Research Report 178

A Sample Dissertation 179

For Further Reading 182

PART IV Quantitative Research Methodologies

Chapter 8

Descriptive Research 184

Descriptive Research Designs 184

Observation Studies 184

Correlational Research 185

A Caution about Interpreting Correlational Results 187

Developmental Designs 188

Survey Research 189

Face-to-Face and Telephone Interviews 190

Questionnaires 191

Planning for Data Collection in a Descriptive Study 191

 **PRACTICAL APPLICATION:** Using Checklists and Rating Scales 192

 **PRACTICAL APPLICATION:** Computerizing Observations 194

 **PRACTICAL APPLICATION:** Planning and Conducting Interviews in a Quantitative Study 194

 **GUIDELINES:** Conducting Interviews in a Quantitative Study 194

An Example in International Relations 195

PRACTICAL APPLICATION: Constructing and Administering a Questionnaire 196

GUIDELINES: Constructing Your Questionnaire 196

GUIDELINES: Using Technology to Facilitate Questionnaire Administration and Data Analysis 200

 **GUIDELINES:** Maximizing Your Return Rate for a Questionnaire 201

 **PRACTICAL APPLICATION:** Using the Internet to Collect Data for a Descriptive Study 205

Choosing a Sample in a Descriptive Study	206
<i>Sampling Designs</i>	207
Probability Sampling	207
Nonprobability Sampling	214
<i>Sampling in Surveys of Very Large Populations</i>	215
<i>Identifying a Sufficient Sample Size</i>	215
<i>Bias in Research Sampling</i>	216
Acknowledging the Probable Presence of Bias	218
PRACTICAL APPLICATION: Population Analysis for a Descriptive Study	219
CHECKLIST: Analyzing a Population	219
Interpreting Data in Descriptive Research	220
Some Final Suggestions	220
A Sample Dissertation	221
For Further Reading	225

Chapter 9

Experimental, Quasi-Experimental, and Ex Post Facto Designs 226

The Importance of Control	227
<i>Controlling for Confounding Variables</i>	229
Overview of Experimental, Quasi-Experimental, and Ex Post Facto Designs	232
Pre-Experimental Designs	233
<i>Design 1: One-Shot Experimental Case Study</i>	233
<i>Design 2: One-Group Pretest–Posttest Design</i>	233
<i>Design 3: Static Group Comparison</i>	234
True Experimental Designs	234
<i>Design 4: Pretest–Posttest Control Group Design</i>	234
<i>Design 5: Solomon Four-Group Design</i>	235
<i>Design 6: Posttest-Only Control Group Design</i>	235
<i>Design 7: Within-Subjects Design</i>	236
Quasi-Experimental Designs	237
<i>Design 8: Nonrandomized Control Group Pretest–Posttest Design</i>	237
<i>Design 9: Simple Time-Series Design</i>	238
<i>Design 10: Control Group, Time-Series Design</i>	238
<i>Design 11: Reversal Time-Series Design</i>	238
<i>Design 12: Alternating Treatments Design</i>	239
<i>Design 13: Multiple Baseline Design</i>	239
<i>Using Designs 11, 12, and 13 in Single-Subject Studies</i>	241
Ex Post Facto Designs	242
<i>Design 14: Simple Ex Post Facto Design</i>	242
Factorial Designs	243
<i>Design 15: Two-Factor Experimental Design</i>	243

Design 16: Combined Experimental and Ex Post Facto Design 244

PRACTICAL APPLICATION: Determining Possible Cause-and-Effect Relationships 248

CHECKLIST: Looking for Confounding Variables 248

Meta-Analyses 249



Conducting Experiments on the Internet 250

Testing Your Hypotheses, and Beyond 251

CONCEPTUAL ANALYSIS EXERCISE: Identifying Research Designs 251

A Sample Dissertation 252

For Further Reading 256

Answers to the Conceptual Analysis Exercise "Identifying Research Designs" 256

Chapter 10

Mixed-Methods Research 258

When Mixed-Methods Designs Are Most Useful and Appropriate 258

Common Mixed-Methods Designs 259

Planning a Mixed-Methods Study 260

Identifying Research Questions and Hypotheses 261

Conducting the Literature Review 262

Addressing Validity Concerns 262

Special Ethical Considerations in Mixed-Methods Research 263

PRACTICAL APPLICATION: Deciding Whether to Use a Mixed-Methods Design 263

CHECKLIST: Pinning Down the Logistics and Feasibility of a Mixed-Methods Study 263

Analyzing and Interpreting Mixed-Methods Data 264

A Sample Dissertation 265

For Further Reading 269

Chapter 11

Strategies for Analyzing Quantitative Data 270

Exploring and Organizing a Data Set 270

Organizing Data to Make Them Easier to Think about and Interpret 272



Using Computer Spreadsheets to Organize and Analyze Data 274

Choosing Appropriate Statistics 276

Functions of Statistics 277

Statistics as Estimates of Population Parameters 278

Considering the Nature of the Data 278

Single-Group versus Multi-Group Data 278

Continuous versus Discrete Variables 279

Nominal, Ordinal, Interval, and Ratio Data 279

Normal and Non-Normal Distributions 280

Choosing between Parametric and Nonparametric Statistics 282

Descriptive Statistics	282
<i>Measures of Central Tendency</i>	283
Curves Determine Means	284
Measures of Central Tendency as Predictors	286
<i>Measures of Variability: Dispersion and Deviation</i>	286
How Great Is the Spread?	287
Using the Mean and Standard Deviation to Calculate Standard Scores	288
Keeping Measures of Central Tendency and Variability in Perspective	290
<i>Measures of Association: Correlation</i>	291
How Validity and Reliability Affect Correlation Coefficients	293
A Reminder about Correlation	293
Inferential Statistics	294
<i>Estimating Population Parameters</i>	294
An Example: Estimating a Population Mean	294
Point versus Interval Estimates	296
<i>Testing Hypotheses</i>	296
Making Errors in Hypothesis Testing	298
Another Look at Statistical Hypotheses versus Research Hypotheses	299
Examples of Statistical Techniques for Testing Hypotheses	300
<i>Meta-Analysis</i>	300
Using Statistical Software Packages	302
Interpreting the Data	303
A Sample Dissertation	304
PRACTICAL APPLICATION: Analyzing Data in a Quantitative Study	308
CHECKLIST: Questions to Consider When Choosing a Statistical Procedure	308
For Further Reading	309



PART V Preparing the Research Report

Chapter 12

Writing the Final Research Report 310



Getting Started	310
<i>Surfing the Internet for Writing Assistance</i>	310
<i>Learn by Looking</i>	311
Planning a Research Report	311
<i>Description of the Research Problem</i>	312
<i>Description of the Method</i>	312
<i>Presentation of the Data</i>	313
<i>Interpretation of the Data</i>	314
<i>Concluding the Report</i>	315
<i>Maintaining Your Academic Integrity</i>	316



Front Matter and End Matter	317
<i>Preliminary Pages</i>	317
<i>Endnotes and Footnotes</i>	318
<i>Reference List</i>	318
<i>Appendix Content</i>	321
Organizing the Research Report	321
Preparing the Report	323
<i>Style of Prose</i>	323
<i>General Principles of Writing</i>	324
PRACTICAL APPLICATION: Developing a Writing Schedule	325
GUIDELINES: Developing a Workable Schedule	325
PRACTICAL APPLICATION: Critiquing a Research Report	327
CHECKLIST: Criteria for Critiquing a Research Report	327
Beyond the Unpublished Research Report: Presenting and Publishing	329
<i>Conference Presentations</i>	329
PRACTICAL APPLICATION: Presenting Your Research at a Professional Conference	329
GUIDELINES: Presenting an Effective Paper or Poster	330
<i>Journal Articles</i>	331
<i>Sharing Authorship</i>	331
<i>Responding to Reviewers' Critiques</i>	332
A Closing Thought	332
For Further Reading	333

Appendix A Using a Spreadsheet: Microsoft Excel 334

Using Excel to Keep Track of Literature Resources	334
Using Excel to Record and Recode Data	336
Reorganizing Data in Excel	339
Using Excel to Perform Simple Statistical Analyses	340

Appendix B Using SPSS 342

Creating a Data Set	342
Computing Basic Descriptive Statistics	344
Computing Inferential Statistics	345

References 348

Index 353

1

The Nature and Tools of Research

In virtually every subject area, our knowledge is incomplete and problems are waiting to be solved. We can address our incomplete knowledge and unsolved problems by asking relevant questions and then seeking answers through systematic research. We have many tools at our disposal to help us do these things—not only physical tools but also mental and social tools.

In everyday speech, the word *research* has numerous meanings, making it a decidedly confusing term for university students, who must learn to use the word in a narrower, more precise sense. From elementary school to college, students hear the word *research* used to describe a variety of activities. In some situations the word connotes finding a piece of information or taking notes and then writing a so-called “research paper.” In other situations it refers to the act of informing oneself about what one does not know, perhaps by rummaging through available sources to locate a few tidbits of information. Such activities have often been called research but are more accurately called other names: information gathering, library skills, self-enlightenment, documentation, or summarization.

Yet when used more appropriately, for many people the word *research* suggests a mystical activity that is somehow exclusive and removed from everyday life. Researchers are sometimes regarded as aloof individuals who seclude themselves in laboratories, scholarly libraries, or the ivory towers of large universities. The general public is often unaware of what researchers do on a day-to-day basis and how their work contributes to people’s overall quality of life and well-being.

In fact, research is often a practical enterprise that—given appropriate tools—*any* rational, conscientious individual can conduct. In this chapter we lay out the nature of true research and describe the general tools that make it possible.

What Research Is Not

We have suggested that the word *research* has been so widely used in everyday speech that few people have any idea of its true meaning. Following are three statements that describe what research is not. Accompanying each statement is an example that illustrates a common misconception about research.

1. *Research is not merely gathering information.* A sixth grader comes home from school and tells her parents, “The teacher sent us to the library today to do research, and I learned a lot about black holes.” For this student, research means going to the library to find a few facts. This might be *information discovery*, or it might be learning *reference skills*. But it certainly is not, as the teacher labeled it, research.

2. *Research is not merely rummaging around for hard-to-locate information.* The house across the street is for sale. You consider buying it and call your realtor to find out much money someone else might pay you for your current home. “I’ll have to do some research to determine the fair market value of your property,” the realtor tells you. What the realtor calls doing “some research” means, of course, reviewing information about recent sales of properties comparable to yours; this information will help the realtor zero in on a reasonable asking price for your own home. Such an activity involves little more than rummaging through files to discover what the realtor

previously did not know. Rummaging—whether through one's personal records or at the public or college library—is not research. It is more accurately called an *exercise in self-enlightenment*.

3. *Research is not merely transporting facts from one location to another.* A college student reads several articles about the mysterious Dark Lady in William Shakespeare's sonnets and then writes a "research paper" describing various scholars' suggestions of who the lady might have been. Although the student does, indeed, go through certain activities associated with formal research—collecting information, organizing it in a certain way for presentation to others, supporting statements with documentation, referencing statements properly, and so on—these activities do not add up to a true research paper. The student has missed the essence of research: the interpretation of data. Nowhere in the paper does the student say, in effect, "These facts I have gathered seem to indicate such-and-such about the Dark Lady." Nowhere does the student interpret and draw conclusions from the facts. This student is approaching genuine research; however, the mere compilation of facts, presented with reference citations and arranged in a logical sequence—no matter how polished and appealing the format—misses genuine research by a hair. Such activity might more realistically be called *fact transcription*, *fact documentation*, *fact organization*, or *fact summarization*. Going a little further, this student would have traveled from one world to another: from the world of mere transportation of facts to the world of interpretation of facts. The difference between the two worlds is the distinction between transference of information and genuine research—a distinction that is critical for novice researchers to understand.

What Research Is

Research is a systematic process of collecting, analyzing, and interpreting information—*data*—in order to increase our understanding of a phenomenon about which we are interested or concerned. People often use a systematic approach when they collect and interpret information to solve the small problems of daily living. Here, however, we focus on *formal research*, research in which we intentionally set out to enhance our understanding of a phenomenon and expect to communicate what we discover to the larger scientific community.

Although research projects vary in complexity and duration, in general, research has eight distinct characteristics:

1. Research originates with a question or problem.
2. Research requires clear articulation of a goal.
3. Research usually divides the principal problem into more manageable subproblems.
4. Research is guided by the specific research problem, question, or hypothesis.
5. Research requires a specific plan for proceeding.
6. Research rests on certain critical assumptions.
7. Research requires the collection and interpretation of data in an attempt to resolve the problem that initiated the research.
8. Research is, by its nature, cyclical or, more exactly, helical.

Let's look at each of these characteristics more closely.

1. *Research originates with a question or problem.* The world is filled with unanswered questions and unresolved problems. Everywhere we look, we see things that cause us to wonder, to speculate, to ask questions. And by asking questions, we strike a spark that ignites a chain reaction leading to the research process. An inquisitive mind is the beginning impetus for research; as one popular tabloid puts it, "Inquiring minds want to know!"

Look around you. Consider the unresolved situations that evoke these questions: What is such-and-such a situation like? Why does such-and-such a phenomenon occur? What does it all mean? These are everyday questions. With questions like these, research begins.

2. *Research requires clear articulation of a goal.* A clear, unambiguous statement of the problem is critical. This statement is an exercise in intellectual honesty: The ultimate goal of the

research must be set forth in a grammatically complete sentence that specifically and precisely answers the question, "What problem do you intend to solve?" When you describe your objective in clear, concrete terms, you have a good idea of what you need to accomplish and can direct your efforts accordingly.

3. *Research usually divides the principal problem into more manageable subproblems.* From a design standpoint, it is often helpful to break a main research problem into several subproblems that, when solved, can resolve the main problem.

Breaking down principal problems into small, easily solvable subproblems is a strategy we use in everyday living. For example, suppose you want to drive from your hometown to a town 50 miles away. Your principal goal is to get from one location to the other as expeditiously as possible. You soon realize, however, that the problem involves several subproblems:

- | | |
|---------------|--|
| Main problem: | How do I get from Town A to Town B? |
| Subproblems: | 1. What route appears to be the most direct one? |
| | 2. Is the most direct one also the quickest one? If not, what route might take the least amount of time? |
| | 3. Which is more important to me: minimizing my travel time or minimizing my energy consumption? |
| | 4. At what critical junctions in my chosen route must I turn right or left? |

What seems like a single question can be divided into several smaller questions that must be addressed before the principal question can be resolved.

So it is with most research problems. By closely inspecting the principal problem, the researcher often uncovers important subproblems. By addressing each of the subproblems, the researcher can more easily address the main problem. If a researcher doesn't take the time or trouble to isolate the lesser problems within the major problem, the overall research project can become cumbersome and difficult to manage.

Identifying and clearly articulating the problem and its subproblems are the essential starting points for formal research. Accordingly, we discuss these processes in depth in Chapter 2.

4. *Research is guided by the specific research problem, question, or hypothesis.* Having stated the problem and its attendant subproblems, the researcher usually forms one or more hypotheses about what he or she may discover. A **hypothesis** is a logical supposition, a reasonable guess, an educated conjecture. It provides a tentative explanation for a phenomenon under investigation. It may direct your thinking to possible sources of information that will aid in resolving one or more subproblems and, as a result, may also help to resolve the principal research problem.

Hypotheses are certainly not unique to research. They are constant, recurring features of everyday life and represent the natural working of the human mind. Something happens. Immediately you attempt to account for the cause of the event by making a series of reasonable guesses. In so doing, you are hypothesizing. As an example, let's take a commonplace event: You come home after dark, open the front door, and reach inside for the switch that turns on a nearby table lamp. Your fingers find the switch. You flip it. No light. At this point, you begin to construct a series of reasonable guesses—hypotheses—to explain the lamp's failure:

1. The bulb has burned out.
2. The lamp is not plugged into the wall outlet.
3. A recent weather event interrupted your electrical service.
4. The wire from the lamp to the wall outlet is defective.
5. You forgot to pay your electric bill.

Each of these hypotheses hints at a direction you might proceed in order to acquire information that may resolve the problem of the malfunctioning lamp. Now you go in search of information to determine which hypothesis is correct. In other words, you look for data that will support one of your hypotheses and enable you to reject others.

1. You get a flashlight from your car, find a new bulb, and put the new bulb in the lamp. The lamp fails to light. (Hypothesis 1 is rejected.)
2. You glance down at the wall outlet and see that the lamp is plugged into it. (Hypothesis 2 is rejected.)
3. You look at your neighbors' homes. Everyone has electrical power. (Hypothesis 3 is rejected.)
4. You lift the cord that connects the lamp to the wall outlet. The lamp lights briefly and then goes out. You lift the cord again. Again the lamp lights briefly. The connecting cord is defective. (Hypothesis 4 is supported. Furthermore, because you clearly do have an active electric current, you can reject Hypothesis 5—your electric bill payments are up to date.)
5. Fortunately, Hypothesis 4 solved the problem. By repairing or replacing the cord, you can count on adequate light from the lamp in the near future.

Hypotheses in a research project are as tentative as those just formed for the malfunctioning lamp. For example, a biologist might speculate that certain human-made chemical compounds increase the frequency of birth defects in frogs. A psychologist might speculate that certain personality traits lead people to show predominantly liberal or conservative voting patterns. A marketing researcher might speculate that humor in a television commercial will capture viewers' attention and thereby will increase the odds that viewers buy the advertised product. Notice the word *speculate* in all of these examples. Good researchers always begin a project with open minds about what they may—or may *not*—discover in their data.

Let's return to a point we made a few paragraphs back, this time emphasizing a particular word: The researcher *usually* forms one or more hypotheses about what he or she may discover. Hypotheses—predictions—are an essential ingredient in certain kinds of research, especially experimental research (see Chapter 9). To a lesser degree, they guide most other forms of research as well, but they are intentionally *not* identified in the early stages of some kinds of qualitative research (e.g., see the discussion of *grounded theory* research in Chapter 6). Yet regardless of whether researchers form specific hypotheses in advance, they must, at a minimum, use their research problem or question to focus their efforts.

5. *Research requires a specific plan for proceeding.* Research is not a blind excursion into the unknown, with the hope that the data necessary to answer the question at hand will somehow fortuitously emerge. It is, instead, a carefully planned itinerary of the route you intend to take in order to reach your final destination—your research goal. Consider the title of this text: *Practical Research: Planning and Design*. The last three words—*Planning and Design*—are especially important ones. Researchers plan their overall research design and specific research methods in a purposeful way so that they can acquire data relevant to their research problem and subproblems. Depending on the research question, different designs and methods are more or less appropriate.

In addition to identifying the specific goal of your research, then, you must also identify how you propose to reach your goal. You cannot wait until you're chin deep in the project to plan and design your strategy. In the formative stages of a research project, much can be decided: Where are the data? Do any existing data address themselves to the research problem? If the data exist, are you likely to have access to them? And if you have access to the data, what will you do with them after you have them? Such questions merely hint at the fact that planning and design cannot be postponed. Each of the questions just listed—and many more—must have an answer early in the research process.¹

Ultimately the research methodology directs the whole research endeavor: It controls the study, dictates how the data are acquired, arranges them in logical relationships, sets up an approach for refining and synthesizing them, suggests a manner in which the meanings that lie

¹As should be apparent in the questions we pose in this paragraph, we are using the word *data* as a plural noun; for instance, we ask "Where *are* the data?" rather than "Where *is* the data?" Contrary to popular usage of the term as a singular noun, *data* (which has its origins in Latin) refers to two or more pieces of information. A single piece of information is known as a *datum*, or sometimes as a *data point*.

below the surface of the data become manifest, and finally yields one or more conclusions that lead to an expansion of knowledge. Thus, research methodology has two primary functions:

1. To dictate and control the acquisition of data
2. To analyze the acquired data in order to extract meaning from them

The second of these functions is what we mean by the phrase *interpretation of the data*.

6. *Research rests on certain critical assumptions.* Whereas a hypothesis involves a prediction that may or may not be supported by the data, an **assumption** is a condition that is taken for granted, without which the research project would be pointless. In research, assumptions are equivalent to axioms in geometry—self-evident truths that any reasonable person might accept. Careful researchers—certainly those conducting research in an academic environment—set forth a statement of their assumptions as the bedrock upon which their study rests.

An example may clarify the point. Imagine that your problem is to investigate whether students learn the unique grammatical structures of a language more quickly by studying only one foreign language at a time or by studying two foreign languages concurrently. What assumptions would underlie such a problem? At a minimum, the researcher must assume that

- The teachers used in the study are competent to teach the language or languages in question and have mastered the grammatical structures of the language(s) they are teaching.
- The students taking part in the research are capable of mastering the unique grammatical structures of any language(s) they are studying.
- The languages selected for the study have sufficiently different grammatical structures that students might reasonably learn to distinguish between them.

Assumptions are often so self-evident that a researcher may consider it unnecessary to mention them. For instance, two assumptions underlie almost all research:

- The phenomenon under investigation is somewhat lawful and predictable; it is *not* comprised of completely random events.
- Certain cause-and-effect relationships can account for the patterns observed in the phenomenon.

Aside from such basic ideas as these, however, careful researchers state their assumptions, so that other people inspecting the research project can evaluate it in accordance with their *own* assumptions. For the beginning researcher, it is better to be overly explicit than to take too much for granted.

7. *Research requires the collection and interpretation of data in an attempt to resolve the problem that initiated the research.* After a researcher has isolated the problem, divided it into appropriate subproblems, posited reasonable hypotheses, identified a suitable design and methodology, and identified the assumptions that underlie the entire effort, the next step is to collect whatever data seem appropriate and to organize them in meaningful ways so that they can be interpreted.²

Events, observations, and measurements are, in and of themselves, *only* events, observations, and measurements—nothing more. The significance of the data depends on how the researcher extracts *meaning* from them. In research, data uninterpreted by the human mind are worthless: They can never help us answer the questions we have posed.

Yet researchers must recognize and come to terms with the subjective and dynamic nature of interpretation. Consider, for example, the many books written on the assassination of U.S. President John F. Kennedy. Different historians have studied the same events: One may interpret them one way, and another may arrive at a very different conclusion. Which one is right? Perhaps

²Some people in academia use the term *research* more broadly to include deriving new equations or abstract principles from existing equations and/or principles through a sequence of mathematically logical and valid steps. Such an activity can be quite intellectually challenging, of course, and is often at the heart of doctoral dissertations and scholarly journal articles in mathematics, physics, and related disciplines. In this book, however, we use the term *research* more narrowly to refer to *empirical* research—research that involves the collection and analysis of new data.

they both are; perhaps neither is. Both may have merely posed new problems for other historians to try to resolve. Different minds often find different meanings in the same set of facts.

Once we believed that clocks measured time and that yardsticks measured space. In one sense, they still do. We further assumed that time and space were two different entities. Then along came Einstein's theory of relativity, and time and space became locked into one concept: the time-space continuum. What is the difference between the old perspective and the new perspective? The way we think about, or interpret, the same information. The realities of time and space have not changed; the way we interpret them has.

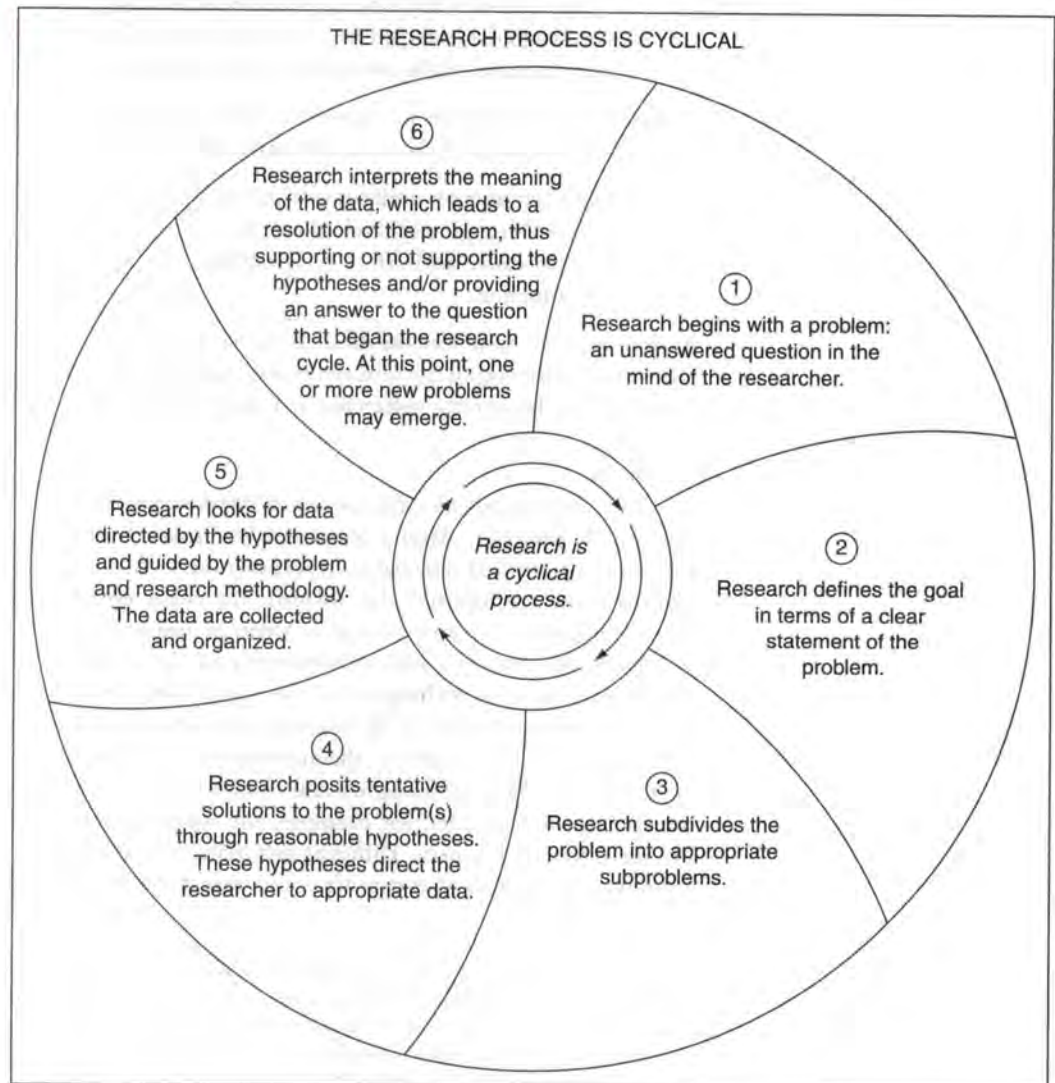
Data demand interpretation. But no rule, formula, or algorithm can lead the researcher unerringly to a correct interpretation. Interpretation is inevitably a somewhat subjective process that depends on the researcher's hypotheses, assumptions, and logical reasoning processes. In subsequent chapters we present a number of potentially useful methods for organizing and interpreting data.

Now think about how we began this chapter. We suggested that certain activities cannot accurately be called research. At this point you can understand why. None of those activities demands that the researcher draw any conclusions or make any interpretation of the data.

8. *Research is, by its nature, cyclical or, more exactly, helical.* Any research project begins simply and then follows a predictable, systematic sequences of steps, as shown in Figure 1.1:

FIGURE 1.1

The research cycle



1. A questioning mind observes a particular situation and asks, Why? What caused that? How come? (This is the internal, mental origin of research.)
2. One or more of these questions become formally stated as a problem. (This is the overt, observable beginning of research.)
3. The problem is divided into several simpler, more specific subproblems.
4. Preliminary information is gathered that appears to bear on the problem. This information may include informal observations of events in one's environment; typically it also includes previous research findings related to the topic at hand. The preliminary information may point to a tentative solution to the problem. A guess is made; a hypothesis or guiding question is formed.
5. A method of collecting data more systematically is identified and carried out in order to address the problem.
6. The body of data is processed and interpreted. A discovery is made, a conclusion reached. The tentative hypothesis is either supported or not supported by the data; the question is either answered (partially or completely) or not answered.

Such is the general format of all research. Different academic disciplines merely use different routes to arrive at the same destination.

Yet only rarely is a research project a one-shot effort that completely resolves a problem. For instance, even with the best of data, hypotheses in a research project are rarely proved or disproved—and thus research questions are rarely answered—beyond a shadow of a doubt. Instead, hypotheses are either *supported* or *not supported* by the data. If the data are consistent with a particular hypothesis, the researcher can make a case that the hypothesis probably has some merit and should be taken seriously. In contrast, if the data run contrary to a hypothesis, the researcher *rejects* the hypothesis and turns to other hypotheses as being more likely explanations of the phenomenon in question. In either case, one or more additional, follow-up studies are called for.

Ultimately, then, most research studies don't bring total closure to a research problem. There is no obvious end point—no point at which a researcher can say “*Voilà!* I've completely answered the question about which I'm concerned.” Instead, research typically involves a cycle, or more accurately, a *belix* (spiral) in which one study spawns additional, follow-up studies. In exploring a topic, one comes across additional problems that need resolving, and so the process must begin anew. Research begets more research.

To view research in this way is to invest it with a dynamic quality that is its true nature—a far cry from the conventional view, which sees research as a one-time act that is static, self-contained, an end in itself. Here we see another difference between true research and the nonexamples of research presented early in the chapter. Every researcher soon learns that genuine research is likely to yield as many problems as it resolves. Such is the nature of the acquisition of knowledge.

Tools of Research

Every professional needs specialized tools in order to work effectively. Without hammer and saw, the carpenter is out of business; without scalpel or forceps, the surgeon cannot practice. Researchers, likewise, have their own set of tools to carry out their plans.

The tools researchers use to achieve their research goals may vary considerably depending on the discipline. The microbiologist needs a microscope and culture media; the attorney requires a library of legal decisions and statute law. By and large, we do not discuss such discipline-specific tools in this book. Rather, our concern here is with general tools of research that the great majority of researchers of all disciplines need in order to collect data and derive meaningful conclusions.

We should be careful not to equate the *tools* of research with the *methodology* of research. A **research tool** is a specific mechanism or strategy the researcher uses to collect, manipulate, or interpret data. The **research methodology** is the general approach the researcher takes in carrying out the research project; to some extent, this approach dictates the particular tools the researcher selects.

Confusion between the tool and the research method is immediately recognizable. Such phrases as “library research” and “statistical research” are telltale signs and largely meaningless terms. They suggest a failure to understand the nature of formal research, as well as a failure to differentiate between tool and method. The library is merely a place for locating or discovering certain data that will be analyzed and interpreted at some point in the research process. Likewise, statistics merely provide ways to summarize and analyze data, thereby allowing us to see patterns within the data more clearly.

Six general tools of research are these:

1. The library and its resources
2. Computer technology
3. Measurement
4. Statistics
5. Language
6. The human mind

In the following sections, we look more closely at each of these general tools.

The Library and Its Resources

Historically, many literate human societies used libraries to assemble and store their collective knowledge. For example, in the seventh century B.C., the ancient Assyrians’ Library of Nineveh contained 20,000 to 30,000 tablets, and in the second century A.D., the Romans’ Library of Celsus in Ephesus housed more than 12,000 parchment and papyrus scrolls.

Until the past few decades libraries were primarily repositories of concrete, physical representations of knowledge—clay tablets, scrolls, manuscripts, books, journals, and so on. For the most part, any society’s collective knowledge expanded rather slowly and could seemingly be contained within masonry walls. But by the latter half of the 20th century, people’s knowledge about their physical and social worlds began to increase many times over, and at the present time it continues to increase at an astounding rate. In response, libraries have evolved in important ways. First, they have made use of many emerging technologies (e.g., microforms, compact disks, online databases) to store information in more compact forms. Second, they have provided increasingly fast and efficient means of locating and accessing information on virtually any topic. And third, many of them have made catalogs of their holdings available on the Internet (e.g., see www.library.unh.edu or library.brown.edu). The libraries of today—especially university libraries—extend far beyond their local, physical boundaries.

We explore efficient use of a library and its resources in depth in Chapter 3. For now, we simply want to stress that the library is—and must be—one of the most valuable tools in any researcher’s toolbox.

Computer Technology



As a research tool, the personal computer is now commonplace. Personal computers have become increasingly compact and portable—first in the form of laptops and more recently in the form of iPads and other tablet computers. And computer software packages have become increasingly user friendly, such that novice researchers can learn to use them quickly and easily. But like any tool—no matter how powerful—computers have their limitations. Yes, computers can certainly calculate, compare, search, retrieve, sort, and organize data more efficiently and more accurately than you can. But in their present stage of development, they depend largely on people to give them directions about what to do.

A computer is not a miracle worker. It cannot do your thinking for you. It can, however, be a fast and faithful assistant. When told exactly what to do, it is one of the researcher’s best friends.

Throughout this book, you will find many “Using Technology” sections that describe specific ways in which, as a researcher, you can use computers to make your job easier. Table 1.1 provides suggestions for how you might use a computer to assist you in the research process.

TABLE 1.1 The computer as a research assistant

<i>Part of the Study</i>	<i>Team of Research Assistants</i>
Planning the study	• Brainstorming assistance—software used to help generate and organize ideas for the research focus, to illustrate how different concepts could be related, and to consider how the process will be conducted. • Outlining assistance—software used to help structure the different aspects of the study and coordinate work efforts. • Project management assistance—software used to highlight and coordinate all the different efforts that need to occur in a timely fashion. • Budget assistance—spreadsheet software to help in outlining, estimating, and monitoring the potential costs involved in the research effort.
Literature review	• Background literature identification assistance—CDs and online databases that identify and describe related published research that should be considered during the formative stages of the research endeavor. • Telecommunication assistance—computer technology used to communicate with other researchers and groups of researchers through e-mail, electronic bulletin boards, list servers, and so on. • Writing assistance—software used to facilitate the writing, editing, formatting, and printing of the literature review.
Study implementation and data gathering	• Materials production assistance—software used for the development and use of instructional materials, graphics, simulations, and so on to be used in experimental interventions. • Experimental control assistance—software used to control the effects of specific variables and restrict the occurrence of other potentially confounding variables. • Survey distribution assistance—database use coupled with word processing to identify and send specific communications to a targeted population. • Data collection assistance—software used to take field notes or to monitor specific types of responses made by the participants in a research study.
Analysis and interpretation	• Organizational assistance—software used to assemble, categorize, code, integrate, and search potentially huge data sets (e.g., open-ended responses to survey questions, qualitative interview data). • Conceptual assistance—software used to write and store ongoing reflections about data or to construct theories that integrate research findings. • Statistical assistance—statistical and spreadsheet software packages used to categorize and analyze various types of data sets. • Graphic production assistance—software used to depict data in graphic form to facilitate interpretation.
Reporting	• Communication assistance—telecommunication software used to distribute and discuss research findings and initial interpretations with colleagues and to receive their comments and feedback. • Writing and editing assistance—word processing software used to write and edit successive drafts of the final report. • Publishing assistance—desktop publishing software used to produce professional-looking documents that can be distributed at conferences and elsewhere to get additional comments and feedback. • Distribution assistance—the Internet and other, more specific networks used to electronically distribute a report of one's findings and to generate discussion for follow-up studies by others in the field.

Measurement

Most researchers strive for objectivity: They believe that their observations should be influenced as little as possible—ideally not at all—by their own perceptions, impressions, and biases. (As we will note in Chapter 6, some qualitative researchers are an exception to this rule.) And one important way of remaining objective is to identify a systematic way of *measuring* a phenomenon being studied.

To measure something, of course, a researcher needs some sort of measurement instrument. Some common, everyday measurement instruments—rulers, scales, speedometers—can occasionally

be helpful in a research project. Such instruments are, of course, designed to measure easily observable physical characteristics, such as length, weight, or speed. More typically, however, a researcher needs more specialized instruments. For example, an astronomer might need a high-powered telescope to detect patterns of light in the night sky, and a neurophysiologist might need a magnetic resonance imaging (MRI) machine to detect and measure neural activity in the brain.

Yet social and psychological phenomena—phenomena that have no concrete physical basis—require measurement as well. For example, an economist might use the Dow-Jones or NASDAQ index to track economic growth over time, a sociologist might use a questionnaire to assess people's attitudes about marriage and divorce, and an educational researcher might use an achievement test to measure the extent to which children are learning at school.

Valid, reliable measurement instruments are critical for any research endeavor. Thus, we explore measurement strategies in some depth when we discuss the research planning process in Chapter 4.

Statistics

As data come to us from the real world, they are unorganized, separate bits of information. They have no focus; they need to be managed in some way. Statistics provide a means to get order out of chaos.

Statistics have two principal functions: to help the researcher (1) describe the data and (2) draw inferences from the data. **Descriptive statistics** summarize the general nature of the data obtained—for instance, how certain measured characteristics appear to be “on average,” how much variability exists among different pieces of data, how closely two or more characteristics are associated with one another, and so on. In contrast, **inferential statistics** help the researcher make decisions about the data; for instance, they might help a researcher decide whether the differences observed between two groups in an experiment are large enough to be attributed to the experimental intervention rather than to a once-in-a-blue-moon fluke. Both of these functions of statistics ultimately involve summarizing the data in some way.

Statistics are typically more useful in some academic disciplines than in others. For instance, researchers use them quite often in such fields as psychology, medicine, and business; they use statistics less frequently in such fields as history, musicology, and literature. But whenever we use statistics, we must remember that statistical values are not—and must not be—the final, ultimate goal of a research endeavor. The ultimate question in research is, *What do the data indicate?*, not *What is their numerical configuration?*—where do they cluster, how broadly do they spread, or how closely are they interrelated? Statistics give us *information* about the data, but a conscientious researcher is not satisfied until the *meaning* of this information is revealed.

Many beginning researchers erroneously think that calculating statistics is the final step in a research project, when in fact all they have done is to arrive at a few numbers that can help them interpret the data. Behind every statistic lies a sizable body of data; the statistic may summarize these data in a particular way, but it cannot capture all the nuances of the data. The entire body of data collected, not any single calculated statistic, is what ultimately must be used to resolve the research problem. There is no substitute for the task the researcher ultimately faces: to discover the meaning of the data and its relevance to the research problem. Any statistical process you may employ is merely ancillary to this central quest.

Furthermore, even the most sophisticated statistical procedures can never make amends for a poorly conceived research study. An editorial in the journal *Research in Nursing and Health* once made this point quite poignantly:

The use of elegant statistics can never compensate for inelegant conceptual bases. The new evaluative procedures are exciting because they enable examination of data in ways previously not possible. The bottom line remains the same, however. One cannot draw large savings out of an account into which little has been deposited. Neither can one draw useful meanings from studies into which less-than-important notions have been entered. (“Use of Elegant Statistics,” 1987, p. iii)

In the process of summarizing data, statistical analyses often create entities that have no counterpart in reality. Let's take a simple example: Four students have part-time jobs on campus.

One student works 24 hours a week in the library, a second works 22 hours a week in the campus bookstore, a third works 12 hours a week in the parking lot, and the fourth works 16 hours a week in the cafeteria. One way of summarizing the students' work hours is to calculate the arithmetic mean. By doing so, we find that the students work, "on average," 18.5 hours a week. Although we have learned something about these four students and their working hours, to some extent we have learned a myth: None of the four students has worked exactly 18.5 hours a week. That figure represents absolutely no fact in the real world.

Apparently, we have solved one problem only to create another. We have created a dilemma. If statistics offer us only an unreality, then why use them? Why create myth out of hard, demonstrable data? The answer lies in the nature of the human mind. Human beings can cognitively think about only so much information at any single point in time.³ Statistics help condense an overwhelming body of data into an amount of information that the mind can more readily comprehend and deal with. In the process, they can help the researcher detect patterns and relationships in the data that might otherwise go unnoticed. More generally, statistics *help the human mind comprehend disparate data as an organized whole*.

Although a book such as this one cannot provide all of the nitty-gritty details of statistical analysis, we give you an overview of potentially useful statistical techniques in Chapter 11.

Language

One of humankind's greatest achievements is language. Not only does it allow us to communicate with one another, but it also enables us to think more effectively. People can often think more clearly and efficiently about a topic when they can represent their thoughts in their heads with specific words and phrases.

For example, imagine that you are driving along a country road. In a field to your left, you see something with the following characteristics:

- Black and white in color, in a splotchy pattern
- Covered with a short, bristly substance
- Appended at one end by an object similar in appearance to a paintbrush
- Appended at the other end by a lumpy thing with four pointy objects sticking upward (two soft and floppy, two hard and curved around)
- Held up from the ground by four spindly sticks, two at each end

Unless you have spent most of your life living under a rock, you would almost certainly identify this object as a *cow*.

Words—even those as simple as *cow*—and the concepts that the words represent enhance our thinking in several ways (J. E. Ormrod, 2012; also see Jaccard & Jacoby, 2010):

1. *Words reduce the world's complexity.* Classifying similar objects and events into categories and labeling those categories in terms of specific words make our experiences easier to understand. For instance, it is much easier to think to yourself, "I see a herd of cows," than to think, "There is a brown object, covered with bristly stuff, appended by a paintbrush and a lumpy thing, and held up by four sticks. Ah, yes, and I also see a black-and-white spotted object, covered with bristly stuff, appended by a paintbrush and a lumpy thing, and held up by four sticks. And over there is a brown-and-white object. . . ."
2. *Words allow abstraction of the environment.* An object that has bristly stuff, a paintbrush at one end, a lumpy thing at the other, and several spindly sticks at the bottom is a concrete entity. The concept *cow*, however, is more abstract: It connotes such characteristics as *female*, *supplier of milk*, and, to the farmer or rancher, *economic asset*. Concepts and the labels associated with them allow us to think about our experiences without necessarily having to consider all of their concrete characteristics.

³If you have studied cognitive psychology, you may recognize that we are talking about the limited capacity of *working memory*.

3. *Words enhance the power of thought.* When you are thinking about an object covered with bristly stuff, appended by a paintbrush and a lumpy thing, held up by four sticks, and so on, you can think of little else (as mentioned earlier, human beings can think about only a very limited amount of information at any one time). In contrast, when you simply think *cow*, you can easily think about other ideas at the same time and perhaps form connections and interrelationships among them in ways you hadn't previously considered.
4. *Words facilitate generalization and inference drawing in new situations.* When we learn a new concept, we associate certain characteristics with it. Then, when we encounter a new instance of the concept, we can draw on our knowledge of associated characteristics to make assumptions and inferences about the new instance. For instance, if you see a herd of cattle as you drive through the countryside, you can infer that you are passing through either dairy or beef country, depending on whether you see large udders hanging down between some of the spindly sticks.

Just as *cow* helps us categorize certain experiences into a single idea, so, too, does the terminology of your discipline help you interpret and understand your observations. The words *tempo*, *timbre*, and *perfect pitch* are useful to the musicologist. Such terms as *central business district*, *folded mountain*, and *distance to k* have special meaning for the geographer. The terms *lesson plan*, *portfolio*, and *charter school* communicate a great deal to the educator. Learning the specialized terminology of your field is indispensable to conducting a research study, grounding it in prior theory and research, and communicating your results to others.

Two outward manifestations of language usage are also helpful to the researcher: (1) knowing two or more languages and (2) writing one's thoughts either on paper or in electronic form.

The value of knowing two or more languages It should go without saying that not all significant research is reported in a researcher's native tongue. Accordingly, many doctoral programs require that students demonstrate reading competency in one or two foreign languages in addition to their own language. The choice of these languages is usually linked to the area of proposed research.

The language requirement is a reasonable one. Research is and always has been a worldwide endeavor. For example, researchers in Japan have made gigantic strides in electronics and robotics. And two of the most influential theorists in child development today—Jean Piaget and Lev Vygotsky—wrote in French and Russian, respectively. Many new discoveries are first reported in a researcher's native language.

The importance of writing To be generally accessible to the larger scientific community and ultimately to society as a whole, all research must eventually be presented as a written document—a *research report*—either on paper or in electronic form. A basic requirement for writing such a report is the ability to use language in a clear, coherent manner.

Although the conventional wisdom is that clear thinking *precedes* clear writing, in fact writing can be a productive form of thinking in and of itself. When you write your ideas down on paper, you do several things:

- You must identify the specific ideas you do and do not know about your topic.
- You must clarify and organize your thoughts sufficiently to communicate them to your readers.
- You may detect gaps and logical flaws in your thinking.

Perhaps it is not surprising, then, that writing about a topic actually enhances the writer's understanding of the topic (Kellogg, 1994; Shanahan, 2004).

If you wait until all of your thoughts are clear before you start writing, you may never begin. Therefore we recommend that you start writing your research proposal or report as soon as possible. Begin with a title and a purpose statement for your study. Commit your title to paper; keep it in plain sight as you focus your ideas. Although you may very well change the title later as your research proceeds, creating a working title in the early stages can provide both focus and direction. And when you can draft a clear and concise statement that begins, "The purpose of this study is . . .," you are well on your way to planning a focused research study.

PRACTICAL APPLICATION Communicating Effectively through Writing

Judging from our own experience, most students have a great deal to learn about what good writing entails. Yet we authors also know that with effort, practice, expert guidance, and regular feedback, students *can* learn to write more effectively. Chapters 3, 5, and 12 present specific strategies for writing literature reviews, research proposals, and research reports. Here we offer general strategies for writing in ways that clearly communicate your ideas and reasoning to others. We also offer suggestions for using word processing software.

GUIDELINES Writing to Communicate

The following guidelines are based on techniques often seen in effective writing. Furthermore, such techniques have consistently been shown to facilitate readers' comprehension of what others have written (e.g., J. E. Ormrod, 2012).

1. *Say exactly what you mean.* Precision is of utmost importance in all aspects of a research endeavor, including writing. Choose your words and phrases carefully so that you communicate your *exact* meaning, not some vague approximation. Many books and other resources offer suggestions for writing clear, concise, and effective sentences and in combining those sentences into unified and coherent paragraphs (e.g., see the sources in the "For Further Reading" list at the end of the chapter).

2. *Continually keep in mind your primary objective in writing your paper, and focus your discussion accordingly.* All too often, novice researchers try to include everything they have learned—both from their literature review and from their data analysis—in their writing. But ultimately, everything you say should relate either directly or indirectly to your research problem. If you cannot think of how something relates, leave it out! You will undoubtedly have enough things to write about as it is.

3. *Provide an overview of what you will be talking about in upcoming pages.* Your readers can more effectively read your work when they know what to expect as they read. Providing an overview of what topics you will discuss and in what order—and possibly also showing how the various topics interrelate—is known as an **advance organizer**. As an example, a doctoral student in educational psychology, Dinah Jackson, was interested in the possible effects of *self-questioning*—that is, asking oneself questions about course material one is studying—on college students' note taking. Jackson began her dissertation's "Review of the Literature" with the following advance organizer:

The first part of this review will examine the theories, frameworks, and experimental research behind the research on adjunct questioning. Part two will investigate the transition of adjunct questioning to self-generated questioning. Specific models of self-generated questioning will be explored, starting with the historical research on question position [and progressing] to the more contemporary research on individual differences in self-questioning. Part three will explore some basic research on note taking, and tie note taking theory with the research on self-generated questioning. (Jackson, 1996, p. 17)

4. *Organize your ideas into general and more specific categories, and use headings and subheadings to guide your readers through your discussion of these categories.* Take a moment to flip through the pages of this book. Notice how often we use headings to let you know what we will be talking about in upcoming paragraphs. In our own experience, students often organize their thoughts (their literature reviews, for example) without communicating their organizational scheme to their readers. Using headings is one simple way to make that scheme crystal clear.

5. *Use concrete examples to make abstract ideas more understandable.* There is a fine line between being abstract and being vague. Even as scholars who have worked in our respective academic disciplines for many years, we authors still find that we can more easily understand something

when the writer gives us a concrete example to illustrate an abstract idea. As an example, we return to Jackson's dissertation on self-questioning and class note taking. Jackson makes the point that how a researcher evaluates, or *codes*, the content of students' class notes will affect what the researcher discovers about those notes. More specifically, she argues that a superficial coding scheme (e.g., counting the number of main ideas included in notes) fails to capture the true quality of the notes. She clarifies her point with a concrete example:

For example, while listening to the same lecture, Student A may record only an outline of the lecture, whereas Student B may record an outline, examples, definitions, and mnemonics. If a researcher only considered the number of main ideas that students included in their notes, then both sets of notes might be considered equivalent, despite the fact that the two sets differ considerably in the *type* of material recorded. (Jackson, 1996, p. 9)

6. *Use figures and tables to help you more effectively present or organize your ideas and findings.* Although the bulk of your research proposal or report will almost certainly be prose, in some cases it might be helpful to present some information in figure or table form. For example, as you read this book, look at the variety of mechanisms we use to accompany our prose, including art, diagrams, graphs, and summarizing tables. We hope you will agree that these mechanisms are critical in helping you understand and organize some of the ideas we present.

7. *At the conclusion of a chapter or major section, summarize what you have said.* Chances are, you will be presenting a great deal of information in any research proposal or report that you write. Summarizing what you have said in preceding paragraphs or pages helps your readers identify the things that are, in your mind, the most important things for them to remember. For example, in a dissertation that examined children's beliefs about the mental processes involved in reading, Debby Zambo summarized a lengthy discussion about the children's understanding of what it means to pay attention:

In sum, the students understand attention to be a mental process. They know their attention is inconsistent and affected by emotions and interest. They also realize that the right level of material, amount of information, and length of time helps their attention. The stillness of reading is difficult for some of the students but calming for others, and they appear to know this, and to know when reading will be difficult and when it will be calming. This idea is contrary to what has been written in the literature about struggling readers. (Zambo, 2003, p. 68)

8. *Anticipate that you will almost certainly have to write multiple drafts.* All too often, we authors have had students submit research proposals, theses, or dissertations with the assumption that they have completed what they set out to do. Such students have invariably been disappointed—sometimes even outraged—when we have asked them to revise their work, usually several times. The necessity to write multiple drafts applies not only to novice researchers but to experienced scholars as well. For instance, we would hate to count the number of times this book has undergone revision—certainly far more often than the label “tenth edition” indicates! Multiple revisions enable you to reflect on and critically evaluate your own writing, revise and clarify awkward passages, get feedback from peers and advisors who can point out where a manuscript lacks clarity, and spend more time ensuring that the final draft is as clear and precise as possible.

9. *Fastidiously check to be sure that your final draft uses appropriate grammar and punctuation, and check your spelling.* Appropriate grammar, punctuation, and spelling are not just bothersome formalities. On the contrary, they help you better communicate your meanings. For example, a colon announces that what follows it explains the immediately preceding statement; a semicolon communicates that a sentence includes two independent clauses (as the semicolon in this sentence does!).

Correct grammar, punctuation, and spelling are important for another reason as well: They communicate to others that you are a careful and well educated scholar whose thoughts and work are worth reading about. If, instead, you misspell many of your words—as we are doing in this sentence—your readers may quickly discredit you as a sloppy researcher who shouldn't be taken seriously!

Many style manuals, such as those in the “For Further Reading” list at the end of this chapter, have sections dealing with correct punctuation and grammar. And dictionaries and word processing spell-check functions can obviously assist you in your spelling.

GUIDELINES Using a Word Processor



Fortunately, word processing software makes the revision process infinitely easier than it was in the days of manual typewriters. Most word processing programs include the following features:

- **Editing features.** Common editing features allow you to enter information quickly, change wording, and delete unwanted letters, words, and paragraphs. As you examine what you have written, it is easy to move sections of text from one location to another. In general, editing features give the researcher more freedom to write, critically examine what has been written, and make modifications as necessary.
- **Formatting features.** Common formatting features provide control over how the words appear on the page. If special emphasis is needed, a word can be highlighted by changing the type size or by underlining, *italicizing*, or using **boldface**. Text can be arranged in columns with various types of margins and alignments. Tables can be set up easily with borders and shading to highlight information. Most word processing programs also let the writer insert graphics quickly and easily into a body of text.
- **Special editing features.** Several special features have proved invaluable to writers using word processors. These include an *outliner* to facilitate the initial planning and organization of the major sections of a writing project; a *spell checker* to call attention to and make suggestions for suspiciously spelled words; a *thesaurus* to help the writer identify potentially more appropriate words and phrases; a *grammar checker* to detect potential problems in how words have been put together; and a *track changes* feature that enables collaborating researchers to identify changes that others make to a coauthored paper.
- **Search-and-replace-features.** These features enable you to scan an entire document very quickly for certain contents—perhaps a particular word, phrase, date, or punctuation mark—and, if desired, replace it with something else. For example, imagine that you want to replace the term *autism* with the broader term *autism spectrum disorder*. A *search-and-replace* command enables you to make the switch for the entire document automatically or, if you prefer, to look at your uses of *autism* on a case-by-case basis. Or imagine, instead, that you discover that you have misunderstood when semicolons are appropriately used. You can search a document for all of its semicolons and, as necessary, change them to more appropriate punctuation.

Word processing software is an invaluable tool throughout the research process; in fact, we authors don't know how we lived without it for as long as we did. For example, as the *study is being planned*, word processing software can be used for brainstorming and organizing ideas. As *literature is being reviewed*, the software provides one means of recording and organizing what others have written related to the topic of investigation. As the *study is being implemented*, the software can be used to generate various types of data collection instruments and to transcribe people's responses to such instruments as interviews and questionnaires. As the *data are being analyzed*, tables and graphics can be developed to help categorize and summarize patterns in the data. Finally, as the *final report is being completed*, it can be written in the proper form for review and potential publication.

We offer three general recommendations for using a word processor effectively:

1. **Save your document frequently.** This seems like such an obvious point that we almost left it on the editing room floor, but then we remembered all the personal horror stories we have heard (and in some cases experienced ourselves) about losing data, research materials, and other valuable information. Every computer user eventually encounters some type of glitch that causes

problems in information retrieval. Whether the electricity goes out before you can save a file, a misguided keystroke leads to a system error, or your personal computer inexplicably crashes, data sometimes get lost. It is imperative that you get in the habit of saving your work. Save multiple copies so that if something goes awry in one place, you will always have a backup in a safe location. Here are a few things to think about:

- Save at least two copies of important files, and save them in different places—perhaps one file at home and another at the office, at a friend's house, or in a safe deposit box. One good option is to save everything on a **flash drive**, an inexpensive device about the size of a package of gum that can hold a mind-boggling amount of information. One of us authors uses a single flash drive to save all of her past work (including the manuscripts of more than 20 books) and any in-progress work; she keeps this flash drive in her purse and takes it everywhere she goes. If her house burns down (she certainly hopes that it won't!), she will at least be able to carry on with her professional life.
- Save your work-in-progress frequently, perhaps every 10 minutes or so; many software programs will do this for you automatically if you give them instructions about whether and how often to do it.
- Save various versions of your work with titles that help you identify each version—for instance, by including the date on which you completed each file.
- If your computer completely dies—seemingly beyond resuscitation—some software programs (e.g., Norton Utilities) may be able to fix the damage and retrieve some or all of the lost material.

2. *Use such features as the spell checker and grammar checker to look for errors, but do **not** rely on them exclusively.* Although computers are marvelous machines, their “thinking” capabilities have not yet begun to approach those of the human mind. For instance, although a computer can detect spelling errors, it does so by comparing each word against its internal “dictionary” of correctly spelled words. Not every word in the English language will be included in the dictionary; for instance, proper nouns (e.g., such surnames as Leedy and Ormrod) will *not* be. Furthermore, it may assume that *abut* is spelled correctly when the word you really had in mind was *about*, and it may very well not know that *there* should actually be *their* or *they're*.

3. *Print out a paper copy for final proofreading and editing.* One of us authors once had a student who turned in a dissertation draft that was chock-full of spelling and grammatical errors—and this from a student who was, ironically, teaching a college-level English composition course at the time. A critical and chastising e-mail message to the student made her irate; she had checked her document quite thoroughly before submitting it, she replied, and was convinced that it was virtually error-free. When her paper draft was returned to her almost bloodshot with spelling and grammatical corrections, she was quite contrite. “I don’t know how I missed them all!” she said. When asked if she had ever edited a printed copy of the draft, she replied that she had not, figuring that she could read her work just as easily on her computer monitor and thereby save a tree or two. But in our own experience, it is *always* a good idea to read a printed version of what you have written. For some reason, reading a paper copy often alerts us to errors we have previously overlooked on the computer screen.

The Human Mind

The research tools discussed so far—the library, computer technology, measurement, statistics, and language—are effective only to the extent that another critical tool comes into play as well. The human mind is undoubtedly the most important tool in the researcher's toolbox. Its functioning dwarfs all other gadgetry. Nothing equals its powers of comprehension, integrative reasoning, and insight.

Over the past few millennia, human beings have developed several general strategies through which they can more effectively reason about and better understand worldly phenomena. Key among these strategies are critical thinking, deductive logic, inductive reasoning, the scientific method, theory building, and collaboration with *other* minds.

Critical Thinking

Before beginning a research project, effective researchers typically look at research studies and theoretical perspectives related to their topic of interest. But they don't just accept research findings and theories at face value; instead, they scrutinize those findings and theories for faulty assumptions, questionable logic, weaknesses in methodology, inappropriate statistical analyses, and unwarranted conclusions. And, of course, effective researchers scrutinize their *own* work for the same kinds of flaws. In other words, good researchers engage in critical thinking.

In general, **critical thinking** involves evaluating the accuracy, credibility, and worth of information and lines of reasoning. Critical thinking is reflective, logical, and evidence-based. It also has a purposeful quality to it—that is, the researcher thinks critically in order to achieve a particular goal (Beyer, 1985; Halpern, 2008; Moon, 2008).

Critical thinking can take a variety of forms, depending on the context. For instance, it may involve any one or more of the following (Halpern, 1998, 2008; Nussbaum, 2008):

- **Verbal reasoning:** Understanding and evaluating persuasive techniques found in oral and written language.
- **Argument analysis:** Discriminating between reasons that do and do not support a particular conclusion.
- **Probabilistic reasoning:** Determining the likelihood and uncertainties associated with various events.
- **Decision making:** Identifying and evaluating several alternatives and selecting the alternative most likely to lead to a successful outcome.
- **Hypothesis testing:** Judging the value of data and research results in terms of the methods used to obtain them and their potential relevance to certain conclusions. When hypothesis testing includes critical thinking, it involves considering questions such as these:
 - Was an appropriate method used to measure a particular outcome?
 - Are the data and results derived from a relatively large number of people, objects, or events?
 - Have other possible explanations or conclusions been eliminated?
 - Can the results obtained in one situation be reasonably generalized to other situations?

To some degree, different fields of study require different kinds of critical thinking. In history, critical thinking might involve scrutinizing various historical documents and looking for clues as to whether things *definitely* happened a particular way or only *maybe* happened that way. In psychology, it might involve critically evaluating the way in which a particular psychological characteristic (e.g., intelligence, personality) is being measured. In anthropology, it might involve observing people's behaviors over an extended period of time and speculating about what those behaviors indicate about the society being studied.

Deductive Logic

Deductive logic begins with one or more *premises*. These premises are statements or assumptions that the researcher initially takes to be true. Reasoning then proceeds logically from these premises toward conclusions that—if the premises are indeed true—must *also* be true. For example,

If all tulips are plants, (Premise 1)
And if all plants produce energy through photosynthesis, (Premise 2)
Then all tulips must produce energy through photosynthesis. (Conclusion)

To the extent that the premises are false, the conclusions may also be false. For example,

If all tulips are platypuses, (Premise 1)
And if all platypuses produce energy through spontaneous combustion, (Premise 2)
Then all tulips must produce energy through spontaneous combustion. (Conclusion)

The if-this-then-that logic is the same in both examples. We reach an erroneous conclusion in the second example, however—we conclude that tulips are likely to burst into flames at unpredictable times—only because both of our premises are erroneous.

Let's look back more than 500 years to Christopher Columbus's first voyage to the New World. At the time, people held many beliefs about the world that, to them, were irrefutable facts: People are mortal, the earth is flat, and the universe is finite and relatively small. The terror that gripped Columbus's sailors as they crossed the Atlantic was a fear supported by deductive logic. If the earth is flat (premise) and the universe is finite and small (premise), then the earth's flat surface must stop at some point. Therefore, a ship that continues to travel into uncharted territory must eventually come to the earth's edge and fall off, and its passengers (who are mortal—another premise) will meet their deaths. The logic was sound; the conclusions were valid. Where the reasoning fell short was in a faulty premise: that the earth is flat.

Deductive logic provides the basis for mathematical proofs in mathematics, physics, and related disciplines. It is also extremely valuable for generating research hypotheses and testing theories. As an example, let's look one final time at doctoral student Dinah Jackson's dissertation project about the possible effects of self-questioning during studying. Jackson knew from well-established theories about human learning that forming mental associations among two or more pieces of information results in more effective learning than does trying to learn each piece of information separately from the others. She also found a body of research literature indicating that the kinds the questions people ask themselves (mentally) and try to answer as they learn (e.g., as they sit in class or read a textbook) affect what they learn and how effectively they remember it. (For instance, a student who is trying to answer the question, "What do I need to remember for the test?" might learn very differently from the student who is considering the question, "How might I apply this information to my own life?") Jackson's reasoning was as follows:

If learning information in an associative, integrative fashion is more effective than learning information piecemeal, (Premise 1)

If the kinds of questions students ask themselves during a learning activity influence how they learn, (Premise 2)

If training in self-questioning techniques influences the kinds of questions that students ask themselves, (Premise 3)

And if learning is reflected in the kinds of notes that students take during class, (Premise 4)

Then teaching students to ask themselves integrative questions as they study class material should lead to class notes that are more integrative in nature. (Conclusion)

Such reasoning led Jackson to form and test the following hypothesis:

Students who have formal training in integrative self-questioning will take more integrative notes than students who have not had any formal training. (Jackson, 1996, p. 12)

Happily for Jackson, the data she collected in her dissertation research supported her hypothesis.

Inductive Reasoning

Inductive reasoning begins not with a preestablished truth or assumption but instead with an observation. For instance, as a baby in a high chair many years ago, you may have observed that if you held a cracker in front of you and then let go of it, it fell to the floor. Hmm, you may have thought, what happens if I do that again? So you took another cracker from the tray on your high chair, held it in front of you, and released it. It, too, fell to the floor. You followed the same procedure with several more crackers, and the result was always the same: The cracker traveled in a downward direction. Eventually you may have performed the same actions on other things—blocks, rattles, peas, milk, and so on—and invariably observed the same result. You probably eventually drew the conclusion that all things fall when dropped—your first inkling about a force called *gravity*. (You may also have concluded that dropping things from your high chair greatly annoyed your parents, but that is another matter.)

In **inductive reasoning**, people use specific instances or occurrences to draw conclusions about entire classes of objects or events. In other words, they observe a sample and then draw conclusions about the population from which the sample has been taken. For instance, an anthropologist might draw conclusions about a certain culture after studying a particular community within that culture. A professor of special education might use a few case studies in which a particular instructional approach is effective with students who have autism to recommend that teachers use the instructional approach with other students who have autism. A sociologist might conduct three surveys (one in 1995, a second in 2005, a third in 2015) asking 1,000 people to describe their beliefs about AIDS and then drawing conclusions about how society's attitudes toward AIDS have changed over that time.

Figure 1.2 graphically depicts the nature of inductive reasoning. Let's look at an example of how this representation applies to an actual research project. Neurologists Silverman, Masland, Saunders, and Schwab (1970) sought the answer to a problem in medicine: How long can a person have a "flat EEG" (i.e., an absence of measurable electrical activity in the brain, typically indicative of cerebral death) and still recover? Silverman and his colleagues observed 2,650 actual cases. They noted that, in all cases in which the flat EEG persisted for 24 hours or more, not a single recovery occurred. All of the data pointed to the same conclusion: *It is unlikely that a recovery might take place for those who exhibit flat EEGs for a period of 24 hours or more.* We cannot, of course, rule out the unexplored cases, but *from the data observed*, the conclusion reached was that recovery seems impossible. The EEG line from *every* case led to that *one* conclusion.

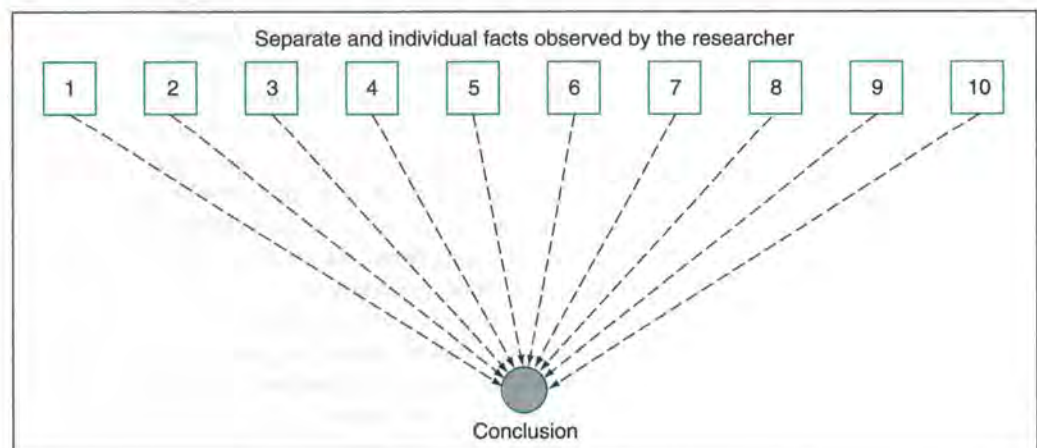
The Scientific Method

During the Renaissance, people found that when data are assembled and studied objectively and systematically, the data may yield previously undiscovered insights. Thus was the scientific method born; the words literally mean "the method that searches after knowledge" (*scientia* is Latin for "knowledge" and derives from *scire*, "to know"). The scientific method gained momentum during the 16th century with such men as Paracelsus, Copernicus, Vesalius, and Galileo.

Traditionally, the **scientific method** is a means whereby insight into the unknown is sought by (1) identifying a problem that defines the goal of one's quest; (2) positing a hypothesis that, if confirmed, resolves the problem; (3) gathering data relevant to the hypothesis; and (4) analyzing and interpreting the data to see whether they support the hypothesis and resolve the question that initiated the research.

Figure 1.1, which depicts research as a cyclical process, is a good illustration of the scientific method in action. We should keep in mind, however, that not all research methodologies follow the steps we have just listed in exactly that sequence. For instance, as you will discover when you read Chapter 6, such approaches as ethnographic research and grounded theory research involve collecting data and *then* developing one or more hypotheses about them. Rather than being a rigid, lock-step procedure, then, the scientific method is a somewhat flexible—although certainly also rigorous—process.

FIGURE 1.2
The inductive process



As you may already have realized, application of the scientific method typically involves both deductive logic and inductive reasoning. Researchers may develop a hypothesis either from a theory (deductive logic) or from observations of specific events (inductive reasoning). Then, using deductive logic, they make predictions about the patterns they are likely to see in the data *if* the hypothesis is true. And often, using inductive reasoning, they generalize from data taken from a sample to describe the characteristics of a larger population.

Theory Building

Psychologists are increasingly realizing that the human mind is a very *constructive* mind. People don't just take in and remember the innumerable pieces of information they acquire in a piece-meal fashion. Instead, they pull together what they learn about the world to form well-organized and integrated understandings about a wide variety of physical and social events. Human beings, then, seem to have a natural tendency to develop *theories* about the world around them (e.g., see Bransford, Brown, & Cocking, 2000; J. E. Ormrod, 2012).

In general, a **theory** is an organized body of concepts and principles intended to explain a particular phenomenon. Even as young children, human beings are inclined to form their own, personal theories about various physical and social phenomena—why the sun “goes down” at night, where babies come from, why certain individuals behave in particular ways, and so on. People's everyday, informal theories about the world are not always accurate. For example, imagine that as an airplane travels forward through the air, it drops a large metal ball. What kind of path will the ball take as it falls downward? The answer, of course, is that it will fall downward at an increasingly fast rate (thanks to gravity) but will also continue to travel forward (thanks to inertia). Thus, its path will have the shape of a parabolic arc. Yet many college students erroneously believe that the ball (a) will fall straight down, (b) will take a straight diagonal path downward, or (c) will actually move *backward* from the airplane as it falls down (McCloskey, 1983).

What distinguishes the theory building of a good researcher is that it is supported by well-documented findings—rather than by naive beliefs and subjective impressions of the world—and by logically defensible reasoning. Thus, the theory-building process involves thinking *actively* and *intentionally* about a phenomenon under investigation. Beginning with the facts known about the phenomenon, the researcher brainstorms ideas about plausible and, ideally, *best* explanations—a process that is sometimes called **abduction** (e.g., Jaccard & Jacoby, 2010; Walton, 2003). Such explanations are apt to involve an interrelated set of concepts and propositions that, taken together, can reasonably account for the phenomenon being studied.

After one or more researchers have developed a theory to explain a phenomenon of interest, the theory is apt to drive further research, in part by posing new questions that require answers and in part by suggesting hypotheses about the likely outcomes of particular investigations. For example, one common way of testing a theory is to use deductive reasoning to make a prediction (hypothesis) about what should occur *if the theory is a viable explanation of the phenomenon under study*. As an example, let's consider Albert Einstein's theory of relativity, first proposed in 1915. Within the context of his theory, Einstein hypothesized that light passes through space as photons—tiny masses of spectral energy. If light has mass, Einstein reasoned, it should be subject to the pull of a gravitational field. A year later, Karl Schwarzschild predicted that, based on Einstein's reasoning, the gravitational field of the sun should bend light rays considerably more than Isaac Newton had predicted many years earlier. In 1919 a group of English astronomers traveled to Brazil and North Africa to observe how the sun's gravity distorted the light of a distant star now visible due to a solar eclipse. After the data were analyzed and interpreted, the results clearly supported the Einstein-Schwarzschild hypothesis—and therefore also supported Einstein's theory of relativity.

As new data emerge that either do or do not support particular hypotheses, a researcher may continue to revise a theory, reworking parts to better account for research findings, filling in gaps with additional concepts or propositions, extending the theory to apply to additional situations, relating the theory to other theories regarding overlapping phenomena, and so on (Steiner, 1988; K. R. Thompson, 2006). Occasionally, when an existing theory cannot adequately account for a

growing body of evidence, a good researcher casts it aside and begins to formulate an alternative theory that better explains the data.

Theory building tends to be a relatively slow process, with any particular theory continuing to evolve over a period of years, decades, or centuries. Often, many researchers contribute to the theory-building effort, testing hypotheses that the theory suggests, suggesting additional concepts and propositions to include, and so on. This last point brings us to yet another strategy for effectively using the human mind: collaborating with *other* minds.

Collaboration with Other Minds

As an old saying goes, two heads are better than one. Typically, three or more heads are even better. Any single researcher is apt to have certain perspectives, assumptions, and theoretical biases—not to mention gaps in his or her knowledge about the subject matter—that will limit how he or she approaches a research project. By bringing one or more professional colleagues onto the scene—ideally, colleagues who have perspectives, backgrounds, and areas of expertise somewhat different from the researcher's own—the researcher brings many more cognitive resources to bear on how to tackle the research problem and how to find meaning in the data obtained (e.g., see Nichols, 1998).

Sometimes these colleagues enter the picture as equal partners. On other occasions they may simply offer suggestions and advice. For example, when a graduate student conducts research for a master's thesis or doctoral dissertation, the student is, of course, the key player in the endeavor. Yet the student typically has considerable guidance from an advisor and, especially in the case of a doctoral dissertation, from a faculty committee. The prudent student selects an advisor and committee members who have the expertise to help shape the research project into a form that will truly address the research question and—perhaps more importantly—will make a genuine contribution to the student's topic of study.

Collaborative interactions don't need to be face to face, of course. Technology offers many ways to facilitate collaboration across far distances. For example, we authors have found e-mail to be an excellent way to collaborate with colleagues in designing research studies and writing journal articles. One researcher will write a first draft of a proposed design or manuscript, send it to a *co-researcher* as an **attachment** to an e-mail message,⁴ who will revise and add to the document and either send it to a third collaborator or back to the first person for inspection and further editing, and so on.

As a general rule, productive researchers keep in regular communication with others who conduct similar research in their field, exchanging ideas, critiquing one another's work, and so on. Such ongoing communication is also a form of collaboration—albeit a less systematic one—in that everyone can benefit from and build on what others are thinking and finding. Here, too, technology plays important roles. For example, some researchers maintain professional **web pages** that describe their research programs and include links to relevant research reports; often you can find these web pages by going to the websites of the researchers' universities or other home institutions. Also of value are **list servers**, which provide a mechanism for electronic discussion groups. A list server (sometimes abbreviated as "listserv") is essentially a mailing list, and any e-mail message sent to it is distributed to everyone who has subscribed to the list. Thousands of list servers on a wide variety of topics are available for subscription, often without charge. Through them, people can easily communicate with one another about topics of common interest. For example, if you like music, you can subscribe to list servers that focus on any number of special musical interests. As e-mail messages are received by a particular list server, you will automatically receive a copy.

At various points in the book we present exercises to help you apply concepts and ideas we have presented. In the first of these exercises, which follows, we ask you to identify ways in which you might use technology to draw on the wisdom of or more directly collaborate with others.



⁴Depending on the e-mail software one uses, an attachment may instead be called an *enclosure*.

CONCEPTUAL ANALYSIS EXERCISE Using the Internet to Facilitate Communication and Collaboration with Others



Following are three scenarios. In each case, think about how the researcher might use the Internet to solve his or her problem. The answers appear after the “For Further Reading” list at the end of the chapter.

1. Arwin is a professor at a small college. Although his research is prominent in his field, few people on campus share his enthusiasm for his specialty—forensic pathology. Although Arwin avidly reads relevant academic journals, he looks forward to the annual meetings of his national organization, where he can exchange ideas with others who have similar interests. He wishes that such exchanges could occur more frequently.
2. Deirdre has a once-in-a-lifetime opportunity to spend 6 months in Australia collecting data about various marine plants of the Great Barrier Reef. Although she is excited about the opportunity, she realizes that the work of her campus research group will suffer. She wants to continue providing feedback on the group’s ongoing projects and papers.
3. Recently Alexis read about a new corrective eye procedure being investigated at a major medical research institution. The work is possibly relevant to her own research, but she has questions about the procedures and long-term results.

As the preceding sections should make clear, we human beings are—or at least have the potential to be—*logical, reasoning* beings. But despite our incredible intellectual capabilities—which almost certainly surpass those of all other species on the planet—we don’t always reason as logically or objectively as we might. For example, sometimes we “discover” what we *expect* to discover, to the point that we do not look objectively at the data we collect. And sometimes we are so emotionally attached to particular perspectives or theories about a phenomenon that we cannot abandon them when mountains of evidence indicate that we should. Figure 1.3 describes some common pitfalls in human reasoning—pitfalls we urge you to be on the lookout for and try to overcome. Good researchers are *reflective* researchers who regularly and critically examine not only their research designs and data but also their own thinking processes.

Reflections on Significant Research

The time: February 13, 1929. The place: St. Mary’s Hospital, London. The occasion: the reading of a paper before the Medical Research Club. The speaker: a member of the hospital staff in the Department of Microbiology. Such was the setting for the presentation of one of the most significant research reports of the early 20th century. The report was about a discovery that has transformed the practice of medicine. Dr. Alexander Fleming presented to his colleagues his research on penicillin. The group was apathetic. No one showed any enthusiasm for Fleming’s paper. Great research has frequently been presented to those who are imaginatively both blind and deaf.

Fleming, however, knew the value of what he had done. The first public announcement of the discovery of penicillin appeared in the *British Journal of Experimental Pathology* in 1929. It is a readable report—one that André Maurois (1959) called “a triumph of clarity, sobriety, and precision.” Get it; read it. You will be reliving one of the great moments in 20th-century medical research.

Soon after Fleming’s paper, two other names became associated with the development of penicillin: Ernst B. Chain and Howard W. Florey (Chain et al., 1940; also see Abraham et al., 1941). Together they developed a pure strain of penicillin. Florey was particularly instrumental in initiating its mass production and its use as an antibiotic for wounded soldiers in World War II (Coghill, 1944; also see Coghill & Koch, 1945). Reading these reports takes you back to

We human beings often fall short of the reasoning capacities with which Mother Nature has endowed us. Following are seven common pitfalls to watch for in your own thinking as a researcher.

1. **Confusing what must logically be true with what seems to be true in the world as we know it—a potential pitfall in deductive reasoning.** Our usual downfall in deductive reasoning is that we have trouble separating logic from everyday experience. For example, consider Isaac Newton's second law of motion: Force equals mass times acceleration ($F = ma$). According to this basic principle of Newtonian physics, any force applied to an object results in acceleration of the object. Using simple algebra—deductive reasoning at its finest—we can conclude that $a = F/m$ and therefore that if there is no acceleration ($a = 0$), then there is no force ($F = 0$), and vice versa. This deduction makes no sense to anyone who has ever tried to push a heavy object across the floor: The object may not move at all, let alone accelerate. What explains the object's stubbornness, of course, is that other forces, especially friction with and resistance from the floor, are counteracting any force that the pusher may be applying.
2. **Making generalizations about members of a category after having encountered only a restricted subset of that category—a potential pitfall in inductive reasoning.** The main weakness of inductive reasoning is that, even if all of our specific observations about a particular set of objects or events are correct, our generalizations about the category as a whole may *not* be correct. For example, if the only tulips we ever see are red ones, we may erroneously conclude that tulips can *only* be red. And if we conduct research about the political or religious beliefs of people who live in a particular location—say, people who live in Detroit—we may draw conclusions that don't necessarily apply to the human race as a whole. Inductive reasoning, then, is most likely to fall short when we gather data from only a small, limited sample.
3. **Looking only for evidence that supports our hypotheses, without also looking for evidence that would disconfirm our hypotheses.** We humans seem to be predisposed to look for confirming evidence rather than disconfirming evidence—a phenomenon known as **confirmation bias**. For many everyday practical matters, this approach serves us well. For example, if we flip a light switch and fail to get any light, we might immediately think, "The light bulb probably burned out." We unscrew the existing light bulb and replace it with a new one—and *voilà!* we now have light. Hypothesis confirmed, problem solved, case closed. However, truly objective researchers don't just look for evidence that confirms what they believe to be true. They also look for evidence that might *disprove* their hypotheses. They hope that they don't find such evidence, of course, but they look for it nevertheless.
4. **Confirming expectations even in the face of contradictory evidence.** Another aspect of our confirmation bias is that we tend to ignore or discredit any contradictory evidence that comes our way. For example, consider the topic of global climate change. Convincing evidence continues to mount to support the ideas that (a) the earth's average temperature is gradually rising and (b) this temperature rise is at least partly the result of carbon emissions and other human activities. Yet some folks have great difficulty looking at the evidence objectively—perhaps the researchers incorrectly analyzed the data, they say, or perhaps the scientific community has a hidden agenda and so is not giving us the straight scoop.
5. **Mistaking dogma for fact.** Although we might be inclined to view some sources of information with a skeptical, critical eye, we might accept others without question. For example, many of us willingly accept whatever an esteemed researcher, scholarly book, or other authority source says to be true. In general, we may uncritically accept anything said or written by individuals or groups we hold in high esteem. Not all authority figures and works of literature are reliable sources of information and guidance, however, and blind, unquestioning acceptance of them can be worrisome.
6. **Letting emotion override logic and objectivity.** We humans are emotional beings, and our emotions often infiltrate our efforts to reason and think critically. We are apt to think quite rationally and objectively when dealing with topics we do not feel strongly about and yet think in decidedly irrational ways about emotionally charged issues—issues we find upsetting, infuriating, or personally threatening.
7. **Mistaking correlation for causation.** In our efforts to make sense of our world, we human beings are often eager to figure out what causes what. But in our eagerness to identify cause-and-effect relationships, we sometimes "see" them when all we really have is two events that just happen to occur at the same time and place. Even when the two events are *consistently* observed together—in other words, when they are *correlated*—one of them does not necessarily cause the other. The ability for a researcher to distinguish between causation and correlation is a critical one, as you will discover in Chapters 8 and 11.

FIGURE 1.3

Common pitfalls in human reasoning

List of pitfalls based on Chapter 8, "Common Sense Isn't Always Sensible: Reasoning and Critical Thinking" in *Our Minds, Our Memories* by J. E. Ormrod, 2011, pp. 151–183. Copyright by Pearson Education, Inc. Used by permission.

the days when the medical urgency of dying people called for a massive research effort to make a newly discovered antibiotic available for immediate use.

October 25, 1945: The Nobel Prize in medicine was awarded to Fleming, Chain, and Florey.

If you wish to know more about the discovery of penicillin, read André Maurois's *The Life of Sir Alexander Fleming* (1959), the definitive biography done at the behest of Fleming's widow. The document will give you an insight into the way great research comes into being.

The procedures of great research are identical to those every student follows in doing a dissertation, a thesis, or a research report. All research begins with a problem, an observation, a question. Curiosity is the germinal seed. Hypotheses are formulated. Data are gathered.

Conclusions are reached. What *you* are doing in research methodology is the same as what has been done by those who have pushed back the barriers of ignorance and made discoveries that have greatly benefited humankind.

Exploring Research in Your Field

Early in the chapter we mentioned that academic research is popularly seen as an activity far removed from everyday living. Even graduate students working on theses or dissertations may consider their task to be meaningless busywork that has little or no relevance to the world beyond the university campus. This “busywork” conception of an academic program’s research requirement is simply not accurate. Conducting the research required to write an acceptable thesis or dissertation is one of the most valuable educational experiences a person can have. Furthermore, a good research project adds to our knowledge about our physical and social worlds and so can ultimately promote the welfare and well-being of ourselves and the planet as a whole.

Even if you plan to become a practitioner rather than a researcher—say, a nurse, social worker, or school principal—knowledge of strong research methodologies and appropriate ways to collect and analyze data is essential for keeping up with advances in your field. The alternative—that is, *not* being well versed in sound research practices—can lead you to base important professional decisions on faulty data, inappropriate interpretations and conclusions, or unsubstantiated personal intuitions. Truly competent and effective practitioners base their day-to-day decisions and long-term priorities on solid research findings in their field.

As a way of getting your feet wet in the world of research, take some time to read articles in research journals in your own academic discipline. You can do so by spending an hour or two in the *periodicals* section of your local college or university library. Some libraries organize these journals alphabetically by title. Others organize them using the Library of Congress classification system, which allows journals related to the same topic to be placed close together (more about the Library of Congress system in Chapter 3).

Your professors should have suggestions about journals that are especially relevant to your academic discipline. Reference librarians can be helpful as well. In addition, especially if you are shy about asking other people for advice, you can get insights about important journals by scanning the reference lists in textbooks in your discipline.

Browse the journals related to your field just to get acquainted with them. Go first to those that pique your interest and skim a few studies that relate to particularly intriguing topics. Then, get acquainted with as many of the journals in your discipline as you can. Competent researchers have general knowledge of the resources available in their field.

Keep in mind that the quality of research you find in your explorations may vary considerably. One rough indicator of the quality of a research study is whether the research report has been *juried* or *nonjuried*. A **juried** (or *refereed*) research report has been judged by respected colleagues in one’s field and deemed to be of sufficient quality and importance to warrant publication. For instance, the editors of many academic journals send submitted manuscripts to one or more reviewers who pass judgment on the manuscripts, and only manuscripts that meet certain criteria are published in the journal. A **nonjuried** (or *nonrefereed*) report is one that appears in a journal or on the Internet without first being screened by one or more experts. Some nonjuried reports are excellent, but others may not be.

PRACTICAL APPLICATION Identifying Important Tools in Your Discipline

In this chapter we have introduced several key tools used by researchers as they go about their work. These tools can, of course, be effective and helpful only to the extent that they are used—and used correctly.

Some of the tools you learn about in this book may be somewhat new to you. How will you learn when, how, and why you should use them? One effective means of learning about research tools is to work closely with an expert researcher in your field. Watch and observe this person in action as he or she uses a variety of research tools.

Take the time to find a person who has completed a few research projects—perhaps someone who teaches a research methods class, someone who has published in several journals, someone who has successfully obtained research grants, or even someone who has recently finished a dissertation. Ideally this individual should be someone in your own field of study. Ask the questions listed in the following checklist and, if possible, observe the person as he or she goes about research work. If you cannot locate anyone locally, it may be possible to contact one or more persons through e-mail.

CHECKLIST

Interviewing an Expert Researcher

- _____ 1. How do you start a research project?

- _____ 2. What specific tools do you use (e.g., library resources, computer software, forms of measurement, statistics)?

- _____ 3. How did you gain your expertise with the various tools you use?

- _____ 4. What are some important experiences you suggest for a novice researcher?

- _____ 5. If I wanted to learn how to become a competent researcher, what specific tools would you suggest I work with?

For Further Reading

General Research Design

- Bouma, G. D., & Ling, R. (2004). *The research process* (5th ed.). New York: Oxford University Press.
- Goodwin, C. J. (2007). *Research in psychology: Methods and design* (5th ed.). New York: Wiley.
- Luczun-Friedman, M. E. (1986). Introduction to research: A basic guide to scientific inquiry. *Journal of Post Anesthetic Nursing*, 1, 64–75.
- McMillan, J. H., & Wergin, J. F. (2010). *Understanding and evaluating educational research* (4th ed.). Upper Saddle River, NJ: Pearson.
- Nieswiadomy, R. M. (2012). *Foundations in nursing research* (6th ed.). Upper Saddle River, NJ: Prentice Hall.
- Repko, A. F. (2008). *Interdisciplinary research*. Thousand Oaks, CA: Sage.
- Rosnow, R. L., & Rosenthal, R. (2008). *Beginning behavioral research: A conceptual primer* (6th ed.). Upper Saddle River, NJ: Pearson.

- Edyburn, D. L. (1999). *The electronic scholar: Enhancing research productivity with technology*. Upper Saddle River, NJ: Merrill/Prentice Hall.
- Mills, S. C., & Roblyer, M. D. (2006). *Technology tools for teachers: A Microsoft Office tutorial* (2nd ed.). Upper Saddle River, NJ: Allyn & Bacon/Pearson.

Effective Writing and Word Processing

- American Psychological Association. (2010). *Concise rules of APA style* (6th ed.). Washington, DC: Author.
- American Psychological Association (APA). (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.
- Beebe, L. (Ed.). (1992). *Professional writing for the human services*. Washington, DC: National Association of Social Workers Press.
- Chicago manual of style* (16th ed.). (2010). Chicago: University of Chicago Press.
- Council of Science Editors. (2006). *Scientific style and format: The CSE manual for authors, editors, and publishers* (7th ed.). Reston, VA: Author.

Using Technology in Research

- Angell, D. (1994). *The elements of e-mail style: Communicate effectively via electronic mail*. Reading, MA: Addison-Wesley.

- Flesch, R. (1974). *The art of readable writing*. New York: Harper & Row.
- Modern Language Association. (2008). *MLA style manual and guide to scholarly publishing* (3rd ed.). New York: Author.
- Sternberg, R. J. (1977). *Writing the psychology paper*. Woodbury, NY: Barron's Educational Series.
- Strunk, W. (2007). *The elements of style* (4th ed.). New York: Penguin.
- Strunk, W., & White, E. B. (2008). *The elements of style*. New York: Longman.
- Williams, J. M. (2006). *Style: Ten lessons in clarity and grace* (9th ed.). Reading, MA: Addison-Wesley Longman.
- Bickle, J., Mauldin, R., & Giere, R. N. (2005). *Understanding scientific reasoning* (5th ed.). New York: Wadsworth.
- Carey, S. S. (2012). *A beginner's guide to scientific method* (4th ed.). Florence, KY: Cengage.
- Jaccard, J., & Jacoby, J. (2010). *Theory construction and model-building skills*. New York: Guilford.
- Jung, S. (1995). *The logic of discovery: An interrogative approach to scientific inquiry*. New York: Peter Lang.
- Poplin, M. S. (1987). Self-imposed blindness: The scientific method in education. *Remedial and Special Education*, 8(6), 31–37.
- Popper, K. (2002). *The logic of scientific discovery*. New York: Routledge.
- Shank, G. D. (2006). *Qualitative research: A personal skills approach* (2nd ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

[See Chapter 7]

Reasoning, the Scientific Method, and Theory Building

- Bicak, L. J., & Bicak, C. J. (1988). Scientific method: Historical and contemporary perspectives. *American Biology Teacher*, 50, 348–353.

ANSWERS TO THE CONCEPTUAL ANALYSIS EXERCISE "Using the Internet to Facilitate Communication and Collaboration with Others" on page 22:

1. Subscribing to one or more list servers might enable Arwin to keep up with current developments in his field and communicate regularly with others who are conducting research in his area of interest.
2. During her time in Australia, Deirdre can stay in regular communication with her colleagues at home via e-mail, and she can easily send papers back and forth as attachments.
3. To gain information about the corrective procedure in which she is interested, Alexis might visit the website of the hospital where the research is taking place. On the website she may find additional information about the procedure and the e-mail addresses of the individuals conducting the research.

Now go to MyEducationLab at www.myeducationlab.com to take assessments to evaluate your mastery of chapter concepts. Assignments and Activities exercises and Building Skills in Reading Research exercises are also available to help you master concepts and skills. Feedback for many of these exercises is provided so that you can see why your answers are correct or incorrect.

2

The Problem: The Heart of the Research Process

The problem or question is the axis around which the whole research effort revolves. The statement of the problem must first be expressed with the utmost precision; it should then be divided into more manageable subproblems. Such an approach clarifies the goals and directions of the entire research effort.

The heart of every research project—the axis around which the entire research endeavor revolves—is the problem. The first step in the research process is to identify the problem with unwavering clarity and to state it in precise and unmistakable terms.

Finding Research Projects

Problems in need of research are everywhere. To get an idea of typical research projects for doctoral dissertations, go to the reference room at your university library, open any volume of *Dissertation Abstracts International*—most university libraries also have these abstracts in an online database—and look at the dissertation abstracts in your academic discipline. To get an online sample of recently published research studies in your area of interest, go to Google Scholar at scholar.google.com; type a topic in the search box and then click on some of the titles that pique your curiosity. As you scan the results of your Google search, especially look for items labeled as **pdf** (referring to “portable document format”); these items are essentially electronic photocopies of articles that have appeared in academic journals and similar sources.

Some research projects are intended to enhance basic knowledge about our physical, biological, psychological, or social world or to shed light on historical, cultural, or aesthetic phenomena. For example, an ornithologist might study the mating habits of a particular species of birds, and a psychologist might study the nature of people’s logical reasoning processes. Such projects, which can advance theoretical conceptualizations about a particular topic, are known as **basic research**.

Other research projects are intended to address issues that have immediate relevance to our society’s current practices, procedures, and policies. For example, a nursing educator might compare the effectiveness of different instructional techniques for training future nurses, and an agronomist might study the effects of various fertilizers on the growth of sunflowers. Such projects, which can inform human decision making about practical problems, are known as **applied research**. Occasionally applied research involves addressing questions in one’s immediate work environment, with the goal of solving an ongoing problem in that environment; such research is known as *action research* (e.g., Cochran-Smith & Lytle, 1993; Mertler, 2009; Mills, 2011).

Keep in mind, however, that the line between basic research and applied research is, at best, a blurry one. Answering questions about basic theoretical issues can often inform current practices in the everyday world; for example, by studying the mating habits of a particular species of birds, an ornithologist might lead the way in saving that species from extinction. Similarly, answering questions about practical problems may enhance theoretical understandings of particular phenomena; for example, the nursing educator who finds that one approach to training nurses is more effective than another may enhance psychologists’ understanding of how, in general, people acquire new knowledge and skills.

Regardless of whether you conduct basic or applied research, a research project is likely to take a significant amount of your time and energy, so whatever problem you study should be *worth* your time and energy. As you begin the process of identifying a suitable research problem to tackle, keep two criteria in mind. First, your problem should address an important question, such that the answer can actually “make a difference” in some way. And second, it should advance the frontiers of knowledge, perhaps by leading to new ways of thinking, suggesting possible applications, or paving the way for further research in the field. To accomplish both of these ends, your research project must involve not only the collection of data but also the *interpretation* of those data.

Some problems are not suitable for research because they lack the interpretation-of-data component; they do not require the researcher to go beyond the data themselves and reveal their meaning. Following are four situations to avoid when considering a problem for research purposes:

1. *Research projects should not be simply a ruse for achieving self-enlightenment.* All of us have large gaps in our education, and filling them is perhaps the greatest joy of learning. But self-enlightenment is not the primary purpose of research. Gathering information to know more about a certain area of knowledge is entirely different from looking at a body of data to discern how it contributes to the solution of the problem.

A student once submitted the following as the statement of a research problem:

The problem of this research is to learn more about the way in which the Panama Canal was built.

For this student, the information-finding effort would provide the satisfaction of having gained more knowledge about a particular topic, but it would *not* have led to *new* knowledge.

2. *A problem whose sole purpose is to compare two sets of data is not a suitable research problem.* Take this proposed problem for research:

This research project will compare the increase in the number of women employed over 100 years—from 1870 to 1970—with the employment of men over the same time span.

A simple table completes the project (*Historical Statistics*, 1975).

	1870	1970
Women employed	13,970,000	72,744,000
Men employed	12,506,000	85,903,000

The “research” project involves nothing more than a quick trip to the library to reveal what is already known.

3. *Simply calculating a correlation coefficient between two sets of data to show a relationship between them is not acceptable as a problem for research.* Why? Because the basic requirement for research is ignored: a human mind struggling to make sense of data. What we see here is a proposal to perform a statistical operation that a computer can do infinitely faster and more accurately than a human being can. A correlation coefficient is nothing more than a statistic that expresses how closely two characteristics or other variables are associated with each other. It tells us nothing about *why* the association might exist.

Some novice researchers think that their work is done when they collect data and, by using a simple statistical procedure, find that two variables are closely related. In fact, their work is *not* done at this point; it has only begun. For example, many researchers have found a correlation between the IQ scores of children and those of their parents. In and of itself, this fact has very little usefulness. It does, however, suggest a problem for research: What is the *cause* of the relationship between children’s and parents’ intelligence test scores? Is it genetic? Is it environmental? Is it a combination of both genetic heritage and environment?

4. *Problems that result only in a yes or no answer are not suitable problems for research.* Why? For the same reason that merely finding a correlation coefficient is unsatisfactory. Both situations simply skim the surface of the phenomenon under investigation, without exploring the mechanisms underlying it.

"Is homework beneficial to children?" That is no problem for research, at least not in the form in which it is stated. The researchable issue is not whether homework is beneficial, but wherein the benefit of homework—if there is one—lies. Which components of homework are beneficial? Which ones, if any, are counterproductive? If we knew the answers to these questions, then teachers could better structure homework assignments to enhance students' learning and classroom achievement.

There is so much to learn—there are so many important questions unanswered—that we should look for significant problems and not dwell on those that will make little or no contribution. When asked about conducting research, Peter Medawar (1979), a Nobel laureate who investigated causes of the human body's rejection of other people's transplanted organs and tissues, gave wise advice to the young scientist:

It can be said with complete confidence that any scientist of any age who wants to make important discoveries must study important problems. Dull or piffling problems yield dull or piffling answers. It is not enough that a problem should be "interesting"—almost any problem is interesting if it is studied in sufficient depth. (p. 13)

Good research, then, begins with identifying a good question to ask—ideally a question that no one has ever thought to ask before. Researchers who contribute the most to our understanding of our physical, biological, psychological, and social worlds are those who pose questions that lead us into entirely new lines of inquiry. To illustrate, let's return to that correlation between the IQ scores of children and those of their parents. For many years, psychologists bickered about the relative influences of heredity and environment on intelligence and other human characteristics. They now know not only that heredity and environment *both* influence virtually every aspect of human functioning but also that they *influence each other's influences* (for a good, down-to-earth discussion of this point, see Lippa, 2002). Rather than ask the question "How much do heredity and environment each influence human behavior?" a more fruitful question—one that is relatively new on the scene—is "How do heredity and environment *interact* in their influences on behavior?"

PRACTICAL APPLICATION Identifying and Describing the Research Problem

How can the beginning researcher formulate an important and useful research problem? Here we offer guidelines both for identifying a particular problem and for describing it in precise terms.

GUIDELINES Finding a Legitimate Problem

As a general rule, appropriate research projects don't fall from trees and hit you on the head. You must be sufficiently knowledgeable about your topic of interest to know what projects might make important contributions to the field. Following are several strategies that are often helpful for novice and expert researchers alike.

1. *Look around you.* In many disciplines, questions that need answers—phenomena that need explanation—are everywhere. For example, let's look back to the early 17th century, when Galileo was trying to make sense of a variety of earthly and celestial phenomena. For example, why did large bodies of water (but not small ones) rise and fall in the form of tides twice a day? Why did sunspots consistently move across the sun's surface from right to left, gradually disappear, and then, about two weeks later, reappear on the right edge? Furthermore,

why did sunspots usually move in an upward or downward path as they traversed the sun's surface, while only occasionally moving in a direct, horizontal fashion? Galileo correctly deduced that the various "paths" of sunspots could be explained by the facts that both the earth and sun were spinning on tilted axes and that (contrary to popular opinion at the time) the earth revolved around the sun rather than vice versa. Galileo was less successful in explaining tides, attributing them to natural "sloshing" as a result of the earth's movement through space, rather than (more accurately) to the moon's gravitational pull (Sobel, 2000).

We do not mean to suggest that novice researchers should take on such monumental questions as the nature of the solar system or oceanic tides. But smaller problems suitable for research exist everywhere. Perhaps you might see them in your professional practice or in everyday events. Continually ask yourself questions about what you see and hear: Why does such-and-such happen? What makes such-and-such tick? and so on.

2. *Read the literature.* One essential strategy is to find out what things are already known about your topic of interest—a topic we address in more detail in Chapter 3. Little can be gained by reinventing the wheel. In addition to telling you what is already known, the existing literature is likely to tell you what is *not* known in the area—in other words, what still needs to be done. For instance, your research project might

- Address the suggestions for future research that another researcher has identified
- Replicate a research project in a different setting or with a different population
- Consider how various subpopulations might behave differently in the same situation
- Apply an existing perspective or theory to a new situation
- Explore unexpected or contradictory findings in previous studies
- Challenge research findings that seem to contradict what you know or believe to be true (Neuman, 2011)

Reading the literature has other advantages as well. It gives you a theoretical base on which to generate hypotheses and build a rationale for your study. It offers potential research designs and methods of measurement. And it can help you interpret your results and relate them to what is already known in the field.

As you read about other people's research related to your topic, *take time to consider how you can improve your own work because of it.* Ask yourself: What have I learned that I would (or would not) want to incorporate into my own research? Perhaps it is a certain way of writing, a specific method of data collection, or a particular approach to data analysis. You should constantly question and reflect on what you read.

We urge you, too, to *keep a running record of helpful journal articles and other sources.* Include enough information that you will be able to track each source down again—perhaps including the author's name, the title and year of the journal or book, key words and phrases that capture the focus of the work, and (if applicable) the appropriate library call number or Internet address. You may think that you will always be able to recall where you found a helpful source and what you learned from it. However, our own experiences tell us that you probably *will* forget a good deal of what you read unless you keep a written record of it.

3. *Seek the advice of experts.* Another simple yet highly effective strategy for identifying a research problem is to ask an expert: What needs to be done? What burning questions are still out there? What previous research findings seemingly don't make sense? Your professors will almost certainly be able to answer each of these questions, as will other scholars you might contact through e-mail or meet at conferences or elsewhere.

Some beginning researchers—including many students—are reluctant to approach well-known scholars for fear that these scholars do not have the time or patience to talk with novices. Quite the opposite is true: Most experienced researchers are happy to talk with people who are just starting out. In fact, they may feel flattered that you are familiar with their work and that you would like to extend or apply it in some way.

4. *Attend professional conferences.* Many researchers have great success finding new research projects at national or regional conferences in their discipline. By scanning the conference program and attending sessions of interest, they can learn "what's hot and what's not" in their field.

Furthermore, conferences are a place where novice researchers can make contacts with more experienced individuals in their field—where they can ask questions, share ideas, and exchange e-mail addresses that enable follow-up communication.

5. *Choose a topic that intrigues and motivates you.* As you read the professional literature, attend conferences, and talk with experts, you will uncover a number of potential research problems. At this point, you need to pick just *one* of them, and your selection should be based on what you personally want to learn more about. Remember, the project you are about to undertake will take you many months, quite possibly a couple of years or even longer. So it should be something you believe is worth your time and effort—even better, one you are truly passionate about. Peter Leavenworth, at the time a doctoral student in history, explained the importance of choosing an interesting dissertation topic this way: “You’re going to be married to it for a while, so you might as well enjoy it.”

6. *Choose a topic that others will find interesting and worthy of attention.* Ideally, your work should not end simply with a thesis, dissertation, or other unpublished research report. If your research adds an important piece to what the human race knows and understands about the world, then you will, we hope, want to share your findings with a larger audience. In other words, you will want to present what you have done at a regional or national conference, publish an article in a professional journal, or both (we talk more about doing such things in Chapter 12). Conference coordinators and journal editors are often quite selective about the research reports they accept for presentation or publication, and they are most likely to choose those reports that will have broad appeal.

Future employers, too, may make judgments about you, at least in part, based on the topic you have chosen for a thesis or dissertation. Your résumé or curriculum vitae will be more apt to attract their attention if, in your research, you are pursuing an issue of broad scientific or social concern or, more generally, a hot topic in your field.

GUIDELINES Stating the Research Problem

Remember, the heart of any research project is the problem. At every step in the process, successful researchers ask themselves: What am I doing? For what purpose am I doing it? Such questions can help you focus your efforts toward achieving your ultimate purpose for gathering data: to resolve the problem.

Researchers get off to a strong start when they begin with an unmistakably clear statement of the problem. After identifying a research problem, therefore, you must articulate it in such a way that *it is carefully phrased and represents the single goal of the total research effort*. Following are some general guidelines to help you do exactly that.

1. *State the problem clearly and completely.* Your problem should be so clearly stated that anyone who reads English can read and understand it. If the problem is not stated with such clarity, you are deceiving yourself that you know what the problem is. Such self-deception will cause you difficulty later on.

You can state your problem clearly only when you also state it completely. At a minimum, you should describe it in one or more *grammatically complete sentences*. As examples of what *not* to do, following are some meaningless half-statements—verbal fragments that only hint at the problem. Ask yourself whether you understand exactly what each student researcher plans to do.

From a student in sociology:

Welfare on children’s attitudes.

From a student in music:

Palestrina and the motet.

From a student in economics:

Busing of schoolchildren.

From a student in social work:

Retirement plans of adults.

All four statements lack clarity. It is imperative to think in terms of specific, researchable goals expressed in complete sentences. We take the preceding fragments and develop each of them into one or more complete sentences that describe a researchable problem.

Welfare on children's attitudes becomes:

What effect does welfare assistance to parents have on the attitudes of their children toward work?

Palestrina and the motet becomes:

This study will analyze the motets of Giovanni Pierluigi da Palestrina (1525?–1594) written between 1575 and 1580 to discover their distinctive contrapuntal characteristics and will contrast them with the motets of his contemporary William Byrd (1542?–1623) written between 1592 and 1597. During the periods studied, each composer was between 50 and 55 years of age.

Busing of schoolchildren becomes:

What factors must be evaluated and what are the relative weights of those several factors in constructing a formula for estimating the cost of busing children in a mid-western metropolitan school system?

Retirement plans for adults becomes:

How do retirement plans for adults compare with the actual realization, in retirement, of those plans in terms of self-satisfaction and self-adjustment? What does an analysis of the difference between anticipation and realization reveal for a more intelligent approach to planning?

Notice that, in the full statement of each of these problems, the areas studied are carefully limited so that the study is of manageable size. The author of the Palestrina-Byrd study carefully limited the motets that would be studied to those written when each composer was between 50 and 55 years of age. A glance at the listing of Palestrina's works in *Grove's Dictionary of Music and Musicians* demonstrates how impractical it would be for a student to undertake a study of all the Palestrina motets. He wrote 392 of them!

2. *Think through the feasibility of the project that the problem implies.* Novice researchers sometimes rush into a problem without thinking through its implications. It's great to have ideas. It's much better to have practical ideas. Before your enthusiasm overtakes you, consider the following research proposal submitted by John:

This study proposes to study the science programs in the secondary schools in the United States for the purpose of . . .

Let's think about that. The United States has more than 40,000 public and private secondary schools. These schools, north to south, extend from Alaska to Florida; east to west, from Maine to Hawaii. Certain practical questions immediately surface. How does John intend to contact each of these schools? By personal visit? Being very optimistic, he might be able to visit two schools per day, thus requiring more than 20,000 visitation days. The number of school days in the average school year is 180, so under the best of circumstances it would take more than one hundred years for John to gather the data. And even if John had exceptional longevity—not to mention exceptional persistence—the financial outlay for his project would be exorbitant.

"But," John explains, "I plan to gather the data by mail with a questionnaire." Fine! Each letter to the more than 40,000 schools, with an enclosed questionnaire and a return postage-paid envelope, would probably cost at least a dollar just for the postage. And we mustn't overlook the fact that John would need a second and perhaps a third mailing. A 50% return on the first mailing would be considered a good return. But for the nonreturnees, one or more follow-up mailings would be required. And we would need to figure in the cost of envelopes, stationery, photocopying, and data analysis.

A faster and less expensive option, of course, would be to conduct the survey by e-mail. In that case, John would need to track down the name and chief administrator of every one of those 40,000-plus schools. How long might it take him to do that? And how many of his e-mail messages might end up in a chief administrator's spam filter and thus never be read?

Obviously, John did not intend to survey every secondary school in the United States, yet that is what he wrote that he would do.

3. *Say precisely what you mean.* When you state your research problem, you should say exactly what you mean. You cannot assume that others will be able to read your mind. People will always take your words at their face value: You mean what you say—that's it. In the academic community, a basic rule prevails: *Absolute honesty and integrity are assumed in every statement a scholar makes.*

Look again at John's problem statement. We could assume that John means to fulfill precisely what he has stated (although we would doubt it, given the time and expense involved). Had he intended to survey only *some* schools, he should have said so plainly:

This study proposes to survey the science programs in *selected secondary schools throughout the United States.*

Or perhaps he could have limited his study to a specific geographical area or to schools serving certain kinds of students. Such an approach would give the problem constraints that the original statement lacked and would communicate to others what John intended to do—what he realistically *could* commit to doing. Furthermore, it would have preserved his reputation as a researcher of integrity and precision.

Ultimately, an imprecisely stated research problem can lead others to have reservations about the quality of the overall research project. If a researcher cannot be meticulous and precise in stating the nature of the problem, others might question whether the researcher is likely to be any more meticulous and precise in gathering and interpreting data. Such uncertainty and misgivings are serious indeed, for they reflect on the basic integrity of the whole research effort.

We have discussed some common difficulties in the statement of the problem, including statements that are unclear or incomplete and statements that suggest impractical or impossible projects. Another difficulty is this one: A researcher *talks about the problem* but never actually *states what the problem is*. Using the excuse that the problem needs an introduction or needs to be seen against a background, the researcher launches into a generalized discussion, continually obscuring the problem, never clearly articulating it. Take, for example, the following paragraph that appeared under the heading "Statement of the Problem":

The upsurge of interest in reading and learning disabilities found among both children and adults has focused the attention of educators, psychologists, and linguists on the language syndrome. In order to understand how language is learned, it is necessary to understand what language is. Language acquisition is a normal developmental aspect of every individual, but it has not been studied in sufficient depth. To provide us with the necessary background information to understand the anomaly of language deficiency implies a knowledge of the developmental process of language as these relate to the individual from infancy to maturity. Grammar, also an aspect of language learning, is acquired through pragmatic language usage. Phonology, syntax, and semantics are all intimately involved in the study of any language disability.

Can you find a statement of problem here? Several problems are suggested, but none is articulated with sufficient clarity that we might put a finger on it and say, "There, that is the problem."

Earlier in this chapter, we invited you to go to *Dissertation Abstracts International* to see how the world of research and the real world of everyday living are intertwined. Now return to those abstracts and notice with what directness the problems are set forth. The problem should be stated in the very first words of an abstract: "The purpose of this study is to . . ." No mistaking it, no background buildup necessary—just a straightforward plunge into the task at hand. All research problems should be stated with the same clarity.

4. *State the problem in a way that reflects an open mind about its solution.* In our own research methods classes, we have occasionally seen research proposals in which the authors state that they intend to *prove* that such-and-such a fact is true. For example, a student once proposed the following research project:

In this study, I will prove that obese adults experience greater psychological distress than adults with a healthy body mass index.

This is not a research question; it is a presumed—and quite presumptuous!—*answer* to a research question. If this student already knew the answer to her question, why was she proposing to study it? Furthermore, as noted in Chapter 1, it is quite difficult to prove something definitively, beyond a shadow of a doubt. We can certainly obtain data consistent with what we believe to be true, but in the world of research we can rarely say with 100% certainty that it *is* true.

Good researchers try to keep open minds about what they might find. Perhaps they will find the result they hope to find, perhaps not. Any hypothesis should be stated as exactly that—a *hypothesis*—rather than as a foregone conclusion. As you will see later in the chapter, hypotheses certainly do have their place in a research proposal. However, they should not be part of the problem statement.

Let's rewrite the preceding research problem, this time omitting any expectation of results that the research effort might yield:

In this study, I will investigate the possible relationship between body mass index and psychological stress, as well as two more specific psychological factors (depression and anxiety) that might underlie such a relationship.

Such a statement clearly communicates that the researcher is open-minded about what she may or may not find.

5. *Edit your work.* You can avoid the difficulties we have been discussing by carefully editing your words. *Editing* is sharpening a thought to a gemlike point and eliminating useless verbiage. Choose your words precisely, ideally selecting simple words, concrete nouns, and active, expressive verbs.

The sentences in the preceding paragraph began as a mishmash of foggy thought and jumbled verbiage. The original version of the paragraph contained 71 words. These were edited down to 41 words, yielding a reduction of about 40% and a great improvement in clarity and readability. Figure 2.1 shows the original version and how it was edited. The three lines under the *c* in *choose* means that the first letter should be capitalized. We present some of the common editing marks when we discuss editing in more detail in Chapter 5.

Notice the directness of the edited copy. We eliminated unnecessarily wordy phrases—"relating to the statement of the problem," "a process whereby the writer attempts to bring what is said straight to the point"—replacing the verbosity with seven words: "sharpening a

FIGURE 2.1

Editing to clarify your writing:
An example

You can avoid the difficulties
~~^ We have been discussing several common difficulties~~
~~relating to the statement of the problem. These can be~~
~~improved or remedied through a careful editing of your~~
~~words. Editing is a process whereby the writer attempts~~
~~to bring what is said straight to the point. Editing also~~
~~eliminates many meaningless expressions. We should~~
~~therefore, choose our words carefully. By editing the words~~
~~words, concrete nouns, and active, expressive verbs.~~
~~we have written our expression will take on new life.~~

by carefully sharpening a thought a gemlike

your precisely, ideally selecting simple

words, concrete nouns, and active, expressive verbs.

thought to a gemlike point." As we edited, we also pinned down what good word choice might involve.

Editing almost invariably improves your thinking and your prose. Many students think that any words that approximate a thought are adequate to convey it to others. This is not so. Approximation is never precision.

The following checklist can help you formulate a research problem that is clear, precise, and accurate.



CHECKLIST

Evaluating the Research Problem

- _____ 1. Write a clear statement of a problem for research.

- _____ 2. Review your written statement and ask yourself the following questions:

- Is the problem stated in a complete, grammatical sentence?
- Is it clear how the area of study will be limited or focused?
- Is it clear that you have an open mind about results that the research effort might yield?

- _____ 3. On the basis of your answers to the questions in Item 2, edit your written statement.

- _____ 4. Look at your edited statement and reflect on the following questions:

- Does the answer to this problem have the potential for providing important and useful answers and information?
- Will the result be more than a simple exercise in gathering information, answering a yes/no question, or making a simple comparison?
- Is the problem focused enough to be accomplished with a reasonable expenditure of time, money, and effort?

- _____ 5. Looking at the statement once more, consider this: Is the problem really what you want to investigate?

- _____ 6. Show other research students your work. Ask them to consider the questions listed in Items 2 and 4 and then to give you their comments. With your compiled feedback, edit and rewrite your problem statement once again:

Dividing the Research Problem into Subproblems

Most research problems are too large or complex to be solved without subdividing them. The strategy, therefore, is to divide and conquer. Almost every problem can be broken down into smaller units. From a research standpoint, these units are easier to address and resolve.

As noted in Chapter 1, the subparts of the main problem are called **subproblems**. By viewing the main problem through its subproblems, the researcher frequently gets a better idea of how to approach the entire research endeavor. So always think of a problem in terms of its component parts.

Subproblems versus Pseudo-Subproblems

The researcher must distinguish subproblems that are an integral part of the main problem from things that look like problems but are really nothing more than procedural issues. The latter, which we might call *pseudo-subproblems*, involve decisions the researcher must make before he or she can resolve the research problem and its subproblems. Consider the following as examples:

- What is the best way to choose a sample?
- How large should a representative sample of a population be?
- What instruments or methods should be used to gather the data?
- What statistical procedures should be used to analyze the data?

Deal with pseudo-subproblems forthrightly by making a reasonable decision about them and then get on with the solution of your research problem. To deal with pseudo-subproblems, you must decide whether (a) a little common sense and some creative thinking might help in solving your “problem” or (b) you simply lack the knowledge to address the difficulty. In the latter case, you have four options:

1. Turn to the index of this text to see whether your pseudo-subproblem regarding sample selection, instrumentation, statistical analysis, or some other issue is discussed.
2. Carefully peruse the “For Further Reading” lists at the end of each chapter in this book to see whether they contain any references that might help you, and consult general research methods books in your discipline.
3. Go to a library—preferably a college or university library—and search for books using such keywords as “Research” and either “Methods” or “Methodology.” Find these books in the library stacks and look in the indexes for possible discussions of your pseudo-subproblems. Also search your library’s online databases to find potentially relevant journal articles. If your library does not have certain periodicals, you can typically obtain any article you need through interlibrary loan.
4. Seek the suggestions and advice of more experienced researchers in your field. Recall a point made in Chapter 1: One of the most effective strategies for using the human mind is *collaborating with other minds*.

Characteristics of Subproblems

Following are four key characteristics of subproblems.

1. *Each subproblem should be a completely researchable unit.* A subproblem should constitute a logical subarea of the larger research undertaking. Each subproblem might be researched as a separate subproject within the larger research goal. The solutions to the subproblems, taken together, are then combined to resolve the main problem.

It is essential that each subproblem be stated clearly and succinctly. Often a subproblem is stated in the form of a question. A question tends to focus the researcher’s attention more directly on the research target of the subproblem than does a declarative statement. As we have seen, a questioning, open-minded attitude is the mark of a true researcher.

2. *Each subproblem must be clearly tied to the interpretation of the data.* Just as is true for the main problem, each subproblem should involve interpretation as well as collection of data. This fact may be expressed as a part of each subproblem statement, or it may be reflected in a separate but related subproblem.

3. *The subproblems must add up to the totality of the problem.* After you have stated the subproblems, check them against the statement of the main problem to see that (a) nothing in excess of the coverage of the main problem is included and (b) all significant areas of the main problem are covered by the subproblems.

4. *Subproblems should be small in number.* If the main problem is carefully stated and properly limited to a feasible research effort, the researcher will find that it usually contains two to six subproblems. Sometimes the inexperienced researcher will come up with as many as 10, 15, or 20 subproblems. When this happens, a careful review of the problem and its attendant subproblems is in order. If you find yourself in this situation, you should study the individual subproblems to see whether (a) some are actually procedural issues (pseudo-subproblems), (b) some might reasonably be combined into larger subproblems, or (c) the main problem is more complex than you originally believed. If the last of these is true, you may want to reconsider whether the solution to the overall research problem is realistically achievable given the time and resources you have.

Identifying Subproblems

Novice researchers often have difficulty identifying the subproblems within the main problem. You should begin with the problem itself. If the problem is correctly written, you will be able to detect subproblem areas that can be isolated for further study. The old axiom that the sum of the parts equals the whole applies here. All of the subproblems must add up to the total problem.

You can use either paper and pencil or brainstorming software to help you identify your subproblems. We briefly describe each of these strategies.

Taking a Paper-and-Pencil Approach

Using this approach, you write the problem on paper and then box off the subproblem areas. More specifically, you might follow these steps:

1. Copy the problem on a clean sheet of paper, leaving considerable space between the lines.
2. Read the problem critically to discover the areas that require in-depth treatment before the problem can be resolved.
3. Make sure every subproblem contains a word that indicates the necessity to interpret the data within that particular subproblem (e.g., *analyze, discover, compare*). Underline this word.
4. Arrange the entire problem, which will now have the subproblems boxed off, in a graphic that shows the research structure of the problem. You now have a structure of the whole research design.

This procedure for finding subproblems should work for any problem in any academic discipline. We use a problem in musicology to illustrate the technique. More specifically, we take the problem of the motets of Palestrina. As presented earlier in the chapter, this problem is as follows:

This study will analyze the motets of Giovanni Pierluigi da Palestrina (1525?–1594) written between 1575 and 1580 to discover their distinctive contrapuntal characteristics and will contrast them with the motets of his contemporary William Byrd (1542?–1623) written between 1592 and 1597. During the periods studied, each composer was between 50 and 55 years of age.

Let's first delete the factual matter, such as life-span dates and the fact that the two men were contemporaries. These facts merely help in giving a rationale for certain elements within the problem. Modified to reflect its essential parts, the motet problem becomes the following:

The purpose of this study will be *to analyze* the motets of Palestrina written between 1575 and 1580 to discover their distinctive contrapuntal characteristics, *to analyze* the same characteristics in the motets of William Byrd written between 1592 and 1597, and to determine what *a comparison of these two analyses* may reveal.

Notice that we have broken up the "will contrast them with" phrase in the original statement into two distinct tasks, *analyzing* Byrd's motets in the same manner that Palestrina's motets have been analyzed, and *comparing* the two analyses. The three italicized phrases in the revised problem statement reflect three subproblems, each of which involves interpretation of data that is necessary for resolving the main research problem.

Let's now arrange the problem so that we may see precisely what the design will be. Figure 2.2 is a graphic depiction of the problem. We have divided the problem into three subproblems. The first and second of these have the same general structural configuration: The analytical aspect of the subproblem is stated in the upper box and the purpose of the analysis is stated in the lower box. Addressing the third subproblem involves comparing the analyses conducted for the two preceding subproblems to determine what similarities and differences may exist. The last of the three subproblems—the comparison step—resolves the original main problem: characterizing Palestrina's motets.



Using Brainstorming Software

Some computer software facilitates the process of breaking problems into subproblems. Computer programs such as Inspiration, BrainStorm, and MindJet allow you to brainstorm research ideas and construct graphic networks of interrelated concepts, terms, and principles. For example, in Inspiration, you put the main problem, idea, or concept inside a box or oval in the middle of your computer screen. As you brainstorm other, related ideas, you put those on the screen as well, and you draw (and perhaps label) arrows to represent how various ideas are interconnected. You can break each concept or problem into subparts and, if helpful, break down each subpart even further. The process is fast and flexible, and you can save and print your final

FIGURE 2.2

A structural representation of the Palestrina-Byrd problem

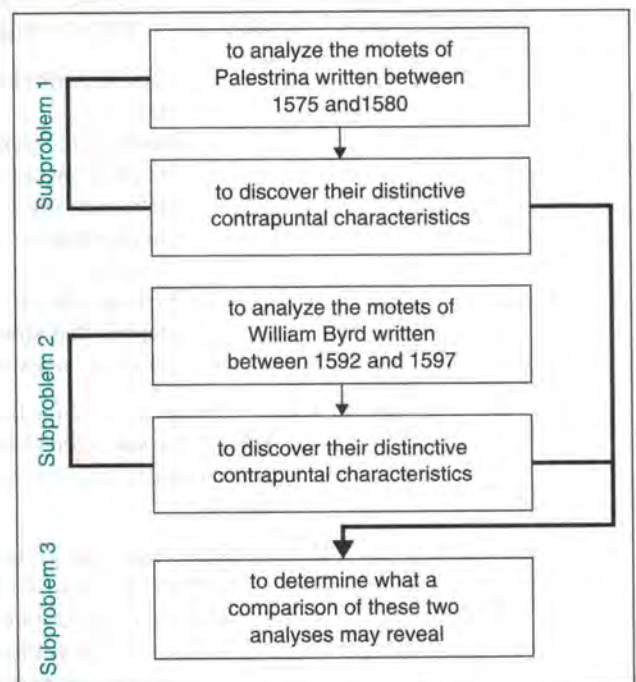


diagram (Figure 3.1, presented in Chapter 3, is an example). Some brainstorming software, such as Inspiration, also allows you to convert your diagram into an outline that lists major topics and various levels of subtopics.

Every Problem Needs Further Delineation

Up to this point, we have been discussing only the problem and its subparts. The statement of the problem establishes the goal for the research effort. The subproblems suggest ways of approaching that goal in a manageable, systematic way. But a goal alone is not enough. To comprehend fully the meaning of the problem, we need other information as well. Both the researcher and those reading the research proposal should ultimately have a clear understanding of every detail of the process.

In any research endeavor, the researcher should eliminate any possibility of misunderstanding by

- *Stating the hypotheses and/or research questions:* Describing the specific hypotheses being tested or questions being asked.
- *Delimiting the research:* Fully disclosing what the researcher intends to do and, conversely, does not intend to do.
- *Defining the terms:* Giving the meanings of all terms in the statements of the problem and subproblems that have any possibility of being misunderstood.
- *Stating the assumptions:* Presenting a clear statement of all assumptions on which the research will rest.

Taken as a whole, these elements comprise *the setting of the problem*. We look at each of them in more detail in the following sections. We also include a section titled "Importance of the Study," as this topic is frequently discussed in dissertations and other research reports.

Stating the Hypotheses and/or Research Questions

We previously discussed hypotheses in Chapter 1. There we pointed out that hypotheses are intelligent, tentative guesses about how the research problem might be resolved. *Research questions* are somewhat different in that, in and of themselves, they do not offer speculative answers related to the research problem. Hypotheses are essential to experimental research (see Chapter 9), whereas research questions are more common in many forms of qualitative research (see Chapter 6). Both hypotheses and research questions provide guidance for the kinds of data the researcher should collect; they also suggest how the researcher should analyze and interpret those data. It is not unusual for a researcher to form hypotheses *and* ask questions related to a research problem.

Research hypotheses and questions may originate in the subproblems. Often a one-to-one correspondence exists between the subproblems and their corresponding hypotheses or questions, in which case there are as many hypotheses or questions as there are subproblems.

Certainly the data from a research study can (and should) answer each research question, and they may support or not support each research hypothesis. But notice how we just said that the data may *support* or *not support* each research hypothesis; we intentionally did *not* say that the data would "prove" a hypothesis. Hypotheses are nothing more than *tentative propositions set forth to assist in guiding the investigation of a problem or to provide possible explanations for the observations made*.

A researcher who deliberately sets out to prove a hypothesis does not have the objective, impartial open-mindedness so important for good research. The researcher might bias the procedure by looking only for those data that would support the hypothesis (recall the discussion of *confirmation bias* in Figure 1.3 of Chapter 1). Difficult as it may be at times, we must let the chips fall where they may. Hypotheses have nothing to do with proof. Rather, their acceptance or rejection depends on what the data—and the data alone—ultimately reveal. If you discover that your data do not support your research hypothesis, do not let such an outcome disturb you. It merely means that your educated guess about the outcome of the investigation was incorrect.

Distinguishing Null Hypotheses from Research Hypotheses

Because we can never really prove a hypothesis, we often set out to cast doubt on—and therefore to *reject*—an opposite hypothesis. For example, imagine that a team of social workers believes that one type of after-school program for teenagers (Program A) is more effective in reducing high school dropout rates than is another program (Program B). The team's research hypothesis is:

Teenagers enrolled in Program A will graduate from high school at a higher rate than teenagers enrolled in Program B.

Because the social workers cannot actually prove their hypothesis, they instead try to discredit an opposite hypothesis:

There will be no difference in the high school graduation rates of teenagers enrolled in Program A and those enrolled in Program B.

If, in their research, the social workers find that there *is* a substantial difference in graduation rates between the two programs—and in particular, if the graduation rate is higher for youth in Program A—they can reject the “no difference” hypothesis and therefore have, by default, supported their research hypothesis.

When we hypothesize that there will be *no* differences between groups, *no* consistent relationships between variables, or, more generally, *no* patterns in the data, we are forming a **null hypothesis**. Null hypotheses are used primarily during statistical analyses; we support a research hypothesis by showing, statistically, that its opposite—the null hypothesis—probably is *not* true. Accordingly, we examine null hypotheses again in our discussion of statistics in Chapter 11.

Identifying the Variables under Investigation

Integral to a researcher's hypotheses and/or research questions are the variables that are the focus of investigation. We have occasionally used the term *variable* in earlier discussions in this chapter and in Chapter 1, but without defining it. We do so now: A **variable** is any quality or characteristic in a research investigation that has two or more possible values. For example, variables in studies of how well seeds germinate might include amounts of sun and water, kinds of soil and fertilizer, presence or absence of various parasites and microorganisms, genetic makeup of the seeds, speed of germination, and hardiness of the resulting plants. Variables in studies of how effectively children learn in classrooms might, for example, include instructional methods used; teachers' educational backgrounds, emotional warmth, and beliefs about classroom discipline; and children's existing abilities and personality characteristics, prior learning experiences, reading skills, study strategies, and achievement test scores.

Whenever a research project involves a possible cause-and-effect relationship, we are, of course, looking at the extent to which one variable (the hypothesized *cause*) influences another variable (the hypothesized *effect*). A variable that the researcher studies as a possible cause of something else—in many cases, this is one that the researcher directly manipulates—is called an **independent variable**. A variable that is potentially influenced by the independent variable—that “something else” just mentioned—is called a **dependent variable**, because it is influenced by and so to some extent *depends* on the independent variable. In research in the social sciences and education, the dependent variable is often some form of human behavior. In medical research, it might be people's physical health or well-being. In agricultural research, it might be quality or quantity of a particular crop. In general, a cause-and-effect relationship can be depicted like this:

Independent variable → Dependent variable

To illustrate the two kinds of variables, let's take an everyday situation. One hot summer morning you purchase two identical cartons of chocolate ice cream at the supermarket. When you arrive home, you put one carton in your refrigerator freezer but absentmindedly leave the other one on the kitchen counter. You then leave the house for a few hours. When you return home, you discover that the ice cream on the counter has turned into a soupy mess. The ice

cream in the freezer is still in the same condition it was when you purchased it. Two things vary in this situation. One, the temperature at which the ice cream is stored, is the independent variable. The other, the consistency of the ice cream, depends on the storage temperature and therefore is the dependent variable.

Now let's consider an example in medical research. Imagine that you want to compare the relative effectiveness of two different drugs that are used to treat high blood pressure. You take a sample of 60 men who have high blood pressure and randomly assign each man to one of two groups: The men in one group take one drug, and the men in the other group take the other drug. Later, you compare the blood pressure measurements for the men in the two groups. In this situation, you are manipulating the particular drug that each man takes; the drug, then, is the independent variable. Blood pressure is the dependent variable. It is presumably influenced by the drug taken and so its measured value depends to some extent on the drug.

A hypothesis or research question may occasionally specify other variables as well. For example, a **mediating variable** (also known as an *intervening variable*) might help explain why a certain independent variable has the effect that it does on a dependent variable. In particular, the independent variable influences the mediating variable, which in turn influences the dependent variable. Thus, the independent variable's influence on the dependent variable is an *indirect one*, as follows:

Independent variable → Mediating variable → Dependent variable

For example, consider the common finding that people who are confident in their ability to perform a particular new task do, on average, actually perform it better than less-confident people, even if the two groups of people had the same ability levels prior to performing the task. Looking at the situation from a simple independent-and-dependent-variables perspective, the situation would be depicted this way:

Confidence level → Performance quality
(independent variable) (dependent variable)

But *why* does this relationship exist? One likely mediating variable is that highly confident people exert more effort in performing the new task than do people with less confidence (e.g., Bandura, 1997; Schunk & Pajares, 2005). The mediating variable, then, is *amount of effort*, as follows:

Confidence level → Amount of effort (mediating variable) → Performance quality

Still another variable of potential interest is a **moderating variable**, a variable that, while not intervening between the independent and dependent variables, influences the nature and strength of their relationship. For example, consider the fact that, on average, children from very-low-income homes are more likely to have difficulties in adolescence and adulthood; for instance, compared to their financially more advantaged peers, they are less likely to complete high school and more likely to get in trouble with the law. Yet some very poor youngsters are resilient to their circumstances: They do quite well in life, sometimes going on to become physicians, lawyers, college professors, or other successful professionals. One factor that apparently increases the odds of resilience—in other words, it *reduces* the cause-and-effect relationship between childhood poverty and later problems—is a warm, supportive mother (Kim-Cohen, Moffitt, Caspi, & Taylor, 2004). Maternal warmth is a moderating variable: It affects the *nature of the relationship* between family income level and adult problems, like so:

Maternal warmth
(moderating variable)
↓
Childhood income level → Problems later in life
(independent variable) (dependent variable)

The distinction between mediating and moderating variables is an important but often confusing one; even some experienced researchers get them confused (Holmbeck, 1997). A helpful way to keep them straight is to remember that an independent variable may potentially influence a mediating variable but does *not*, in and of itself, influence a moderating

variable. For example, in the earlier *mediating variable* example, a high confidence level might increase the amount of effort exerted, but in the *moderating variable* example, we would certainly not suggest that having a low income increases (i.e., causes) a mother's warmth toward her children. Rather, moderating variables provide potential *contexts or conditions* that alter—that is, they *moderate*—an independent variable's effects. When researchers refer to *risk factors* or *protective factors* in their research reports, they are talking about moderating variables—variables that affect the likelihood that certain cause-and-effect relationships will come into play.

Not all research studies require explicit identification of independent and dependent variables and any mediating or moderating variables under consideration. Such variables are most often identified in quantitative research, especially in experimental studies (see Chapter 9) and certain kinds of descriptive studies (see Chapter 8).

Identifying independent and dependent variables is often quite helpful both in choosing (a) an appropriate research design and (b) an appropriate statistical analysis. However, an important caution is in order here. In particular, *identifying independent and dependent variables does not guarantee that the research data will support the existence of a cause-and-effect relationship*. We return to this point in the discussion of correlational research in Chapter 6.

CONCEPTUAL ANALYSIS EXERCISE Identifying Independent, Dependent, Mediating, and Moderating Variables

Following are eight proposed research problems. Each one of them implies one or more independent variables and one or more dependent variables. Some of them also imply one or more mediating or moderating variables. Identify the independent and dependent variables—and, if applicable, any mediating and/or moderating variables—in each problem. We warn you that some of these scenarios may challenge you, as the writer's hypotheses may lie well below the surface of the words. We encourage you, then, to try to put yourself in each researcher's mind and guess what the person is probably thinking about a possible cause-and-effect relationship in the phenomenon under investigation. The answers appear after the "For Further Reading" list at the end of the chapter.

1. In this study, I will examine the possible effects of regular physical exercise on the health and longevity of laboratory rats.
2. In this study, I will investigate the extent to which placing recycling bins in convenient locations in classroom buildings affects college students' recycling behaviors.
3. In this study, I will examine the relationship between amount of cell phone use while driving and the frequency of car accidents.
4. I propose to study the degree to which test anxiety may influence test performance by increasing the frequency of distracting thoughts.
5. This investigation will examine the extent to which a supportive student–teacher relationship reduces the negative emotional impact of peer bullying on a child's emotional well-being.
6. I will investigate the degree to which male and female adolescents choose gender-stereotypical careers in three different countries: Canada, Lebanon, and Japan.
7. This study will investigate the extent to which a particular tumor-suppressing gene reduces the risk of getting melanoma [a potentially deadly form of skin cancer] after a history of frequent exposure to sunlight.
8. In this study, I will investigate the possible relationship between body mass index and psychological stress, as well as two more specific psychological factors (depression and anxiety) that might underlie such a relationship. (You previously saw this problem statement in the guidelines for "Stating the Research Problem" earlier in the chapter.)

Delimiting the Research

The statement of the research problem describes precisely what the researcher intends to do. But it is also important to know precisely what the researcher does *not* intend to do. What the researcher is not going to do is stated in the *delimitations*. The limits of the problem should be as carefully bounded for a research effort as a parcel of land is for a real estate transfer.

Research problems typically emerge from larger contexts and larger problem areas. The researcher can easily be beguiled and drawn off course by addressing questions and obtaining data that lie beyond the precincts of the problem under investigation. For instance, in the Palestrina-Byrd problem, it is possible that, because the two men were contemporaries, Byrd may have met Palestrina or at least come in contact with some of his motets. Such contact may have been a determinative influence on Byrd's compositions. But given how the problem has been stated, the researcher does not need to be concerned with *influences* on the motets of the two composers. He or she should be primarily interested in the *characteristics* of the motets, including their musical style, musical individualism, and contrapuntal likenesses and differences. Study the contrapuntal characteristics—that is what a researcher of this problem will do. What the researcher will *not* do is become involved in any data extraneous to this goal—no matter how enticing or interesting such an exploratory safari might be.

Only a researcher who thinks carefully about the problem and its focal center can distinguish between what is relevant and what is not relevant to the problem. All irrelevancies to the problem must be firmly ruled out in the statement of delimitations. Figure 2.3 may make the matter of delimitations more understandable.

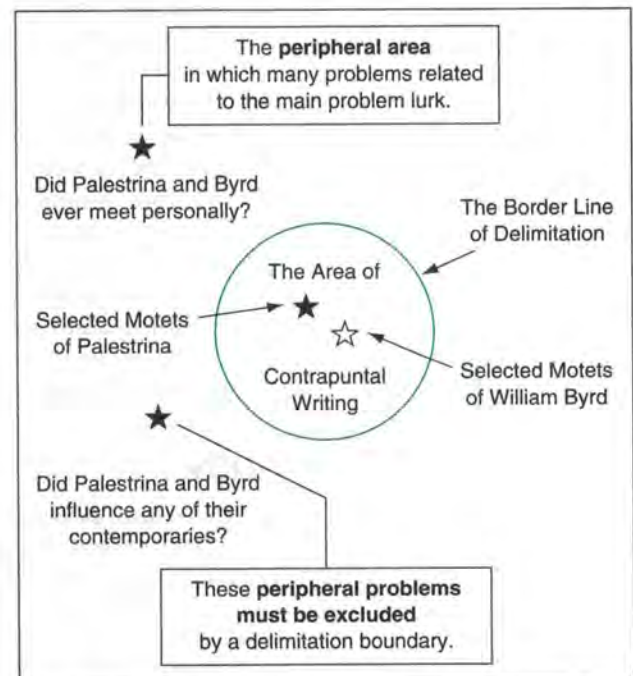
Defining Terms

What precisely do the terms in the problem and the subproblems mean? For example, if we say that the purpose of the research is to analyze the harmonic characteristics of motets, what are we talking about? What are *harmonic characteristics*? Without knowing explicitly what a term means, we cannot evaluate the research or determine whether the researcher has carried out what was proposed in the problem statement.

Sometimes novice researchers rely on dictionary definitions, which are rarely either adequate or helpful. Instead, each term should be defined *as it will be used in the researcher's project*. In defining a term, the researcher makes the term mean whatever he or she wishes it to mean.

FIGURE 2.3

Delimitation of a problem



within the context of the problem and its subproblems. Other individuals who read the researcher's research proposal or report must know how the researcher defines the term. Those individuals won't necessarily agree with such a definition, but as long as they know what the researcher means when using the term, they can understand the research and appraise it appropriately.

A formal definition contains three parts: (a) the *term* to be defined; (b) the *genera*, or the general class to which the concept being defined belongs; and (c) the *differentia*, the specific characteristics or traits that distinguish the concept being defined from all other members of the general classification. For example, *harmonic characteristics* (the term to be defined) might be defined as *the manner* (the genera) in which tonal values are combined *to produce individualized polyphonic patterns associated with the works of a particular composer* (the differentia: telling what particular "manner" the researcher means).

The researcher must be careful to avoid *circular definitions*, in which the terms to be defined are used in the definitions themselves. For instance, if we were to define *harmonic characteristics* in a circular manner, we might describe them as "those characteristics that derive from the harmonic patterns found in the works of a particular composer." Here the words *characteristics* and *harmonic* are used to define harmonic characteristics, giving others little or no assistance in understanding what the researcher means by the term.

Especially when talking about phenomena that have no cut-and-dried, easy-to-pinpoint manifestation in the physical world, it is often helpful to include an **operational definition**. That is, the researcher defines a characteristic or variable in terms of how it will be measured in the research study. For instance, a researcher might, for purposes of his or her study, define *intelligence* as a score on a certain intelligence test or define *popularity* as the number of peers who specifically identify an individual as being a desirable social partner. As another example, let's return to the first scenario in the earlier Conceptual Analysis Exercise: examining the possible effects of regular physical exercise on the health and longevity of laboratory rats. Longevity is easily defined and measured: It is simply the length of a rat's life span in days or some other unit of time. Somewhere in the research proposal, however, the researcher will need to be more specific about how he or she will define and measure physical exercise and health, thereby providing operational definitions for these terms. For example, physical exercise might involve putting a treadmill in some rats' cages but not in others. Health might be measured in any number of ways—for instance, through measurement of hypertension or analyses of blood or hair samples.

Stating the Assumptions

We briefly discussed assumptions in Chapter 1. Assumptions are so basic that, without them, the research problem itself could not exist. For example, suppose we are attempting to determine, by means of a pretest and a posttest, whether one method of classroom instruction is superior to another. A basic assumption in such a situation is that the pretest and posttest measure knowledge of the subject matter in question.¹ We must assume, too, that the teacher(s) in the study can teach effectively and that the students are capable of learning the subject matter. Without these assumptions, our research project would be meaningless.

In research, we try to leave nothing to chance in order to prevent any misunderstandings. All assumptions that have a material bearing on the problem should be openly and unreservedly set forth. If others know the assumptions a researcher makes, they are better prepared to evaluate the conclusions that result from such assumptions.

To discover your own assumptions, ask yourself: What am I taking for granted with respect to the problem? Ideally, your answer should bring your assumptions into clear view.

Importance of the Study

In dissertations or research reports, researchers frequently set forth their reasons for undertaking the study. Such a discussion may be especially important in a research proposal. Some studies seem to go far beyond any relationship to the practical world. Of such research efforts

¹Alternatively, we might make no such assumption; instead, we might set out to determine the *validity* of the tests as measures in this situation. We discuss the nature of validity of measurement in Chapter 4.

one inwardly, if not audibly, asks, “Of what *use* is it? What *practical value* does the study have?”

Contemplating the six manned moon explorations conducted between 1969 and 1972, many citizens asked, “What good is it? What’s the use of it all? How will spending all this money on space flights benefit anyone?” Perhaps those engaged in space research did not set forth clearly and succinctly enough the reasons the missions were undertaken. Only now are we beginning to appreciate the practical value of those early missions.

Ordering the Topics in a Research Proposal

To some degree, you may find a one-to-one correspondence between the discussions in this text and the sequence of topics that typically appear in a research proposal or research report. In any document, the first order of business is to present the problem and its setting. Generally, the document opens with a statement of the problem for research. This is followed by subproblems, hypotheses, and questions presented in a logical order.

Once the problem and its component parts have been articulated, the remaining items comprising the setting of the problem are presented, typically including a statement of delimitations, definitions of terms, and assumptions. A discussion of the importance of the study might have its own section or, alternatively, might be integrated into early paragraphs that introduce the research problem.

In a proposal or research report, such items often comprise the first chapter or section. The document then generally continues with a discussion of investigations that others have done, usually titled “Review of the Related Literature” or something of that nature. We discuss this topic in the next chapter.

PRACTICAL APPLICATION Writing the First Sections of a Proposal

In a checklist earlier in this chapter, you stated your main problem for research. In doing so, you took the first step in creating a research proposal. Now you can add the subproblems and identify the setting of the problem by doing the following exercise.

1. *State the subproblems.* On a blank sheet of paper, write the research problem statement you developed earlier. Allow considerable space between the lines. Now inspect your problem carefully and do the following:
 - a. Within the problem, box off those areas that must receive in-depth treatment if the problem is to be fully explored. Consecutively number the boxed-in areas.
 - b. Underline the words that indicate your intention to interpret the data (e.g., *analyze*, *compare*).
 - c. Below the problem, which has been thus treated, write the several subproblems of your study in complete sentences. Make sure each subproblem includes a word that reflects data interpretation.
2. *Write your hypotheses/questions.* Read again what we have said about hypotheses and research questions in this chapter. Then write your own hypothesis/question related to each of your subproblems.
3. *Write the delimitations.* Review an earlier section of the chapter, titled “Delimiting the Research.” Now write down the specific things that your own research project will *not* address.
4. *Write the definitions of terms.* Before writing your definitions, reread the section “Defining the Terms” earlier in the chapter. After writing your definitions, it may be helpful to box in the specific parts of each definition, labeling each box as “term,” “genera,” or “differentia.” (Delete these labels in the final draft of your proposal.)

5. *Write the assumptions.* Reread the section “Stating the Assumptions.” Now write a list of the specific assumptions you will be making as you design and carry out your research project—perhaps assumptions related to the people you will be studying, your measurement techniques, the relevance (or nonrelevance) of the context in which you will be conducting your study, and so on.
6. *Describe the importance of the study.* In a short paragraph or two, explain why your study is important. Eventually you may want to move this discussion to an earlier point in your proposal where you introduce your topic and provide an overall context for it (more on this point in Chapter 5). For now, however, keeping it in a separate section with its own label can help you remember that *talking* about your study’s importance is important in its own right.
7. *Type your proposal.* Ideally, use word processing software so that you will easily be able to make future edits (there will be many!). Set margins at least an inch wide, and double-space the entire document; double spacing makes proofreading easier and allows room for handwritten edits.

Now that you have written the first sections of a proposal, reflect on your proposed project using the following checklist.



CHECKLIST

Evaluating Your Proposed Research Project

1. Have you read enough literature relevant to your topic to know that your research project is worth your time and effort?
 - Will the project advance the frontiers of knowledge in an important way?

 - Have you asked an expert in your field to advise you on the value of your research effort?

2. Have you looked at your research problem from all angles to minimize unwanted surprises?
 - What is good about your potential project?

 - What are the potential pitfalls of attempting this research effort?

3. What research procedure will you follow?
 - Do you have a tentative plan to review the literature?

 - Do you have a tentative plan for data collection?

 - Do you have a tentative plan for data analysis?

- Do you have a tentative plan to interpret the data you collect?

4. What research tools are available for you to use? Make a list and check their availability. Determine how you will use them.

5. Ask two or three peers to read your proposal. Do they understand what you are proposing to do? What questions do they have? What concerns do they express?

- I have discussed this plan with _____, and _____.
- They have the following questions and concerns:

PRACTICAL APPLICATION Reappraising a Proposed Research Problem

In this chapter we have given you numerous suggestions for identifying an appropriate problem or question for your research. Because the problem is the center and driving force of any research project, we have devoted considerable space to its discussion. We cannot overemphasize that if the problem is not correctly selected and stated, you may put considerable time, energy, and resources into an endeavor that is much less than what it could be.

GUIDELINES Fine-Tuning Your Research Problem

Earlier in the chapter we presented guidelines for identifying and stating an appropriate research problem. Here we offer a few general suggestions for fine-tuning the problem you have identified:

1. *Conduct a thorough literature review.* You have presumably already looked at some of the literature related to your research problem. A next critical step is to make sure you know enough about your topic that you can ask important questions and then make solid decisions about how you might answer them through your research endeavor. You may find that you need to revise your research plan significantly once you have delved deep into the literature related to your topic. We address strategies for conducting an in-depth literature review in the next chapter.

2. *Try to see the problem from all sides.* What is good about this potential project? What is not? Try to take a critical view of what you are proposing to do. Such a perspective will help minimize unwanted surprises.

3. *Think through the process.* Once you have brought your research problem into clear focus, imagine walking through the whole research procedure, from literature review through data collection, data analysis, and interpretation. You can gain valuable insights as you mentally walk through the project. Pay close attention to specific bottlenecks and pitfalls that might cause problems later on.

4. *Use all available tools and resources at your disposal.* Remember that research is always a learning experience. Allow time for learning about new tools or for learning how to use old tools in new ways.

5. *Discuss your research problem with others.* Frequently, beginning researchers need to clarify their problem statement. One good way to do this is to present it to other people in as clear a fashion as possible. If they do not understand, further explanation and clarity are needed. One can learn a great deal from trying to explain something to someone else.

6. *Show the first draft of your written proposal to others for their examination and critique.* Do not hide it because you are afraid someone else may not like your idea or may want to steal it from you. Rarely will either of these events happen.

Continually ask for feedback from others. Ask other people questions about your research problem, and ask them to ask *you* questions about it. Do not be overly discouraged by a few individuals who may get some sense of satisfaction from impeding the progress of others. Many great discoveries have been made by people who were repeatedly told that they could not do what they set out to do.

7. *Remember that your project will take time—lots of time.* All too often, we have had students tell us that they anticipate completing a major research project, such as a thesis or dissertation, in a semester or less. In the vast majority of cases, such a belief is unrealistic. Consider all the steps involved in research: formulating a research problem, conducting the necessary literature search, collecting and interpreting the data, describing what you have done in writing, and improving on your research report through multiple drafts. If you think you can accomplish all of these things within 2 or 3 months, you are almost certainly setting yourself up for failure and disappointment. We would much rather you think of any research project—and especially your first project—as something that is a valuable learning experience in its own right. As such, it is worth however much of your time and effort it takes to do the job well.

8. *Remember that the first drafts of whatever you write will almost certainly not be your last ones.* Good researchers continually revise their thinking and, as a result, their writing. Furthermore, as mentioned in Chapter 1, writing about one's project often helps to clarify and enhance one's thinking. So get used to writing . . . and rewriting . . . and rewriting once again.

Nevertheless, by putting your problem statement on paper early in your research project, you have begun to focus your research efforts.

In this chapter we have given you numerous suggestions for identifying an appropriate problem or question for your research. Because the problem is the center and driving force of any research project, we have devoted considerable space to its discussion. We cannot overemphasize that if the problem is not correctly selected and stated, you may put considerable time, energy, and resources into an endeavor that is much less than what it could be.

For Further Reading

- Cooper, H. (2006). Research questions and research designs. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 849–877). Mahwah, NJ: Erlbaum.
- Dominick, J. R., & Wimmer, R. D. (2005). *Mass media research: An introduction* (8th ed.). Belmont, CA: Wadsworth. [See Chapter 2.]
- Gay, L. R., Mills, G. E., & Airasian, P. (2009). *Educational research: Competencies for analysis and application* (9th ed.). Upper Saddle River, NJ: Merrill/Pearson Education. [See Chapter 2.]
- Holmbeck, G. N. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology*, 65, 599–610.
- McBurney, D. H. (1995). The problem method of teaching research methods. *Teaching of Psychology*, 22(1), 36–38.
- McMillan, J. H., & Schumacher, S. (2006). *Research in education: Evidence-based inquiry* (6th ed.). Upper Saddle River, NJ: Prentice Hall. [See Chapter 3.]
- Medawar, P. B. (1979). *Advice to a young scientist*. New York: Harper & Row.
- Neuman, W. L. (2011). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Upper Saddle River, NJ: Pearson.
- Schram, T. H. (2006). *Conceptualizing and proposing qualitative research* (2nd ed.). Upper Saddle River, NJ: Merrill/Prentice Hall. [See Chapter 5.]

ANSWERS TO THE CONCEPTUAL ANALYSIS EXERCISE “Identifying Independent, Dependent, Mediating, and Moderating Variables” on page 42:

1. The words “effects of . . . on” tell us the direction of a hypothesized cause-and-effect relationship. Amount of physical exercise is the independent variable. Health and longevity are two dependent variables.
2. Placement of recycling bins is the independent variable; for example, the researcher might vary the number of recycling bins available and/or their proximity to classrooms and high-student-traffic areas. Recycling behavior is the dependent variable; for example, the researcher might count the number of recyclable objects (aluminum cans, sheets of paper, etc.) found in recycling bins each day or week.
3. The problem statement uses the term *relationship* without necessarily implying that this is a cause-and-effect relationship; for instance, the statement does not include a term such as *influence* or *affect*. However, we can reasonably guess that the researcher is hypothesizing that cell phone usage *increases the risk of* an accident, in which case amount of cell phone use is the independent variable and accident rate is the dependent variable. (To some degree, car accidents lead to cell phone use as well—someone in an accident is likely to call a family member or 911—but that cause-and-effect relationship hardly seems worthy of a research study.)
4. Don’t let the sequence of variables mentioned in the problem statement lead you astray here. Text anxiety is the independent variable; test performance is the dependent variable. The third variable mentioned—distracting thoughts—is hypothesized to be the mediating variable: Level of anxiety (independent variable) affects the degree to which one has distracting thoughts (mediating variable), which in turn affects test performance (dependent variable).
5. The word *impact* implies a possible causal connection between bullying (independent variable) and emotional well-being (dependent variable). The nature of a student’s relationship with his or her teacher can influence the impact of bullying; thus the student–teacher relationship is a moderating variable.
6. The problem statement includes no words to suggest the direction of a relationship. Certainly, however, career choices *cannot* affect one’s gender, so any possible causal relationship must go in the other direction: from gender (independent variable) to career choice (dependent variable). The comparative aspect of the problem statement suggests that the researcher suspects that gender might have more of an influence on career choice in some countries (presumably those that adhere to traditional ideas about occupations appropriate for men and women) than in others. Country of residence, then, would be a moderating variable affecting the strength of the gender–career choice relationship.
7. The cause-and-effect relationship between frequent exposure to sunlight (independent variable) and melanoma (dependent variable) is well established in the medical literature. The presence or absence of a particular gene is hypothesized to be a moderating variable: The chances of sunlight leading to melanoma may be reduced—that is, the cause-and-effect relationship may be considerably weaker or possibly nonexistent—if a person has the tumor-suppressing gene.
8. Once again the problem statement talks only about a *relationship*, without using verbs such as *cause*, *affect*, or *influence* to imply causation. However, the mention of two psychological factors that *underlie* the relationship suggests that the researcher is assuming that either body mass index affects psychological stress or vice versa. Although the problem statement does not clarify which of these two variables is the independent variable and which is the dependent variable, two other variables—levels of depression and anxiety—are apparently hypothesized to be *mediating variables*. Perhaps a higher body mass index (independent variable) increases depression and anxiety (mediating variables) that, in turn, increase psychological stress (dependent variable). Or perhaps, instead, greater psychological stress (independent variable) increases depression and

anxiety (mediating variables) that, in turn, lead to more food consumption and/or less physical exercise (two more, unstated and presumably unmeasured mediating variables), which in turn increase body mass index (dependent variable).

Now go to MyEducationLab at www.myeducationlab.com to take assessment to evaluate your mastery of chapter concepts. Assignments and Activities exercises and Building Skills in Reading Research exercises are also available to help you master concepts and skills. Feedback for many of these exercises is provided so that you can see why your answers are correct or incorrect.

3

Review of the Related Literature

Those who have conducted research before you belong to a community of scholars, each of whom has journeyed into the unknown to bring back an insight, a truth, a point of light. What they have recorded of their journeys and findings will make it easier for you to explore the unknown: to help you also discover an insight, a truth, or a point of light.

As noted in Chapter 2, reading the literature related to your topic of interest can help you formulate a specific research problem. There are several other ways in which related literature can help you as well. In this chapter we discuss the importance of the literature review and give you suggestions for reviewing the related literature thoroughly and efficiently.

Understanding the Role of the Literature Review

Research proposals and research reports typically have a section—in the case of a thesis or dissertation, often an entire chapter—that describes theoretical perspectives and previous research findings related to the problem at hand. Its function is to review—to “look again” at (*re + view*)—what others have done in areas that are similar, though not necessarily identical to, one’s own topic of investigation.

As a researcher, you should ultimately know the literature related to your topic *very, very well*. An extensive literature review has numerous benefits:

1. It can help you ascertain whether other researchers have already addressed and answered the research problem you are proposing (see Chapter 2) or at least some of its subproblems.
2. It can offer new ideas, perspectives, and approaches that may not have occurred to you.
3. It can inform you about other individuals who conduct work in this area—individuals whom you may wish to contact for advice or feedback.
4. It can show you how others have handled methodological and design issues in studies similar to your own.
5. It can reveal sources of data you may not have known existed.
6. It can introduce you to measurement tools that other researchers have developed and used effectively.
7. It can reveal methods of dealing with difficulties similar to those you are facing in tackling your research problem.
8. It can help you interpret and make sense of your findings and, ultimately, help you tie your results to the work of those who have preceded you.
9. It can bolster your confidence that your topic is one worth studying, because you will find that others have invested considerable time, effort, and resources in studying it.

Simply put, the more you know about investigations and perspectives related to your topic, the more effectively you can address your own research problem.

In most instances, researchers begin their review of the literature quite early in the game, and they draw on existing theories and prior research studies to help them pin down their research problem and accompanying hypotheses and questions. Only on rare occasions—when they intentionally choose *not* to let existing perspectives and research findings bias their data

collection and interpretation—do researchers postpone a thorough literature review until relatively late in the research process (e.g., see the discussion of grounded theory research in Chapter 6).

Strategies for Locating Related Literature

You might find literature related to your topic in a number of places—for instance, in books, journals, newspapers, government publications, conference presentations, and Internet websites. Obviously you cannot simply wander aimlessly through the library stacks or Internet with the hope that you will eventually stumble on items that may help you; you must focus your search from the very beginning.

A good way to start is to identify one or more **keywords**—words or short phrases summarizing your research topic—that can point you toward potentially useful resources. A prime source of such keywords is your statement of your research problem. For example, imagine that you want to investigate possible reasons why some children bully other children at school. Obvious keywords for this topic are *peer relationships*, *bullying*, and *aggression*. These are very general concepts, of course, but they should get you started. They will lead you to thousands of potential resources, however, and so you will soon want to identify more specific keywords. As you begin to look at books, journal articles, subject matter indexes, and other resources related to your topic and initial set of keywords, you should come across words and phrases that more closely capture what you want to study—for the bullying problem, these might include such terms as *social goals*, *social cognition*, and *cyberbullying*—and may also help you focus your research problem a bit more.

Armed with your initial keywords—which you will undoubtedly continue to revise—you can proceed in several directions. In the following sections, we describe seven good starting points: the library catalog, the library's index and abstract collections, online databases, reference librarians, the Internet, government publications, and other researchers' citations and reference lists.

Using the Library Catalog

The library catalog has come a long way from the tool it was in the mid-20th century. If you were a student in, say, 1960, after entering the library you would go straight to the card catalog—a set of free-standing dressers-of-sorts with many small drawers containing 3-by-5 index cards. The catalog would have three cards for every book in the library—one card each for a book's title, author, and general topic. You would rifle through the cards in search of books relevant to your topic and then write down the call numbers of books you wanted to track down in the library's numerous shelves of books (i.e., the “stacks”). If you were conducting an extensive literature review, the process might involve going through drawer after drawer in the card catalog, writing down a lengthy list of books and call numbers, and then heading to the stacks to determine whether or not each book you wanted was currently available. The whole process could be incredibly tedious and time-consuming.

In today's college library, a researcher's plan of attack is entirely different. Although you may occasionally find a small public library that still uses a physical card catalog, college and university libraries rely almost exclusively on electronic catalogs of their collections. In place of those rows upon rows of index-card drawers are computer terminals at which users can quickly locate a library's holdings related to particular authors, titles, keywords, or call numbers. The database will tell you on what floor of the library—and, if relevant, in what building or on what branch campus—a particular book can be found. (Note that some widely used books are kept in the library's reserved books section rather than in the stacks; you must read these books in the library itself, as they cannot be checked out.) The database will also tell you the status of a book—whether it's currently available or, if not, when it is due to be returned. If you have any questions about how to use the library catalog and its many features, don't be afraid of “looking stupid”—ask a librarian to show you the basics.

A good college or university library will almost certainly have a number of books relevant to your research topic. Some books will be written entirely by one or two individuals. Others may be edited collections of chapters written by a variety of experts in the field. And don't overlook general textbooks in your discipline. A good textbook can give you a broad overview of a topic, including important concepts, theoretical perspectives, a sampling of relevant research, and critical references.

The library's collection of academic journals, popular magazines, and newspapers—collectively known as **periodicals**—is another indispensable resource. The library catalog will tell you which periodicals the library owns, where each one is located, and the one or more forms (paper, electronic, microform) in which particular volumes and issues can be found. For instance, if the library has a periodical in paper form, you will find most volumes in the library stacks—usually in a section of the library devoted specifically to periodicals—but you are apt to find recently acquired, unbound issues (say, from the past year or two) on a separate set of shelves near the main desk for the periodicals section. Some university libraries organize and shelve their paper periodicals by call number; this approach enables a researcher to find periodicals about any single topic close together, but the researcher must, of course, know the call number(s) relevant to his or her discipline and topic. Other university libraries organize and shelve paper periodicals alphabetically by title; this approach enables a researcher to find any particular periodical without having to consult the library catalog, but he or she will most likely go to many different shelves to retrieve all articles relevant to a particular literature review.

University libraries typically also have access to many periodicals in electronic form, which you can retrieve from a computer terminal (more about accessing electronic copies in the upcoming section on online databases). Finally, your library may have some periodicals (especially older ones) in *microform*. The microforms area of a library is easy to spot, as it will have numerous file cabinets containing *microfilm*, *microfiche*, and the like, along with several large devices for viewing them. The devices may seem intimidating to a novice researcher, but they are quite easy to use once you have had a little practice. Don't be afraid to ask someone behind the periodicals desk to demonstrate how to use them.

As a rule of thumb, use books and periodicals with recent copyright dates. The more recently a book or article has been written, the more likely it is to give you a sense of current perspectives in your field and alert you to recent research findings that may be pertinent to your research problem. You should ignore this rule, of course, if you are specifically interested in how perspectives about your topic have changed over the years.

If you have access to the Internet from your home computer, then you already have access to countless online library catalogs around the world. You can explore how online catalogs work by visiting some of them on the Internet. Following are three examples:



Brown University

library.brown.edu

University of New Hampshire

library.unh.edu

University of Northern Colorado

library.unco.edu

Typically the Internet home page for your own institution will have a link to the library and its catalog.

A few words about call numbers. The *call numbers* referred to earlier are the unique identification codes that books and similar items are given. A book's call number provides an "address" that enables you to find it in the library stacks. Books are coded and arranged on the library shelves in accordance with one of two principal classification systems, which divide areas of human knowledge in somewhat different ways:

- *The Dewey decimal (DD) classification system.* Books are cataloged and shelved according to ten basic areas of knowledge and subsequent subareas, each divided decimally. The

Dewey decimal system is the principal classification system in most public libraries and many other libraries and is probably the most generally accepted system throughout the world.

- *The Library of Congress (LC) classification system.* Books are assigned to particular areas of knowledge that are given special alphabetical categories. This system is widely used in college and university libraries.

Table 3.1 provides a rough overview of how the two systems generally classify many traditional academic subject areas. For each subject area listed in the table, the entries in the DD column to its left and the LC column to its right provide either starting points or general ranges for the Dewey decimal and Library of Congress designations, respectively.

Please be aware, however, that neither the Dewey decimal system nor the Library of Congress system is as simple and cut-and-dried as Table 3.1 might suggest, in part because virtually any academic discipline includes many topics and draws from many research areas and—often—from other disciplines. Furthermore, we authors have found that books in our own areas of expertise are not always classified exactly as we ourselves might have classified them.

You can find descriptions of more specific categories of the Dewey decimal and Library of Congress systems on many Internet websites. A simple search on Google or Yahoo! using such phrases as “dewey decimal system chart” and “library of congress categories” can help you find these websites.

Browsing the library’s holdings. Although keywords and knowledge of specific book titles and authors can get you off to a good start in locating helpful volumes in your library, they will give you *only* a start, because you probably won’t be able to think of every potentially useful keyword, and you certainly won’t be aware of every book and author relevant to your topic. We therefore suggest that you also browse the library, either physically by walking among the stacks or electronically by “browsing” the entries in the library’s computer catalog. In many cases, when

TABLE 3.1

A general conversion chart: Dewey decimal classification system (DD) versus the Library of Congress classification system (LC) for various subject areas

DD	Subject	LC	DD	Subject	LC
630	Agriculture	S	070	Journalism	PN
301	Anthropology	GN	400	Language	P
930	Archaeology	CC	340	Law	K
700	Art	N	020	Library and Information Sciences	Z
520	Astronomy	QB	800	Literature	P
920	Biography	CT	510	Mathematics	QA
570	Biology	QH	610	Medicine and Public Health	QS–QZ, W
580	Botany	QK	355	Military Science	U
650	Business	HF	780	Music	M
540	Chemistry	QD	100	Philosophy	B
004–006	Computer Science	QA	530	Physics	QC
550	Earth Sciences	QE	320	Political Science	J
330, 380	Economics and Commerce	HB–HJ	150	Psychology	BF
370	Education	L	200	Religion	B
620	Engineering	T	500	Science (General)	Q
910	Geography	G	301	Sociology	HM
350	Government	JF, JK, JS	790	Sports and Recreation	GV
930–995	History	D, E, F	600	Technology	T
640	Hospitality	TX	590	Zoology	QL

one goes to a library shelf to get a particular book or journal, the information most useful is found not in the material that was originally targeted, but rather in one or more volumes nearby.

Remember, too, that most academic disciplines are becoming increasingly interdisciplinary in both their problems and their methodologies (e.g., Eisenhart & DeHaan, 2005; Miksa, 1987). For example, marketing researchers often draw on sociologists' and geographers' concepts and data collection techniques to identify the needs and shopping patterns of different populations, and psychologists can learn a great deal about human thought processes by using the positron emission tomography (PET) and magnetic resonance imaging (MRI) technologies of neurophysiologists. Thus, you are apt to find helpful sources under more than one range of call numbers. Good researchers are flexible and creative in their searches for relevant literature.

Using Indexes and Abstracts

Although the library catalog will tell you which periodicals your library owns and in what form it owns them, it won't tell you the specific articles that each volume of a journal contains. Instead, you may need to consult one or more indexes or abstract compilations in paper or electronic form. An **index** lists articles, research reports, and other documents in certain specified areas. An **abstract** is a summary of an article or research study; it also gives the source of the original study, should a reader wish to refer to it. The examples of indexes and abstract compilations presented in Table 3.2 should give you an idea of the wide variety of topics these resources address.

The paper forms of such resources are located in the library's reference section. A reference librarian can show you how to use them effectively and tell you which ones are available in electronic form as well as—or perhaps only instead of—paper form. In fact, their electronic versions, in the form of online databases, can greatly enhance the efficiency and effectiveness of your library searches.

Using Online Databases



Virtually all college and university libraries provide access to a wide variety of **online databases**—not only indexes and abstracts but also encyclopedias, dictionaries, and other general reference tools. These databases enable you to locate sources of information available either at the campus library or at other libraries and institutions around the world. Table 3.3 lists examples.

Online databases allow searches of thousands of journals and such other sources as books, chapters in edited books, dissertations, government documents, technical reports, and

TABLE 3.2

Examples of indexes and abstracts

<i>Indexes</i>	<i>Abstracts</i>
<i>Applied Science and Technology Index</i>	<i>Biological Abstracts</i>
<i>Biography Index</i>	<i>Dissertation Abstracts International</i>
<i>Book Review Index</i>	<i>International Political Science Abstracts</i>
<i>Business Periodicals Index</i>	<i>Psychological Abstracts</i>
<i>Essay and General Literature Index</i>	<i>Sage Race Relations Abstracts</i>
<i>Humanities Index</i>	<i>Sage Urban Studies Abstracts</i>
<i>Index to Legal Periodicals</i>	<i>Women Studies Abstracts</i>
<i>The Music Index</i>	
<i>New York Times Index</i>	
<i>Poole's Index to Periodical Literature</i>	
<i>Readers' Guide to Periodical Literature</i>	
<i>Social Sciences Index</i>	
<i>United Nations Documents Index</i>	
<i>Wall Street Journal Index</i>	

TABLE 3.3

Examples of online
databases

<i>Database</i>	<i>Subject Area(s) Covered</i>
Academic Search Premier	Education, humanities, multicultural issues, sciences, social sciences
America: History and Life	History of the United States and Canada
Applied Science Full Text	Applied sciences and technology (e.g., computing, engineering, resource management, telecommunications, transportation)
Art Full Text	Broad range of art topics (e.g., advertising, architecture, art history, folk art, graphic arts, video)
Biological Abstracts	Biology, medicine
Business Source Premier	Business, economics
Dissertation Abstracts International	All disciplines (dissertations)
EconLit	Economics
ERIC (Educational Resources Information Center)	Education and education-related topics
Historical Abstracts	World history (excluding the United States and Canada; for these, use America: History and Life)
IngentaConnect	All disciplines
JSTOR	Business, fine arts, humanities, sciences, social sciences
Linguistics and Language Behavior Abstracts (LLBA)	Language
MathSciNet	Mathematics (pure and applied), statistics
Medline	Dentistry, health care, medicine, veterinary medicine
National Criminal Justice Reference Service Abstracts	Courts, crime, justice, law enforcement, victimization
PAIS (Public Affairs Information Service) International	Public and social policy, social sciences
ProQuest Dissertations and Theses: Full Text	All disciplines
PsycINFO	Psychology and psychological aspects of other disciplines (e.g., physiology, sociology, anthropology, education, medicine, business)
Sociological Abstracts	Sociology and related topics in the social and behavioral sciences
SPORT Discus	Physical education, physical fitness, recreation, coaching, sports medicine
Web of Science	Humanities, sciences, social sciences
WorldCat	All disciplines

newspapers. A typical database allows you to limit your search in a variety of ways—perhaps by keywords, title, author, year, source (e.g., journal title), language, or any combination of these. Many databases focus on particular disciplines and subject areas. As an example, let's consider PsycINFO, a database that includes information not only about sources in psychology but also about psychology-related sources in such disciplines as physiology, sociology, anthropology, education, medicine, and business. As this edition of *Practical Research* goes to press, PsycINFO works as follows:

1. When you enter the database, you type one to three words or phrases in boxes at the top of the screen. In pull-down menus to the right of the boxes, you can indicate whether each word or phrase you have typed is a keyword, author, title, word or phrase in the abstract, or some other entity.
2. In pull-down menus to the left of the second and third boxes at the top of the screen, you can tell the computer to

- a. Identify only those items that include *all* of the words/phrases you have entered (for this, you select the “and” option).
 - b. Identify items that include *any* of the words/phrases you have entered (for this, you select the “or” option).
 - c. Exclude items that have one of the words/phrases you have entered (for this, you select the “not” option).
3. Options in the lower portion of the computer screen allow you to limit your search results still further, perhaps by specifying a particular year, journal title, population, age-group, or language.
 4. Once you have limited your search to some degree (at a minimum by completing Step 1), you click on “Search” near the top of your computer screen.
 5. The next screen will either (a) give you one or more references or (b) tell you that it has come up empty-handed (“Your query did not find any matching records”). If references appear, you can click on their titles to view abstracts and, in some cases, see and download the entire articles or other texts. If your search has been unsuccessful, you probably need to eliminate one or more of the limitations you imposed on your original search—you should also check for spelling errors in what you have typed—and click on “Search” once again.
 6. Each time you identify a potentially useful source, you can “mark this record,” and at the end of your search of PsycINFO, you can “save, print, [or] e-mail” the complete list of sources you have identified.

As is true for some of the articles in PsycINFO, some databases provide entire documents. For example, the *New York Times Article Archive* allows you to search—and then also to read—news articles, editorials, letters to the editor, birth announcements, obituaries, advertisements, and virtually any other entry in any issue dating back to its first issue in 1851. Another good general resource is JSTOR (pronounced “jay-stor”), which contains electronic copies of articles from many journals in the sciences, social sciences, humanities, business, law, and education.

One easy way to access a university library’s online databases is through computer terminals located throughout the library building. Often a library’s Internet home page will provide a link to its online databases, and users may be able to access them on their home computers as well as at the library. However, because a library pays large annual fees for its online databases, it typically restricts off-site use of them to current students and employees. Hence, students who want to use a database at home may need to enter a preassigned user name and password before gaining access to it. A reference librarian at your own library can give you the details.

Researchers not currently connected to a university have other possible ways to access online databases. Many professional associations give current members access to electronic copies of articles published in the associations’ journals. Some online databases are available without charge on the Internet. An example is Google Scholar (scholar.google.com), through which you can search the general topics and contents of books, journal articles, and other scholarly works in a wide range of disciplines. Some of the websites it identifies provide complete articles and other documents you can download and print (e.g., look for sites labeled “pdf”), whereas others provide abstracts and links to companies that charge a service fee for the complete documents. Another, more specialized database—one especially helpful for researchers interested in medicine and related topics—is that of the National Library of Medicine (www.nlm.nih.gov). Also, check out Google Books (books.google.com), which provides excerpts from—and in some cases the entire texts of—out-of-print books.

One especially helpful database during a literature search is Web of Science, which can tell you what publications cite *other* publications. For example, imagine that you are particularly intrigued by a 1999 article in the journal *Nature Neuroscience* indicating that the human brain is not fully mature until its owner reaches adulthood in the 20-something age range (Sowell, Thompson, Holmes, Jernigan, & Toga, 1999). Given the rapid-pace advances in neuroscience in recent years, this article is a fairly “old” one, and so you want to find more up-to-date articles on the same topic. In the Web of Science database, the “Cited Reference Search” option allows you to search the reference lists of all other articles in its database and find more recently published

articles that cite the article by Sowell and colleagues. If you were to use the Web of Science for this purpose (as we did), you would find that the Sowell and colleagues' article has been cited by hundreds of other researchers and so obviously has been an influential one in neuroscience.

Another invaluable database is **WorldCat**, which combines the library catalogs of thousands of academic libraries, large public libraries, and other specialized collections throughout the world. Through this database, you can identify libraries that have particular books, periodicals, visual materials, sound recordings, and other items that might be rare and hard to come by.

Our list of databases and their features is hardly exhaustive. Databases become more sophisticated with each passing year. We urge you to consult with a reference librarian about databases that may be especially suitable for your research purposes.

Consulting with Reference Librarians

When you visit the reference section of your library—and we urge you to do this very early in your literature search—you will almost certainly see one or more librarians sitting at the reference desk. These individuals are there for one reason only: to help you and others find needed information. They can show you reference materials you never dreamed existed. They can also demonstrate how to use the computer catalog, hard-bound reference volumes, online databases, or any of the library's other resources.

Some new researchers are reluctant to approach a reference librarian for fear of looking foolish or stupid. Yet the reality is that library resources are changing so quickly that most of us cannot possibly keep up with them all. Whatever you do, do *not* be afraid to ask librarians for assistance. Even as seasoned researchers, we authors sometimes seek the advice of these individuals; by doing so, we can often save ourselves a great deal of time and aggravation.

The best way to master the library as a research tool is to use it! Go in, explore, take stock of its resources; experiment with the various search options in its computer terminals; browse in the reference room; go into the stacks and browse some more. You may be surprised at what a magnificent research tool the library really is.

Surfing the Internet



We have already mentioned the Internet as a source of such free-access online databases as Google Scholar and Google Books. With each passing year, the Internet becomes an increasingly valuable source of information to researchers. As most of our readers already know, an Internet search typically begins with a **search engine** at a website such as Google or Yahoo! These websites typically have a box in which you can type one or more keywords to start your search. Following are some general strategies to keep in mind when using search engines:

1. Use at least two keywords to limit your search. (For example, to locate research about children with autism, you might type the words *children* and *autism*.)
2. Type a plus sign (+) before any keyword you definitely want used in your search. (For example, to limit your search only to children who have autism, you should type "+children" and "+autism". Otherwise, you may get a listing of all resources involving children *or* autism, which would undoubtedly be a long list indeed.)
3. If you want to look for a phrase rather than a single word, put quotation marks around the phrase. (For example, if you are looking for the home page of the Autism Society, you should type "Autism Society" within quotation marks. This way, your search will be restricted to items specifically about that particular organization.)

Surfing the Internet will lead you to a wide variety of locations. For instance, it may lead you to government office websites that can provide helpful information, including the U.S. Census Bureau (www.census.gov), Department of Education (www.ed.gov), Department of Labor (www.dol.gov), National Aeronautics and Space Administration (www.nasa.gov), National Network of Libraries of Medicine (www.nlm.gov), and U.S. Geological Survey (www.usgs.gov). Most professional associations have websites as well, and these sites often provide a wealth of information about their areas of expertise.

One site to which an Internet search will often lead you is Wikipedia (www.wikipedia.org), an online encyclopedia that virtually anyone can add to and edit. Wikipedia contains millions of entries on a diverse range of topics, with people adding new ones every day. In our experience, Wikipedia provides good general overviews of many topics and can help a novice researcher identify key concepts and issues related to a topic. Keep in mind, however, that *its contents are not juried*—that is, there is no oversight of any entry's accuracy by experts in the subject matter at hand. Accordingly, although you might use ideas you find in Wikipedia to guide your subsequent searches, as a general rule *you should not use Wikipedia as an authoritative source about your topic*.

An Internet search may also lead you to research articles and opinion papers that individual researchers have made available on the Internet, and you can typically print such documents or download them to your own computer. We caution you to keep in mind that such articles and papers vary widely in quality. Although most academic publications have a review process that enhances the quality of the research articles they include, many research reports posted on the Internet have not yet been reviewed or judged by professional colleagues. Obviously you will want to read *any* research report with a somewhat critical eye, but you should be especially careful when you find research reports on the Internet that you cannot verify as the work of credible scholars.

In your searches of the Internet, you will probably have to wade through many listings that are not terribly helpful. On the plus side, however, if you have access to the Internet from a home computer, you can browse anytime day or night—weekends, holidays, even 3:00 a.m. if you like. Libraries are sometimes closed, but the Internet is always open.

Whenever you find information on the Internet, you should make a note of the address (Uniform Resource Locator, or URL) at which you found it and the date on which you did so. Alternatively, some online documents posted within the past few years have a Digital Object Identifier number, or DOI—a unique number that enables you to find a document even if its precise location on the Internet has changed in the meantime. (For more information, see Chapter 12, or go to www.doi.org). Given the dynamic, ever-changing nature of the Internet, you may need to include such information in your reference list (more on this point in Chapter 12).

Locating Relevant Government Publications

We have already mentioned the potential use of government office websites. Another vast source of information you should not overlook is the depository libraries of the U. S. government. The U.S. Government Printing Office is the largest publisher of printed material in the world, and its publications are available through more than one thousand federal depository libraries (including many university libraries) throughout the United States and its territories. A few examples of the many publications available are *Statistical Abstract of the United States*, *Survey of Current Business*, and *Public Papers of the President*. Other publications provide information about environment and weather, science and technology, nutrition, health care, and education. You can obtain more information about government publications and designated libraries by going to the office's website (www.gpoaccess.gov). And you can purchase paper copies at the Government Printing Office's online bookstore (bookstore.gpo.gov).

Using Citations and Reference Lists of Those Who Have Gone Before You

No library or computer search—no matter how extensive—is foolproof. Ultimately any search depends on the particular keywords you use and the particular databases you include in your search. One additional—in our minds, *essential*—resource is the literature reviews of researchers whose own writings you have consulted. Such reviews, especially if they have been published recently, can give you valuable guidance about seminal research studies and cutting-edge ideas related to your research topic. As a rule of thumb, we urge you to track down *any references you see cited by three or more other researchers*. Such references are clearly influencing current work in your field and should not be overlooked.

The preceding paragraph brings us to another important point: Don't depend on what other authors say about a particular reference. Too often we have seen two or more authors misrepresent

the work of a particular researcher in the same, particular way; apparently, they are reading one another's descriptions of that researcher's work rather than reading the researcher's own words! Whenever possible, *go to the original source and read it yourself*.

Considering all of the resources we have described in this chapter, you might be thinking that you will be spending the next ten years conducting your literature review! Don't worry. In the two Practical Application sections that follow, we describe (a) how to conduct your literature search in an organized, efficient, and time-saving fashion and (b) how to distinguish between research reports that are and are not worth taking seriously.

PRACTICAL APPLICATION Conducting a Literature Search

In Chapter 2 we discussed how to select a research problem or question. We also suggested that most problems, taken as a whole, are fairly complex and can be more easily solved when they are divided into two or more subproblems.

The main problem and subproblems provide a way to focus your attention as you read the literature. One concrete and effective approach, using either paper and pencil or brainstorming software such as Inspiration, involves the following steps:

1. Write the problem in its entirety at the top of the page or computer screen.
2. Write each subproblem in its entirety as well.
3. Identify the important words and phrases in each subproblem.
4. Translate these words and phrases into specific topics you must learn more about. These topics become your "agenda" as you read the literature.
5. Go to the library catalog, its online databases, and the Internet to seek out resources related to your agenda.
6. Read!

Let's take an example. For his dissertation research, doctoral student Arthur Benton wanted to develop a means of using an existing measurement instrument, the Strong Vocational Interest Blank (SVIB), to identify potential cartographers for the federal government. The SVIB assesses a person's interests in a wide variety of activities; the profile of interests that it generates is then compared with the interests of people in various occupations to identify career paths in which the person might find satisfaction and success. At the time the study was conducted, interest scales for 54 different occupational groups had been developed for the SVIB, but none had been developed for cartographers. The SVIB was published in two versions, the SVIB for Men and the SVIB for Women; to limit the scope of the project, Mr. Benton focused only on the SVIB for Men. In his dissertation proposal, he presented the following research problem:

This researcher proposes to identify and evaluate the existing discrete interests among Federally employed male cartographers and to develop a scale for the revised Strong Vocational Interest Blank to aid recruitment of cartographers into Federal employment.

He then divided his problem into three subproblems:

The first subproblem is to determine whether male cartographers employed by the Federal Government have a discrete pattern of interests different from those of men in general, as measured by the Strong Vocational Interest Blank for Men.

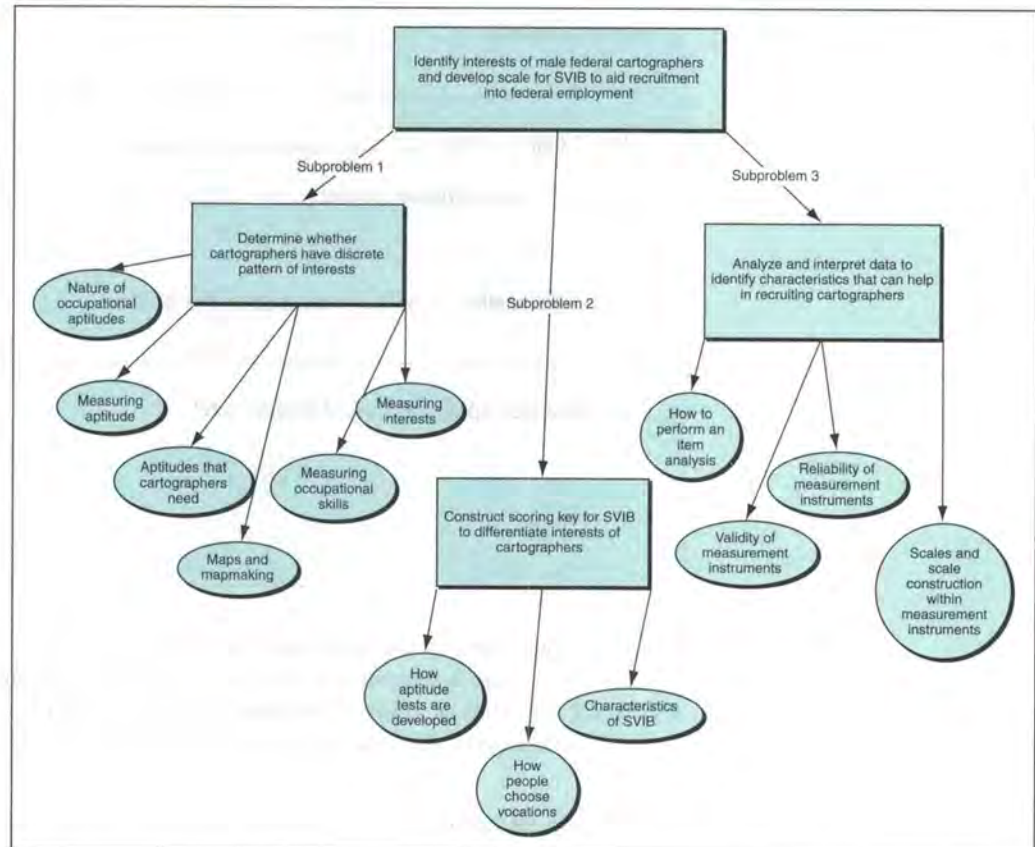
The second subproblem is to construct a scoring key for the Strong Vocational Interest Blank to differentiate the interests of cartographers from those of men in general and also from the interests of other occupational groups.

The third subproblem is to analyze and interpret the treated data so as to evaluate the discovered interests in terms of their discreteness in recruiting cartographers.

Figure 3.1 shows the literature-review agenda we have identified for the research problem and three subproblems. We used Inspiration software to create the agenda. The four rectangles

FIGURE 3.1

Using Inspiration software to prepare for a review of the literature



represent the research problem and its three subproblems. The ovals reflect the specific things we need to learn more about with respect to each subproblem.

The sequence of steps just described can help you keep your literature review within the realm of your research problem. It prevents you from wandering into other areas of the literature that, although potentially quite intriguing, may be of little or no use to you when the time comes to write your literature review.

Now that you have an idea of what to search for, let's discuss how to make your search efforts efficient and productive.

GUIDELINES Using Your Library Time Efficiently

Make no mistake about it: Conducting a thorough literature review takes a great deal of time. And almost certainly you will *not* be able to conduct your entire literature review from your home computer. So plan on going to the library—not just once but many times. Following are several suggestions for maximizing your efficiency at the library.

1. *Go to the library armed with data-gathering tools.* You will need a systematic way of recording the information you will be gathering—ideally one that will also allow you to easily organize that information. As is true for many aspects of the research process, you can use either paper and pencil or a computer (a laptop or tablet) to record what you find at the library.

A paper-and-pencil approach involves taking note paper or index cards on which to write what you find—not only information relevant to your topic but also information about the sources in which you find it: titles, authors, journal titles (if applicable), publication dates, and so on. Some people going the paper-and-pencil route have found it helpful to print up a large number of note-taking forms similar to that shown in Figure 3.2.

If, instead, you bring a laptop or tablet computer to the library, you can enter the information you gather directly into one or more computer files. Computerizing the information you obtain

later—are either unintelligible or so lacking in essential information that they are practically useless. If you are using paper to record your findings, write legibly or print clearly. If you can't distinguish between undotted *i*'s and closed-up *e*'s, or between carelessly made *a*'s and *o*'s, you may eventually find yourself scurrying back to sources for such details.

In addition to recording any essential bibliographic information you have not previously obtained, you will want to take sufficient notes about the content of the source to enable you to recreate its ideas when, later, you are writing your literature review. As you do so, you may find it helpful to keep track of the pages on which you have found certain ideas; furthermore, page numbers are essential if you are quoting a source word for word. In some cases, of course, a source will have so much information that you will simply want to make a photocopy of relevant pages¹ or, in the case of a book, check it out and take it home.



6. *Identify strategies for obtaining sources that are not immediately available.* We can almost guarantee that you won't find everything you need the first time around. Some books may be currently checked out to other users. Other books and a few journals may not be among the library's holdings. Following are several strategies to consider in such situations:

- *Put a hold on a checked-out book.* If you discover that a book has been checked out but is due to be returned shortly, some university catalog systems will let you put a *hold* on the book so that you are the next one in line to check it out when it comes in. If your catalog system doesn't provide this option, ask the librarian at the circulation desk to put a hold on the book for you. The library will contact you (typically by e-mail) as soon as the book is returned and will save it for you for a few days at the circulation desk.
- *Recall the book.* Many university libraries allow professors to check out books for an entire semester or academic year, but they must return them earlier if someone else wants to use them. If you discover that a book you need is checked out for a long period of time—perhaps for several months or longer—you can ask for the book to be *recalled* so that you can use it. Some university catalog systems let you recall a book yourself; for example, if you see a “request” button on the catalog page for the book, clicking on it will probably initiate a recall notice to the person who currently has it.
- *Submit a request for the source through interlibrary loan.* Almost all libraries have cooperative arrangements to exchange resources with other libraries. In all likelihood, you will be able to order a book or journal article through such **interlibrary loan** using your library's catalog system. A “low-tech” alternative is to make the request through your library's interlibrary loan office. Typically you can get books and journal articles from other libraries at little or no charge except, perhaps, for photocopying.
- *Check Google Books (books.google.com) for older, out-of-print books.* Some books available on Google Books are no longer protected by copyright and are typically available free of charge. Those still under copyright are available only to the extent that the copyright holders have given permission; in such cases, you may be able to view excerpts from the book but may need to pay a small fee to see the entire book.
- *Order books through a bookstore.* Most bookstores will order any book that is currently in print and obtain it for you within a week or two. You can also order both new and used books through such online booksellers as Amazon (www.amazon.com) and Barnes and Noble (www.barnesandnoble.com).
- *Use an online document delivery service.* Some of the online databases listed in Table 3.3 (e.g., Academic Search Premier, JSTOR, PsycINFO) provide electronic copies of articles from selected journals. Others (e.g., ERIC) provide electronic copies of conference papers and other nonpublished works. If you find one or more doctoral dissertations that pertain to your research problem, you can order complete copies through ProQuest Dissertations and Theses: Full Text (www.proquest.com). Some of these services may be available to you free through your university library; others (such as ProQuest) will charge you a fee.

¹In the United States, federal copyright law allows one copy for personal use.

As you conduct your literature review, you will undoubtedly need to repeat this cycle of steps several times. With each go-round, however, you will become more and more of an expert about the topic you are pursuing. You will also become increasingly knowledgeable about the library and its resources.

PRACTICAL APPLICATION Evaluating the Research of Others

An important skill for any researcher is the ability to review the work of others and evaluate the quality of their methods, results, and conclusions. *Never take other people's conclusions at face value; determine for yourself whether their conclusions are justified based on the data presented.* Critically examining what others have done has three distinct benefits:

- It can help you determine which ideas, research findings, and conclusions you should take seriously and which you can reasonably discredit or ignore.
- It can help you reconcile inconsistent findings obtained in previous research studies.
- It can give you some ideas about how you might improve your own research efforts.

As you proceed through the rest of this book and learn more about research methodology, you will become increasingly knowledgeable about the kinds of conclusions that are and are not warranted from various methodologies and types of data. At this point, you may be able to judge the work of other researchers only in a fairly superficial fashion. Even so, there's no better time than the present to begin examining other researchers' work with a critical eye. We suggest that you begin to sharpen your evaluation skills by locating several research articles relevant to your interests. As you read and study the articles, consider the questions in the following checklist.



CHECKLIST

Questions to Consider When Evaluating Research

- _____ 1. In what journal or other source did you find the research article? Was it reviewed by experts in the field before it was published? That is, was the article in a *juried* (refereed) publication?

- _____ 2. Does the article have a stated research question or problem? That is, can you determine the focus of the author's work?

- _____ 3. Does the article describe the collection of new data, or does it describe and synthesize previous studies in which data were collected?

- _____ 4. Is the article logically organized and easy to follow? What could have been done to improve its organization and readability?

- _____ 5. Does the article contain a section that describes and integrates previous studies on this topic? In what ways is this previous work relevant to the research problem?

_____ 6. If the author explained procedures that were followed in the study, are these procedures clear enough that you could repeat the work and get similar results? What additional information might be helpful or essential for you to replicate the study?

_____ 7. If data were collected, can you describe how they were collected and how they were analyzed? Do you agree with what was done? If you had been the researcher, what additional things might you have done?

_____ 8. Do you agree with the interpretation of the results? Why or why not?

_____ 9. Finally, think about the entire article. What is, for you, most important? What do you find most interesting? What do you think are the strengths and weaknesses of this article? Will you remember this article in the future? Why or why not?

Organizing the Information You Collect

Imagine entering a new library to locate some books and articles, but this library doesn't have an organized system for storing its resources. Printed materials are placed on the most convenient shelves, rather than in any designated location. The librarians leave it up to the users to look through all this material until they find what they need. Naturally you would very quickly become frustrated trying to find what you want. You will encounter similar frustration—albeit on a smaller scale—if you don't organize the information you gather from your literature search. Thus, as you begin to accumulate a substantial amount of information about your topic, you will want to begin organizing it as well.

Putting your sources and notes on a computer, either while you are at the library or at a later time, can often help you keep track of and organize your sources more quickly and efficiently than using paper-and-pencil notes. Computer databases are one extremely useful tool in this regard.

Creating a Computer Database for Your Related Literature



Earlier we spoke of the databases available through online searches. In fact, you can create your *own* database for the literature you read. In general, a **database** is a collection of information that is organized in an easily accessible manner.

A simple analogy might be helpful. Most of us have an address book or file in which we keep track of the names, addresses, and telephone numbers of friends, relatives, and other important people in our lives. For some of us it takes the form of an actual notebook or card file, whereas for others it is a digital address book on a cell phone, personal digital assistant (PDA), tablet computer (e.g., iPad), or other electronic device. Regardless of its form, an address book has specific, designated spots for each kind of information—name, address, telephone number, and so on—for every individual we include in it.

You might think of an electronic database as an electronic card file. The entire file—known as a *data file*—contains many “cards,” or *records*. There is one record for each separate entity you want to keep track of. If you wanted to use a database as an address book, you would have a separate record for each person; if you wanted to use a database to keep track of your various resources (books, articles, etc.), you would have a separate record for each resource. Each record, in turn, contains specified kinds of information (e.g., name, address, phone number, etc.; or author, title, date, etc.); the “slots” for these various pieces of information are called *fields*.

Several database software programs are available specifically for creating bibliographic databases. Some are available commercially; examples are Biblioscape (www.biblioscape.com), EndNote (www.endnote.com), OneNote (office.microsoft.com/en-us/onenote), and ProCite (www.procite.com). By searching the Internet, you can also find and download “freeware” (free software) that will serve your purpose; as this book goes to press, two examples are BiblioExpress (www.biblioscape.com/biblioexpress.htm) and Bibliography Writer (available at many Internet websites; simply google “Bibliography Writer software”).

Bibliographic database programs typically have fields for all the information you would need in order to create a reference list—author, title, date, journal title, page numbers, and so on. In fact, these programs can also *create your reference list for you!* Most programs allow you to use whatever format your institution or academic discipline requires (e.g., APA style, MLA style; more on such styles in Chapter 12). Furthermore, these programs have additional fields for call numbers, keywords, any notes you take while reading an item, and (often) any graphics you want to include.

By putting the information you collect in an electronic database, you can find anything you want in a matter of seconds. For example, perhaps you want to find books and articles written by a specific author. You can search the author field in the entire database; the program will identify every record for that author. In addition, you can rapidly sort the data by whatever field you choose. For example, if you want all sources listed by publication date, the date field on each record within the database is identified and the records are rearranged in chronological order.

Knowing When to Quit

Certainly you shouldn't read only one or two articles and think you are done. Get used to looking for and reading new research reports; for a researcher, this is a lifelong endeavor. Always, always look for additional things you can learn.

But at some point, of course, you must bring your literature review to a close. When might you know that you have completed a reasonably thorough literature review? In theory, the answer might be “Never.” As long as people continue to conduct research related to your topic, there will always be additional information that might be helpful to you. From a practical standpoint, however, you must finish your search, use what you have obtained, and get on with the study itself. How will you know when that time has arrived?

The best advice we can give you is this: *Look for repetitive patterns in the materials you are finding and reading.* As you read more and more sources, eventually familiar arguments, methodologies, and findings will start to appear. Perhaps you will see the same key people and studies cited over and over. You will get a feeling of *déjà vu*—“I've seen this (or something very similar to it) before.” When you are no longer encountering new viewpoints, you can be reasonably sure that you are familiar with the critical parts of the literature.

Organizing and Synthesizing the Literature into a Cohesive Review

Too many literature reviews do nothing more than report what other people have done and said. Such reviews, which are typically written by novice researchers, go something like this:

In 1995, Jones found that such-and-such. . . . Also, Smith (1998) discovered that such-and-such. . . . Black (2006) proposed that so-on-and-so-forth. . . .

We learn nothing new from such a review; we would be better off reading the original books, articles, and other sources for ourselves.

In a good literature review, the researcher doesn't merely report the related literature. He or she also *evaluates, organizes, and synthesizes what others have done*. A checklist earlier in the chapter gave you a taste of what the *evaluation* component involves. But in addition to evaluating what you read, you must also *organize* the ideas you encounter during your review. The subproblems within your main problem can, in many cases, provide a general organizational scheme you can use. Looking at how other authors have organized literature reviews related to your topic can be helpful as well.

Finally, and perhaps most importantly, you must *synthesize* what you have learned from your review. In other words, you must pull together the diverse perspectives and research results you have read into a cohesive whole. Here are some examples of what you might do:

- Show how approaches to the topic have changed over time.
- Compare and contrast varying theoretical perspectives on the topic.
- Describe general trends in research findings.
- Identify discrepant or contradictory findings, and suggest possible explanations for such discrepancies.
- Identify general themes that run throughout the literature.

When you write a literature review that does such things, you have contributed something new to the knowledge in the field even *before* you have conducted your own study. In fact, a literature review that makes such a contribution is often publishable in its own right. (We talk more about writing for publication in Chapter 12.)

PRACTICAL APPLICATION Writing the Literature Review

Once you have read, evaluated, organized, and synthesized the literature relevant to your research problem, you will, of course, eventually need to begin writing the section or chapter that describes the literature you have examined. We offer several guidelines to help you in the process.

GUIDELINES Writing a Clear and Cohesive Literature Review

As university professors, we authors have written many literature reviews ourselves. We have also read countless master's theses and dissertations written by novice researchers. From such experiences, we have developed the following general guidelines for writing a solid review of the related literature.

1. *Get the proper psychological orientation.* Be clear in your thinking. Know precisely what you are trying to do. The review of the related literature section is a discussion of the studies, research reports, and scholarly writings that bear directly on your own research effort.

You might think of your written review of related literature as a description for one or more peers about what others have written in relation to what you plan to do. Viewing the literature section in this way can help both you and your readers see your own effort within the context of the efforts of researchers who have preceded you.

2. *Have a plan.* Writing a review of the related literature takes planning and organization. Before beginning to write the review of the related literature, create an outline of the topics you intend to address and the points you intend to make. Perhaps one of the best guides for such an outline is the problem itself, as well as its subproblems. A careful consideration of the problem and subproblems should suggest relevant areas for discussion and the order in which they should be addressed (e.g., look once again at Figure 3.1).

Begin your discussion of the related literature from a comprehensive perspective, like an inverted pyramid—broad end first. Then, as you proceed, you can deal with more specific ideas and studies and focus in more and more on your own particular problem.

Throughout your discussion of the related literature, your organizational scheme should be crystal clear to both you and your readers. For example, start off with an *advance organizer*—an overview of the topics you will discuss and the sequence in which you will discuss them (see Chapter 1). And use headings and subheadings throughout your literature review to alert readers to the particular topics that each section addresses.

Early in the review, you will probably want to consider the classic works—those groundbreaking studies that have paved the way for much of the research about the topic. Such studies give an overall historical perspective and provide a context for your own efforts. As an example, the scientists who developed the laser built on the efforts and writings of those who explored relativistic quantum mechanics and the nature of matter and light in electromagnetic fields.

3. *Emphasize relatedness to your research problem.* Keep your reader constantly aware of how the literature you are discussing is related to your own research study. Point out precisely what the relationship is. Remember that you are writing a review of the *related* literature.

Literature reviews should never be a chain of isolated summaries of the writings of others (Jones says . . . ; Smith says . . . ; Black says . . .). Such is not a discussion of related literature; no attempt is made to demonstrate the relatedness of the literature to the problem being researched. If you cannot identify a relationship, you would do well to consider whether you should include the source at all.

As a way of helping you make frequent connections between your discussion of the literature and your own research project, we suggest the following procedure:

- a. Write your research problem prominently at the top of a sheet of paper or word processing document. In this location, it will be a constant reminder of the central axis around which everything else revolves.
- b. Divide the problem by numbering its various subparts.
- c. Divide the remainder of the page or word processing document into two columns. If you are using paper, you can simply draw a line down the middle. If you are using a word processing document, you might want to use the "insert table" function to create a table with two columns and numerous rows.
- d. In the left-hand column, list each source (book, article, etc.) in your review.
- e. In the right-hand column, opposite each source, note the particular subdivision(s) of the problem to which the source relates and your rationale for including it in the literature review.
- f. Add additional pages (for a paper list) or table rows (for a word processing document) as you need them until you have listed every source.

4. *Provide transitional phrases, sentences, or paragraphs that help your readers follow your train of thought.* If one idea, paragraph, or section leads logically to the next, say so! Furthermore, give readers some sort of signal when you change the course of your discussion. For instance, in a doctoral dissertation examining the various thinking processes that students might use when listening to a lecture, Nancy Thrailkill finished a discussion of the effects of visual imagery (mental "pictures" of objects or events) and was making the transition to a more theoretical discussion of imagery. She made the transition easy to follow with this sentence:

Although researchers have conducted numerous studies on the use and value of imagery in learning, they seem to have a difficult time agreeing on why and how it works. (Thrailkill, 1996, p. 10)

The first clause in this transitional sentence recaps the discussion that immediately preceded it, whereas the second clause introduces the new (albeit related) topic.

5. *Know the difference between describing the literature and plagiarizing it.* Our own experiences tell us—and research confirms our observations—that many novice researchers don't fully understand the various forms that plagiarism might take (Cizek, 2003; McGue, 2000). In particular, **plagiarism** involves either (a) presenting another person's work as being one's own or (b) insufficiently acknowledging and identifying the sources from which one has drawn while writing. Reproducing another person's work word-for-word without crediting

that person constitutes plagiarism, of course. But so, too, is making small, insignificant wording changes in someone else's sentence or paragraph a form of plagiarism. For example, early in Chapter 2 we say:

Some research projects are intended to enhance basic knowledge about our physical, biological, psychological, or social world or to shed light on historical, cultural, or aesthetic phenomena. . . . Such projects, which can advance theoretical conceptualizations about a particular topic, are known as **basic research**.

You would be plagiarizing our work if you said something like this without giving your source proper credit:

Basic research is intended to enhance basic knowledge about our physical, biological, psychological, or social world or to shed light on historical, cultural, or aesthetic phenomena. Such research can advance theoretical conceptualizations about a particular topic.

All you would have done here is to replace "Some research projects are" with "Basic research is" at the beginning of your first sentence and make a few tiny adjustments to the second sentence.

6. *Always give credit where credit is due.* Note the second part of our earlier definition of plagiarism: *insufficiently acknowledging and identifying the sources from which one has drawn while writing*. In writing the literature review, you must always, *always* credit those people whose ideas you are using or whose research results you are reporting. Such is true regardless of whether you are making use of printed materials, Internet resources, conference presentations, or informal conversations with others in your field. Omitting this crucial step leads your readers to infer that certain ideas are your own rather than those of someone else. Recall that all-important rule we presented in Chapter 2: *Absolute honesty and integrity are assumed in every statement a scholar makes*.

Citing other sources also means citing them *correctly*. Making major errors in citations—even if you do so unintentionally—constitutes plagiarism. For example, you must take care not to cite "Smith and Jones (2005)" when the correct citation is "Greene and Black (2007)." Sloppiness in your record keeping is no excuse. Although you haven't meant to, you are plagiarizing from Greene and Black's work.

The specific way in which you give credit to scholars whose ideas you are presenting—for instance, whether you use footnotes or, as we do, citations in parentheses within the body of the text—must depend on the particular style manual you are using, which, in turn, depends on your particular discipline. We provide more details about various style manuals in Chapter 12 (e.g., see Table 12.1).

7. *Minimize your use of direct quotations from other people's work.* Sometimes you may decide that someone else's prose captures an idea so well or so vividly that you want to present it word for word. Occasionally, too, certain excerpts provide examples of a point you are trying to make about the literature in general; such is the case when we authors present excerpts from students' dissertations and master's theses in this book. You can legitimately use other people's words if you present them within quotation marks (for a phrase or sentence) or in an indented passage (for a longer quotation). For example, we used the indentation strategy earlier when we presented Thrailkill's transitional sentence in Guideline 4. Notice that we immediately cited the source of the sentence. Consistent with our use of APA style in this book, we gave the author's last name, the date of her dissertation, and the page number on which we found the sentence. That information would be sufficient for any of our readers to find the exact source in our reference list (located near the end of the book) and, after obtaining a copy of Thrailkill's dissertation, finding the actual sentence on page 10.

You should use lengthy quotations only when you have a very good reason—for instance, when the specific words that an author uses are as important as the ideas that the author presents. The law allows what is known as *fair use* of a quotation, but some publishers have their own rules of thumb about how much material you can quote without their permission. When in doubt, check with the publisher or other copyright holder.²

²Many publishers now include on their websites their guidelines about what and how much you can use without seeking their permission. If you do need their permission for what you want to use, you can often submit a permission request online.

All too often, however, we have seen literature reviews that appear to be little more than a sequence of quoted excerpts from various published sources. *Use quotations only when you have a very good reason.* Consistently using other people's words, even when you give those researchers appropriate credit, can convey the impression that you aren't willing to take the time to write a clear, cohesive literature review on your own.

As important as what others say about their research, and perhaps even more important, is what *you* say about their research. Your emphasis should always be on how a particular idea or research finding relates to your own problem—something only *you* can discuss.

8. *Summarize what you have said.* Perhaps the most important question any researcher can ask—and *should* ask throughout the progress of the research study—is, What does it all mean? In a thesis or dissertation, every discussion of related literature should end with a brief summary section in which you gather up all that has been said and describe its importance in terms of the research problem. Under the heading "Summary," you can condense your review into a synopsis of how the literature on your topic contributes to an understanding of the problem you are trying to address.

9. *Remember that your first draft will almost certainly not be your last draft.* Here we are simply repeating a point we made in Chapter 1—a point that applies to your literature review as well as to any other part of your research report. First drafts almost inevitably leave a lot to be desired, in part because (as we also noted in Chapter 1) the human mind can handle only so much information at any single point in time.

Imperfections in a first draft are virtually unavoidable. In fact, we urge you to write a first draft even before you have completely finished your literature review. Writing a first, incomplete draft can help you identify parts of the literature that are still unclear to you and places where you may need additional information or citations. One strategy we authors use as we write a literature review is to leave blanks for information we realize we still need, mark the blanks in bold, red font or with Post-it notes, and then make a final visit to the library (either to the actual building or to its online resources) in order to fill them in.

Even when you have obtained all the information you need for a complete review, you will typically not be able to express your thoughts with total clarity the first time around. Write the review, print it out, and let it sit for a few days. Then reread it with a critical eye, looking for places where you have been ambiguous, incomplete, or self-contradictory.

10. *Ask others for advice and feedback.* In this book we frequently suggest that you seek feedback from other people, and your literature review is no exception. Talk with others about what you have found, ask others to read an early draft, and get ideas about additional avenues you may need to explore. Use e-mail to contact people who have an interest in this area of study (e.g., contact the authors of studies that have influenced your own work). Explain where you are working and what you are working on, send them a copy of what you have written, and ask for their feedback and suggestions. You will be amazed at how helpful and supportive people can be when you tell them you have read their work and would appreciate their opinion.

A Sample Research Proposal

At this point, it may be helpful to look at excerpts from what is, in our view, a well-written review of the related literature for a doctoral dissertation proposal. The author of the review, Kay Corbett, wanted to identify possible relationships between cognitive development and motor development (i.e., between the development of children's thinking abilities and that of their movement patterns), especially between ages 4 and 8. Thus, the literature review focused on both the cognitive and motor development of young children.

Two qualities of the proposal are particularly worth noting. First, the author does not present the studies she has read in a piecemeal, one-at-a-time fashion; instead, she continually synthesizes the literature into a cohesive whole. Second, the author's organizational scheme is obvious throughout; she uses an advance organizer, numerous headings and subheadings, and transitional paragraphs to help the reader follow her as she moves from one topic to the next.

Excerpts from the proposal itself appear on the left-hand side and our commentary appears on the right. The ellipses (. . .) indicate places where we have omitted portions of the text. In some cases, we have summarized the content of what we've omitted within brackets.

Dissertation ANALYSIS

1

REVIEW OF LITERATURE

The literature review will include three areas: (a) empirical studies relating motor and cognitive development, (b) motor development, and (c) the neo-Piagetian theories of development as they relate to both motor and cognitive development. The present review is limited to investigations of children within the four- to eight-year-old age range. Studies targeting children with special needs are excluded.

[The remainder of the chapter is divided into three main sections: "Motor and Cognitive Development," "The Development of Gross Motor Skills," and "The Neo-Piagetian Theories of Development." We pick up the chapter midway through the section on "The Development of Gross Motor Skills."]

The Development of Gross Motor Skills

... [T]he early childhood period is when many fundamental motor patterns are most efficiently learned. During this age period, children must have daily practice and participation in movement education programs to develop the fundamental movement skills to a mature pattern (Gallahue, 1993, 1995b, 1996; Halverson & Robertson, 1984; Haubenstricker & Seefeldt, 1986; Haywood, 1993; Miller, 1978, cited in Gallahue, 1989; Williams, 1983). If opportunity for this practice is not provided, children may move into adolescence with immature motor patterns that will hinder their ability to enter games or sports activities (Gallahue, 1995a; Haubenstricker & Seefeldt, 1986). Mature patterns can be acquired later in the developmental life span, but it requires much more time and practice to relearn the patterns.

... The fundamental patterns for the 4- to 8-year-old age range include four categories of movements: (a) locomotor movements, (b) stability movements, (c) manipulative movements, and (d) axial movements (Gallahue, 1995b).

The locomotor movements acquired and/or refined during this period of childhood are running, jumping, hopping, galloping and sliding, leaping, skipping, and climbing (Gallahue, 1995b). These movements "involve a change in location of the body relative to a fixed point on the surface" (Gallahue, 1989, p. 46).

Stability movements refer to the "ability to maintain one's balance in relationship to the force of gravity even though the nature of the force's application may be altered or parts of the body may be placed in unusual positions" (Gallahue, 1989, p. 494). Stability movements include weight transfer skills (Haywood, 1993). Weight transfer skills include inverted supports, in which the body assumes an upside-down position for a number of seconds before the movement is discontinued. "Stabilization of the center of gravity and maintenance of the line of gravity within the base of support apply to the inverted posture as well as to the erect standing posture" (Gallahue, 1989, p. 275).

Comments

The author begins with an advance organizer that outlines the upcoming chapter and describes the scope of the literature review.

Notice how the three sections correspond roughly to the "a," "b," and "c" that the author described in the first paragraph.

Notice how the author integrates and summarizes the results of several studies—an approach that is quite appropriate when researchers have all come to a similar conclusion. Several of the studies are (in 1997, the year the proposal is written) quite recent, communicating the (probably accurate) impression that the author is presenting an up-to-date perspective on the topic. A citation such as "Miller, 1978, cited in Gallahue, 1989" should be used only when the original source (in this case, Miller, 1978) is difficult to obtain.

Notice how this sentence alerts the reader to the organizational structure that follows.

To indicate that she is using Gallahue's definition of locomotor movements, the author uses quotation marks and, within the citation, lists the page on which she found the definition.

The author quotes Gallahue several times. As a general rule, you should limit your quotations to situations in which an author's presentation of ideas or information is exceptionally vivid, precise, or in some other way highly effective. Otherwise, just paraphrase what your sources have said, giving them appropriate credit, of course, for their ideas.

Other stability movements are dodging, one-foot balancing, beam walking, and rolling.

The manipulative movements involve giving force to objects and receiving force from them (Gallahue, 1989). Movements practiced during childhood are overhand throwing, catching, kicking, striking, dribbling, ball rolling, trapping (feet or body is used to absorb the force of the ball instead of the hands and arms), and volleying.

The axial movements are "movements of the trunk or limbs that orient the body while it remains in a stationary position" (Gallahue, 1989, p. 271). Bending, stretching, twisting, turning, swinging, swaying, reaching, and lifting are all axial movements. They are used in combination with other movements to execute more complex movement skills.

Researchers investigating the development of fundamental movement skills focus on qualitative changes as children's developing movement patterns become more smooth and efficient. The following section will review studies investigating the development of fundamental movement patterns in children four to eight years of age.

Development of Locomotor Skills

The locomotor skills, from earliest acquisition until mature patterns are established, develop through qualitatively different stages (e.g., Gallahue, 1995b; Haywood, 1993; Haubenstricker & Seefeldt, 1986). The studies reviewed investigated qualitative changes that occur as fundamental locomotor patterns are developed.

Walking. The mature walking pattern is achieved between the fourth and seventh years (Eckert, 1987; Guttridge, 1939; Wickstrom, 1983; Williams, 1983). At this level, there is a reflexive arm swing, narrow base of support (feet are placed no further apart than the width of the shoulders), the gait is relaxed, the legs lift minimally, and there is definite heel-toe contact (Gallahue, 1989). Although the mature pattern is achieved during the early childhood period, walking is not targeted in movement education programs as a skill needing concentrated focus (Gallahue, 1989, 1996; Werder & Bruininks, 1988).

Running. Many investigators have studied the running pattern. Robertson and Halverson (1984) document the development of running by rating arm action separately from leg action but base the documentation on earlier work (Wickstrom, 1983; Seefeldt et al., 1972, cited in Gallahue, 1989). Gallahue (1995b) proposes a whole-body sequence of development based on the same earlier work. Running patterns develop from flat-footed, uneven patterns with arms swinging outward to smoother patterns with step length increased, and a narrower base of support. The mature pattern includes a reflexive arm swing, narrow base of support, relaxed gait, minimal vertical lift, and a definite heel to toe contact. Several University of Wisconsin studies of children between 1.5 and 10 years of age have documented the qualitative changes in the running pattern (Haywood, 1993).

Jumping. Early developmentalists defined age norms for children's jumping achievements (Wickstrom, 1983). The children step down from a higher surface from one foot to the other, before jumping off the floor with both feet. Then they learn to jump from progressively greater heights onto both feet. Later, they can jump forward, and over objects (Haywood, 1993).

This paragraph helps the reader follow the author's train of thought as she makes the transition from one topic to another, related one.

Here the headings "Walking," "Running," "Jumping," and so on, under the more general "Development of Locomotor Skills" heading, communicate quite clearly how the section is organized.

Notice how, in this paragraph, the author synthesizes what previous researchers have found. She intentionally does not describe studies one by one because they all point to the same conclusion. The result is a smooth-flowing, easy-to-read, summary of work that has been done related to the topic.

In the second and third sentences of the "Running" paragraph, the verbs document, base, and proposes should be documented, based, and proposed (past tense). In general, use past tense (e.g., proposed or has proposed) to describe what has been done in the past. Use present tense to represent general ideas that are not restricted to a single time period. For instance, present tense is appropriately used in the paragraph's fourth sentence ("Running patterns develop from...").

Developmental sequences in both the horizontal and vertical jump are based on research on the standing long jump (Clark & Phillips, 1985; Hellebrandt et al., 1961; Seefeldt et al., 1972, cited in Gallahue, 1989; Wickstrom, 1983; Robertson, 1984; Robertson & Halverson, 1984). The one-footed takeoff is one salient characteristic of the earliest jump pattern and persists in some children well into their elementary school years (Robertson, 1984). The jumping motor patterns develop during the ages from two to seven years (Haubenstricker & Seefeldt, 1986). Some elements of the jumping pattern remain stable across ages and type of jump, specifically, 3-, 5-, 7-, and 9- year olds and adults all use the same pattern of leg coordination. All people do not obtain a mature pattern in childhood. In fact many immature patterns are found in adults (Haywood, 1993)....

[The author devotes additional sections to "Hopping," "Gallop and Sliding," "Skipping," and "Leaping and Climbing." She then proceeds to the development of other categories of motor skills and, eventually, to a discussion of the third major topic of the chapter—neo-Piagetian theories.]

NOTE: Excerpt is from a research proposal submitted by Katherine E. Corbett to the University of Northern Colorado, Greeley, in partial fulfillment of the requirement for the degree of Doctor of Philosophy. Reprinted with permission.

In this paragraph the author clarifies the types of studies (i.e., research on the standing long jump) on which certain conclusions have been drawn. By doing so, she helps the reader put the conclusions in perspective and, perhaps, judge the quality of those conclusions.

Throughout the chapter, various levels of headings continue to be important guideposts that reflect this overall organizational scheme.

For Further Reading

- Boote, D. N., & Beile, P. (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. *Educational Researcher*, 34(6), 3–15.
- Chan, L. M. (1999). *A guide to the Library of Congress classification* (5th ed.). Englewood, CO: Libraries Unlimited.
- Fink, A. (2010). *Conducting research literature reviews: From the Internet to paper* (3rd ed.). Thousand Oaks, CA: Sage.
- Galvan, J. L. (2009). *Writing literature reviews: A guide for students of the social and behavioral sciences* (4th ed.). Los Angeles: Pyrczak.
- Janes, J. (2003). *Introduction to reference work in the digital age*. New York: Neal-Schuman.
- Katz, W. A. (2002). *Introduction to reference work: Basic information sources* (Vol. 1, 8th ed.). New York: McGraw-Hill.
- Reed, J. G., & Baxter, P. M. (2003). *Library use: A handbook for psychology* (3rd ed.). Washington, DC: American Psychological Association.
- Stebbins, L. F. (2005). *Student guide to research in the digital age: How to locate and evaluate information sources*. Westport, CT: Libraries Unlimited.
- Taylor, A. G. (2006). *Introduction to cataloging and classification* (10th ed.). Westport, CT: Libraries Unlimited.

Now go to MyEducationLab at www.myeducationlab.com to take assessments to evaluate your mastery of chapter concepts. Assignments and Activities exercises and Building Skills in Reading Research exercises are also available to help you master concepts and skills. Feedback for many of these exercises is provided so that you can see why your answers are correct or incorrect.

4

Planning Your Research Project

Before the construction of a building, an architect develops a detailed set of plans. These plans enable the builder to erect a strong, well-designed structure. Researchers should be no less detailed and precise in the planning of a research design. Plans, specifications, and design: All of these serve well the architect, the builder, and the researcher alike.

Architectural planning and research planning have much in common. Each requires a conceptualization of the overall organization of a project and a detailed specification of the steps to be carried out; only after such meticulous planning can work on the project actually begin. For successful completion, a building requires plans that are clearly conceived and accurately drawn. A research project should be no less completely conceptualized and precisely detailed.

Planning a General Approach

When we talk about a general strategy for solving a research problem, we are talking about a **research design**. The research design provides the overall structure for the procedures the researcher follows, the data the researcher collects, and the data analyses the researcher conducts. Simply put, research design is *planning*.

Nothing helps a research effort be successful so much as carefully planning the overall design. More research effort is wasted by going off half-prepared—with only a vague set of ideas and procedures—than in any other way. You will be much more efficient and effective as a researcher if you identify your resources, your procedures, and the forms that your data will take—always with the central goal of solving your research problem in mind—at the very beginning of your project.

The Basic Format of All Research

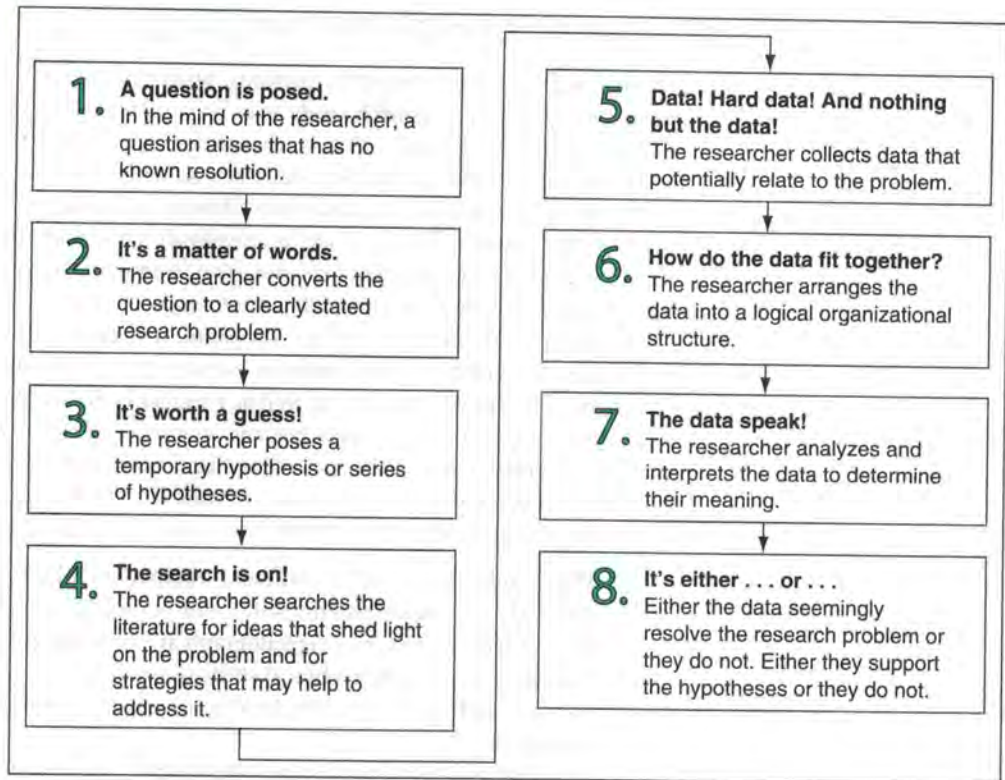
The research process follows a basic format. No matter which academic discipline gives rise to the research endeavor, the general research procedure is fundamentally the same. Although you will sometimes hear people refer to “social science research,” “nursing research,” or “marketing research,” in fact the search for information to solve a problem or to answer a question seldom fits squarely within the bounds of a single academic discipline.

In planning a research design, the researcher in quest of new knowledge cannot be shackled by discipline-specific methodological restraints. The course of a research project will frequently lead the investigator into new and unfamiliar territories that have historically been associated with other content areas. The sociologist trying to resolve a problem in sociology may come face to face with problems that are psychological or economic. The educational researcher exploring the causes of a learning disability may need to consider the domains of neurophysiology, psychopathology, endocrinology, and family counseling. On the way to finding a solution for a problem in criminology, the student in criminal justice may venture into the realms of abnormal psychology and behavioral genetics. Any good researcher must be *eclectic*, willing to draw on whatever sources seem to offer productive methods or evidence for resolving the research problem.

Figure 4.1 presents the basic format for the research process. It is important to note, however, that a research project does not always follow these steps in the exact sequence depicted

FIGURE 4.1

The basic format of the research process



here. Especially in the case of qualitative research—an approach we describe later in the chapter—a researcher may order the steps somewhat differently. For instance, in some forms of qualitative research, the researcher formulates a hypothesis and reads much of the literature only *after* collecting a substantial amount of data (more on this point in Chapter 6).

Instead of limiting their thinking to departmentalized knowledge, researchers might do much better to think of problems as arising out of broad generic areas within whose boundaries all research falls: people, things, records, thoughts and ideas, and dynamics and energy. Let's briefly consider some research problems that may fall within each of these areas.

- **People.** In this category are found research problems relating to children, senior citizens, families, communities, cultural groups, ancestors, employees, management, mental and physiological processes, learning, motivation, social and educational problems, crime, rehabilitation, medical treatments, nutrition, language, religion.
- **Things.** In this category are found research problems relating to animal and vegetable life, viruses and bacteria, inanimate objects (rocks, soil, buildings, machines), matter (molecules, atoms, subatomic matter), stars, galaxies.
- **Records.** In this category are found research problems relating to newspapers, journals, letters, manuscripts, registers, speeches, interviews, minutes, legal documents, financial and corporate statements, census reports, archeological remains, sketches, drawings, paintings, music.
- **Thoughts and ideas.** In this category are found research problems relating to concepts, theories, perceptions, opinions, beliefs, reactions, issues, semantics, poetry, cartoons, caricatures.
- **Dynamics and energy.** In this category are found research problems relating to human interactions, metabolism, chemical reactions, radiation, radio and microwave transmissions, quantum mechanics, thermodynamics, hydrodynamics, hydrologic cycles, atomic and nuclear energy, wave mechanics, atmospheric and oceanic energy systems, solar energy, black holes.

We do not intend the preceding lists to be mutually exclusive or all-inclusive. We merely present them to give you an idea of the many research possibilities that each category suggests.

Research Planning versus Research Methodology

Do not confuse overall research planning with research methodology. Whereas the general approach to *planning* a research study may be similar across disciplines, the techniques one uses to collect and analyze data—that is, the *methodology*—may be specific to a particular academic discipline. Such is the case because data vary so widely in nature. You cannot deal with a blood cell in the same way that you deal with a historical document, and the problem of finding the sources of Coleridge's "Kubla Khan" is entirely different from the problem of finding the sources of radio signals from extragalactic space. You cannot study chromosomes with a questionnaire, and you cannot study attitudes with a microscope.

In planning the research design, therefore, it is extremely important for the researcher not only to choose a viable research problem but also to consider the kinds of data that an investigation of the problem will require, as well as reasonable means of collecting and interpreting those data. Many beginning researchers become so entranced with the glamour of the problem that they fail to consider practical issues related to data availability, collection, and interpretation.

Comparing the brain wave patterns of children who are gifted versus those of average ability may be an engaging project for research, but consider the following issues:

- Will you be able to find a sufficient number of children who are willing to participate in the study and whose parents will grant permission for their children to participate?
- Do you have an electroencephalograph at your disposal?
- If so, do you have the technical skills to use it?
- Are you sufficiently knowledgeable to interpret the electroencephalographic tracings you obtain?
- If so, do you know how you would interpret the data and organize your findings so that you could draw conclusions from them?

Unless the answer to all of these questions is *yes*, it probably is better that you give up this project in favor of one for which you have the knowledge, skills, and resources to carry through to completion. Your research should be *practical* research, built on precise and realistic *planning* and executed within the framework of a clearly conceived and feasible *design*.

General Criteria for a Research Project

A good research project has four important qualities you can use as criteria in evaluating your own research plan. Here, briefly, are those qualities.

Universality. The research project should be one that might be carried out by *any* competent person. As a researcher, you are a director and producer—an agent whose function is to collect, organize, and report what the collected data seem to indicate—but another, equally knowledgeable individual might take your place and complete the project with essentially the same results. (We will find an exception to this criterion later in the chapter, when we discuss the nature of data collection in qualitative research.)

Replication. The research should be repeatable. Any other competent researcher should be able to take your problem and, collecting data under the same circumstances and within the same parameters you have used, achieve results comparable to those you have obtained.

Control. The researcher must in some way isolate, or *control*, those factors that are central to the research problem. Control is important for replication: An experiment should be repeated under the identical conditions and in the identical way in which it was first carried out. Control is also important for consistency within the research design. For instance, if you want to compare the effects of two different treatments on some other factor, then you should keep everything else (aside from the specific treatments you are studying) as similar as possible.

Control is more easily achieved in some areas than in others. In the physical sciences, for example, such factors as temperature, pressure, electrical potential, humidity, and the like are

highly controllable. Control is more challenging—but certainly possible—in research areas concerned with human behavior.

Control is especially important in experimental research designs. Accordingly, we talk more about this issue in Chapter 9.

Measurement. The data should be able to be measured in some way. This, again, is often easily accomplished in the physical sciences. Measurement is typically less precise and less accurate in the humanities and social sciences. Later in the chapter we discuss strategies for enhancing measurement procedures in these areas.

The Nature and Role of Data in Research

Research is a viable approach to a problem only when data exist to support it. The term *data* is plural (singular is *datum*) and comes from the past participle of the Latin verb *dare*, which means “to give.” Data are those pieces of information that any particular situation *gives* to an observer.

What Are Data?

Researchers must always remember that data are not absolute reality—the pure, naked Truth that underlies all the phenomena we observe. Rather, data are *manifestations* of that reality. No one has ever looked upon Truth itself. We are like those individuals who live in a dungeon, across the floor of which a beam of sunlight passes. That light gives us an idea of what the sun must be like, but if we can never behold the sun, we shall never know the difference between it and the shaft of light on the dungeon floor.¹

The researcher is in a factual dungeon. He or she will never be able to see the original source of the data. For instance, we often see what other people do—the behaviors they exhibit, the things they create, and the effects of their actions on others. But the actual people “inside”—those individuals we will never know!

Research seeks, through data, to discover underlying truths. Yet such is probably an endless pursuit. Experienced researchers are constantly aware that the Truth they most ardently seek is forever just beyond what is represented by the data and, hence, just beyond human grasp. For instance, the scientist probing the nature of subatomic matter may detect tiny entities that make up larger bits of matter. Yet the scientist knows, too, that such entities are probably made up of sub-sub-subentities that beckon yet evade measurement and investigation.

The mind yearns to understand the Truth. To pursue that goal, we have chosen the path of research. But the path always ends at the farthest reaches of the data, which are at the brink of the canyon in whose depths lies the inaccessible, ultimate Truth.

Data Are Transient and Ever Changing

Whenever we look at data analytically, we gain new insights. But at the same time, we also discern new problems that demand further research.

Data are not only elusive but also transient. Data that the researcher is permitted to glimpse may exist for only a split second. Consider, as an example, a sociologist who plans to conduct a survey in order to learn about people’s attitudes and opinions in a certain city. The sociologist’s research assistants begin by administering the survey in a particular city block. By the time they move on to the next block, the data they have collected are already out of date. Some people in the previous block who indicated that they held a particular opinion have now changed their minds and have a somewhat different opinion. They may have seen a television program or heard a discussion that has changed it. Some people have moved away, and others have moved in; some

¹For readers interested in philosophy, our dungeon analogy is based loosely on Plato’s Analogy of the Cave, which he used in Book VII of *The Republic*.

have died, and others have been born. Tomorrow, next week, next year—what we thought we had “discovered” may have changed completely.

Thus is the transient nature of data. We catch merely a fleeting glance of what seems to be true at one point in time but is not necessarily true the next. Even the most carefully collected data may have an elusive quality about them; at a later point in time they may have no counterpart in reality whatsoever. Data are volatile: They evaporate quickly.

Primary Data versus Secondary Data

The researcher's only perceptions of Truth are various layers of truth-revealing facts. In the layer closest to the Truth are **primary data**; these are often the most valid, the most illuminating, the most truth-manifesting. Farther away is a layer consisting of **secondary data**, which are derived not from the Truth itself, but from the primary data.

Earlier we used the analogy of the researcher as someone who sits in a dungeon and tries to understand the sun only by looking at a shaft of sunlight that falls on the floor. This direct beam of sunlight represents the *primary data*. Although the shaft is not the sun itself, it has come directly from the sun. But now imagine that the imprisoned researcher sees the sunlight, not as a direct beam, but as a pattern of shimmering light on the floor. The sunlight (primary data) has fallen onto a shiny surface and then been reflected—distorted by the imperfections of the shiny surface—to an image that is in some ways similar but in other ways dissimilar to the original shaft of light. This reflection of light is secondary data.

As another example, consider the following incident: You see a car veer off the highway and into a ditch. You have witnessed the entire event. Afterward, the driver says he had no realization that an accident might occur until the car went out of control. Neither you nor the driver will ever be able to determine the absolute truth underlying the accident. Did the driver have a momentary seizure of which he was unaware? Did the car have an imperfection that the damage from the accident obscured? Were other factors involved that neither of you noticed? The answers lie beyond an impenetrable barrier. The true cause of the accident may never be known, but the things you witnessed, incomplete as they may be, are primary data that emanated directly from the accident itself.

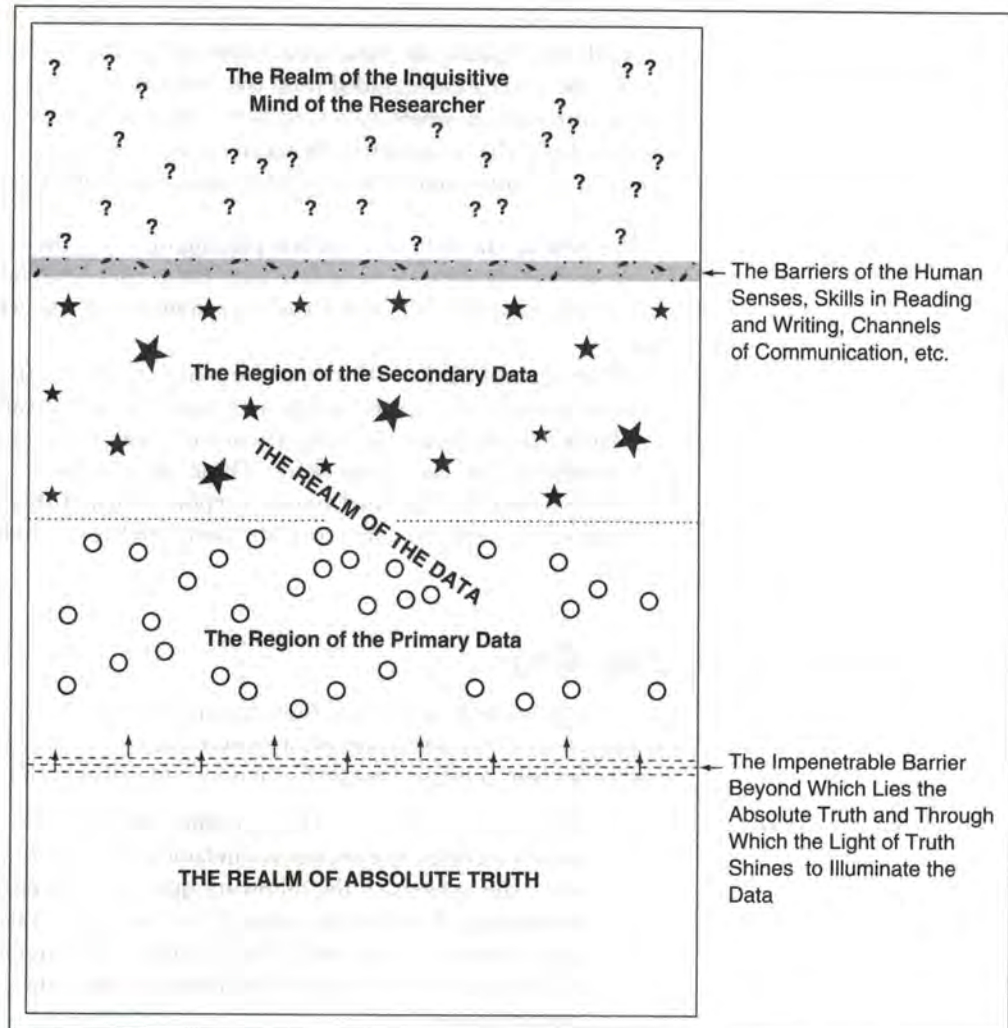
Now along comes a newspaper reporter who interviews both you and the driver. The reporter writes an account of the accident for the local paper. When your sister reads the account in the paper the next morning, she gets, as it were, the sunlight-reflection-on-the-floor version of the event. The newspaper article provides secondary data. The data are, of necessity, distorted—perhaps only a little, perhaps quite a bit—by the channels of communication through which they must pass to her. The reporter's writing skills, your sister's reading skills, and the inability of language to reproduce every nuance of detail that a firsthand observation can provide—all of these factors distort what you actually observed.

What we have been saying about data and their relation to Truth is represented by Figure 4.2. Lying farthest away from the researcher—and, hence, least accessible—is The Realm of Absolute Truth. It can be approached by the researcher only by passing through two intermediate areas that we have labeled The Realm of the Data. Notice that a barrier exists between The Realm of Absolute Truth and The Region of the Primary Data. Small bits of information leak through the barrier and manifest themselves as data. Notice, too, the foggy barrier between The Realm of the Data and The Realm of the Inquisitive Mind of the Researcher. This barrier is comprised of many things: the distortions and insensitivities of the human senses, the weaknesses of instrumentation, the inability of language to communicate thoughts exactly, the inability of two human beings to witness the same event and report it in precisely the same way, and so on.

Researchers must never forget the overall idea underlying Figure 4.2. Keeping it in mind can prevent them from making exaggerated claims or drawing unwarranted conclusions. No researcher can ever glimpse Absolute Truth, nor can anyone even perceive the data that reflect that Truth except through imperfect senses and imprecise channels of communication. Such a humiliating awareness helps researchers be cautious in the interpretation and reporting of research findings—for instance, by using such words and phrases as *perhaps*, *it seems*, *one might conclude*, *it would appear to be the case*, and *the data are consistent with the hypothesis that*.

FIGURE 4.2

The relation between data and Truth



Criteria for the Admissibility of Data

Not all data that come to the researcher's attention are acceptable for use in a research project. Data can be defective. If they are, they may affect the validity of the researcher's conclusions. The imperfections in the data stem from the imperfections and irregularities of nature. If researchers include in the mass of data those that are imperfect or irregular—those that are distorted, inaccurate, hopelessly entangled with irrelevant variables, and so on—they corrupt the entire body of data.

As previously mentioned, any research effort should be *replicable*; that is, it should be able to be repeated by any other researcher at any other time under precisely the same conditions. To ensure such precision, certain criteria must be adopted, certain limits established, and certain standards set up that all data must meet in order to be admitted for study.

For example, imagine that an agronomist wants to determine the effect of ultraviolet light on growing plants. *Ultraviolet* is a vague term: It encompasses a range of light waves that vary considerably in nanometers. The agronomist must narrow the parameters of the data so that they will fall within certain specified limits. What precisely does she or he mean by *ultraviolet light*? Within what nanometer range will ultraviolet emission be acceptable? At what intensity? For what length of time? At what distance from the growing plants? What precisely does the researcher mean by the phrase "effect of ultraviolet light on growing plants"? All plants? A specific genus? A particular species? The agronomist must be quite specific about all of these things so that another researcher can achieve very similar conditions in replicating the study.

By prescribing such criteria and insisting on standards, we can control the types of data admitted and regulate the conditions under which the research effort proceeds. Any data not meeting the criteria are excluded from the study. It is, of course, much easier to control data in the physical sciences, where we can measure a phenomenon fairly objectively, but such control is also possible in the humanities, the social sciences, and the arts. In these latter disciplines, the criteria are usually established by definition, and admissible data must meet the definitive parameters.

We hem in the data on all sides, placing on them the restrictions of criterion after criterion, so that we are able to isolate only those data acceptable for our use. The rest of the data are inadmissible. The restrictions we identify are the **criteria for the admissibility of data**.

When we standardize the data, admitting only those that comply with our criteria, we can better control the research effort and conclude with greater certainty what appears to be true. Therefore, to ensure the integrity of the research, we must set forth *beforehand* precisely what standards the data must meet. These criteria must be set forth clearly in both the research proposal and the final research report. Only in this way can the researcher and the consumer of the research come to a common, intelligent understanding of the investigation at hand.

Planning for Data Collection

Basic to any research project are four fundamental questions about the data. To avoid serious trouble later on, the researcher must answer these questions specifically and concretely. Clear answers can help bring any research planning and design into focus.

1. *What data are needed?* This question may seem like an overly simple one, but in fact a specific, definitive answer is fundamental to any research effort. On a sheet of paper, write the answers to the following questions: To resolve the problem, what data are mandatory? What is their nature? Are they historical documents? Interview excerpts? Questionnaire responses? Observations? Measurements made before and after an experimental intervention? Specifically, what data do you need, and what are their characteristics?
2. *Where are the data located?* Those of us who have taught courses in research methodology are constantly awed by the fascinating problems that students identify for research projects. But then we ask a basic question: "Where will you get the data to resolve the problem?" Some students either look bewildered and remain speechless or else mutter something such as, "Well, they must be available *somewhere*." Not *somewhere*, but *precisely where*? If you are doing a study of documents, where are the documents you need? At exactly which library and in what collection will you find them? What society or what organization has the files you must examine? Where are these organizations located? Specify geographically—by town, street address, and postal code! Suppose a nurse or a nutritionist is doing a research study about Walter Olin Atwater, whose work has been instrumental in establishing the science of human nutrition in the United States. Where are the data on Atwater located? The researcher can go no further until that basic question is answered.
3. *How will the data be obtained?* To know where the data are located is not enough; you need to know how you might acquire them. With privacy laws, confidentiality agreements, and so on, obtaining the information you need might not be as easy as you think. You may indeed know what data you need and where you can find them, but an equally important question is, How will you get them? Careful attention to this question marks the difference between a viable research project and a pipe dream.
4. *How will the data be interpreted?* This is perhaps the most important question of all. The three former hurdles have been overcome. You have the data in hand. But you must also spell out precisely what you intend to do with the data to solve the research problem or one of its subproblems.

Now go back and look carefully at how you have worded your research problem. Will you be able to get data that might adequately provide a solution to the problem? And if so, might they reasonably lend themselves to interpretations that shed light on the problem? If the answer to either of these questions is *no*, you must obviously rethink the nature of your problem. If, instead, both answers are *yes*, a next important step is to consider how you might best *measure* what you observe.

Identifying Good Measurement Instruments

As you zero in on the nature of the data necessary for solving your research problem, you will probably discover that you must pin down your observations by measuring them in some way. In some cases you will be able to use one or more existing instruments—perhaps an oscilloscope to measure patterns of sound, a published personality test to measure a person's tendency to be either shy or outgoing, or a rating scale that a previous researcher has developed to assess parents' childrearing practices. In other situations you may have to develop your *own* measurement instruments—perhaps a survey to assess people's opinions about welfare reform, a paper-and-pencil test to measure what students have learned from a particular instructional unit, or a checklist to evaluate the quality of a new product.

Measurement instruments provide a basis on which the entire research effort rests. Just as a building with a questionable foundation is unlikely to be safe for habitation, so, too, will a research effort employing faulty measurement tools provide little of value in solving the problem under investigation.

Defining Measurement

What exactly *is* measurement? Typically we think of measurement in terms of such objects as rulers, scales, gauges, and thermometers. In research, **measurement** takes on a somewhat different meaning:

Measurement is limiting the data of any phenomenon—substantial or insubstantial—so that those data may be interpreted and, ultimately, compared to a particular qualitative or quantitative standard.

Let's look more closely at this definition. The first five words are *measurement is limiting the data*. When we measure something, we set a limit that constrains the data in some way. We erect a barrier beyond which those data cannot go. What is a foot, a mile, a pound? Each is a unit of measure governed by a numerical constraint: 12 inches constrain a foot; 5,280 feet, a mile; and 16 ounces, a pound.

Now let's look at the next six words: *of any phenomenon—substantial or insubstantial*. This phrase is all-inclusive. Nothing exists that the researcher cannot measure in some way. In some cases, observable physical entities are measured. These are **substantial** measurements; that is, the things being measured have physical substance, an obvious basis in the physical world. An astronomer measures patterns and luminosity of light in the night sky; a neurologist measures intensity and location of activity in the brain; a chemist measures the mass of a compound both before and after transforming it in some way. All of these are attempts to measure substantial phenomena. Some devices designed to measure substantial phenomena, such as high-powered telescopes and MRI machines, are highly specialized and used only in particular disciplines. Others, such as balance scales and tape measures, are applicable to many fields of inquiry.

We can also measure those things—if “things” they be—that are **insubstantial**, that exist only as concepts, ideas, opinions, feelings, or other intangible entities. For example, we might attempt to measure the economic “health” of business, the degree to which students have

"learned," or the extent to which people "value" physical exercise. We seek to measure these intangibles, not with tape measures or scales, but with the Dow-Jones index, achievement tests, questionnaires, or interviews.²

We continue with the next seven words of our definition of measurement: *so that those data may be interpreted*. We cannot emphasize this point enough: Research involves not only the collection but also the interpretation of data—the transformation of data into new discoveries, revelations, and enlightenments.

Now we finish our definition: *and, ultimately, compared to a particular qualitative or quantitative standard*. A researcher must have a goalpost, a true north, a point of orientation. In research, we call these standards *norms, averages, conformity to expected statistical distributions, goodness of fit, accuracy of description*, and the like.

Measurement is ultimately a comparison: a thing or concept measured against a point of limitation. We compare the length of an object with the scale of a ruler or a measuring tape. We "measure" an ideology against the meaning of it as articulated by its originator. For example, the essence of a philosophy arises from the writings and teachings of its founder: Platonism from Plato, Marxism from Karl Marx, and romanticism, perhaps, from Jean Jacques Rousseau. The essence of a religious belief lies in its sacred writings, in the precepts of its great teachers, and in its creed. The meaning of freedom is articulated in many political documents—for instance, in the Declaration of Independence and the Constitution of the United States. Against these original sources, it is possible to measure the thoughts and ideas of others and to approximate their similarity to or deviance from those sources.

Data examined statistically are constantly being interpreted in comparison with statistical norms: the normal curve, probability tables, and other accepted statistical standards. Data analyzed in a nonnumerical, qualitative manner are compared across data sources, across methods, and across time.

As you can see, then, our definition of measurement implies much more than an everyday understanding of measurement might suggest. Measurement provides an important tool with which data may be inspected, analyzed, and interpreted so that the researcher may probe the meaning that lies below the surface.

Measuring Insubstantial Phenomena: An Example

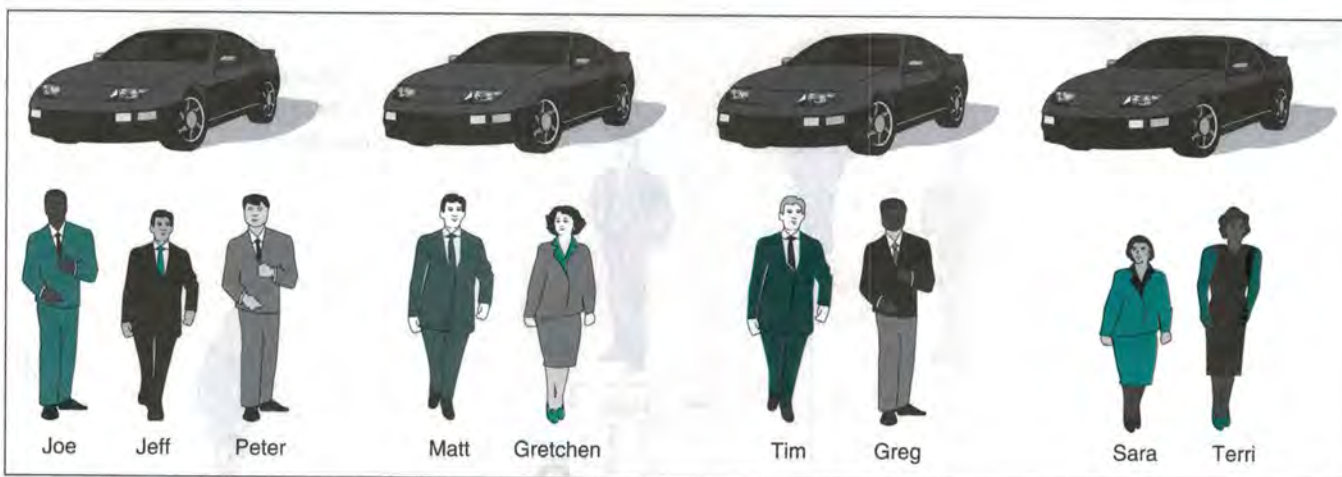
Measuring insubstantial phenomena—those phenomena that have no obvious, concrete basis in the physical world—can sometimes involve considerable creativity on the part of the researcher. For example, imagine that we want to examine—and hence also to *measure*—the interpersonal dynamics among a small group of people. Let's take a group of nine people, shown in Figure 4.3, who work together in the human resources department of a large corporation. They are to attend a recognition dinner at an exclusive hotel.

They arrive in four cars. They enter the hotel in the following order: Terri, Sara, Greg, Tim, Gretchen, Matt, Peter, Jeff, and Joe. They greet one another and have time for a brief conversation before dinner. Most of them position themselves in conversation groups, as shown in Figure 4.4.

To the perceptive observer, the interpersonal dynamics within the group soon become apparent. Who greets whom with enthusiasm or with indifference? Who joins in conversation with whom? Who seems to be a relative outsider? However, *to merely observe the behavior of individuals in a particular situation is not to measure it*.

One possible approach to measuring the group's interpersonal dynamics is to give each person in the group a slip of paper on which to record three choices: (a) one or more individuals in

²You may sometimes see the substantial–insubstantial distinction referred to as *manifest variables* (which can be directly observed and measured) versus *latent variables* (which lie below the surface and can be measured only indirectly through their effects on another, observable entity; e.g., see Bartholomew, 2004).

**FIGURE 4.3**

Recognition dinner participants

the group whom the person likes most, (b) one or more individuals whom the person likes least, and (c) one or more individuals for whom the person has no strong feeling one way or the other. When using this method, we should poll each person in the group individually and guarantee that every response will be kept confidential.

We can then draw a chart, or **sociogram**, of these interpersonal reactions, perhaps in the manner depicted in Figure 4.5. We might also assign “weights” that place the data into three numerical categories: +1 for a positive choice, 0 for indifference, and -1 for a negative reaction. Categorizing the data in this way, we can then construct a sociometric matrix. To create a matrix, we arrange the names of each person twice: vertically down the left side of a grid and horizontally across the top of the grid. The result is shown in Table 4.1. The dashes in the grid reflect the fact that the people can choose other individuals but cannot choose *themselves*.

Certain relationships begin to emerge. As we represent group dynamics in multiple forms, certain clusters of facts suggest the following conclusions:

- Jeff is the informal or popular leader (sometimes called the “star”) of the group. He received five choices and only one rejection (see the “Jeff” column in Table 4.1). The sociogram confirms Jeff’s popularity with his colleagues.
- Probably some factions and possible tension are present in this group. Notice that Peter, Sara, and Terri form a subclique, or “island,” that is separated from the larger clique that Jeff leads. The apparent liaison between these two groups is Joe, who has mutual choices with both Jeff and Peter.
- Friendship pairs may lend cohesion to the group. Notice the mutual choices: Matt and Gretchen, Gretchen and Jeff, Jeff and Joe, Joe and Peter, Peter and Terri, Terri and Sara. The sociogram clearly reveals these alliances.

FIGURE 4.4

Conversation groups

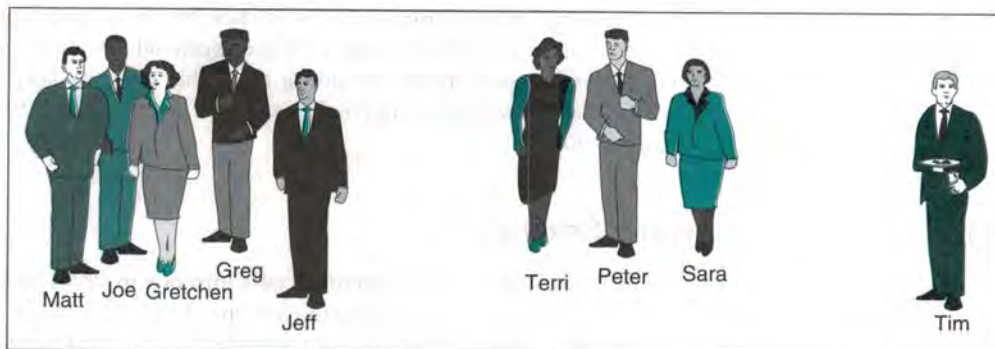
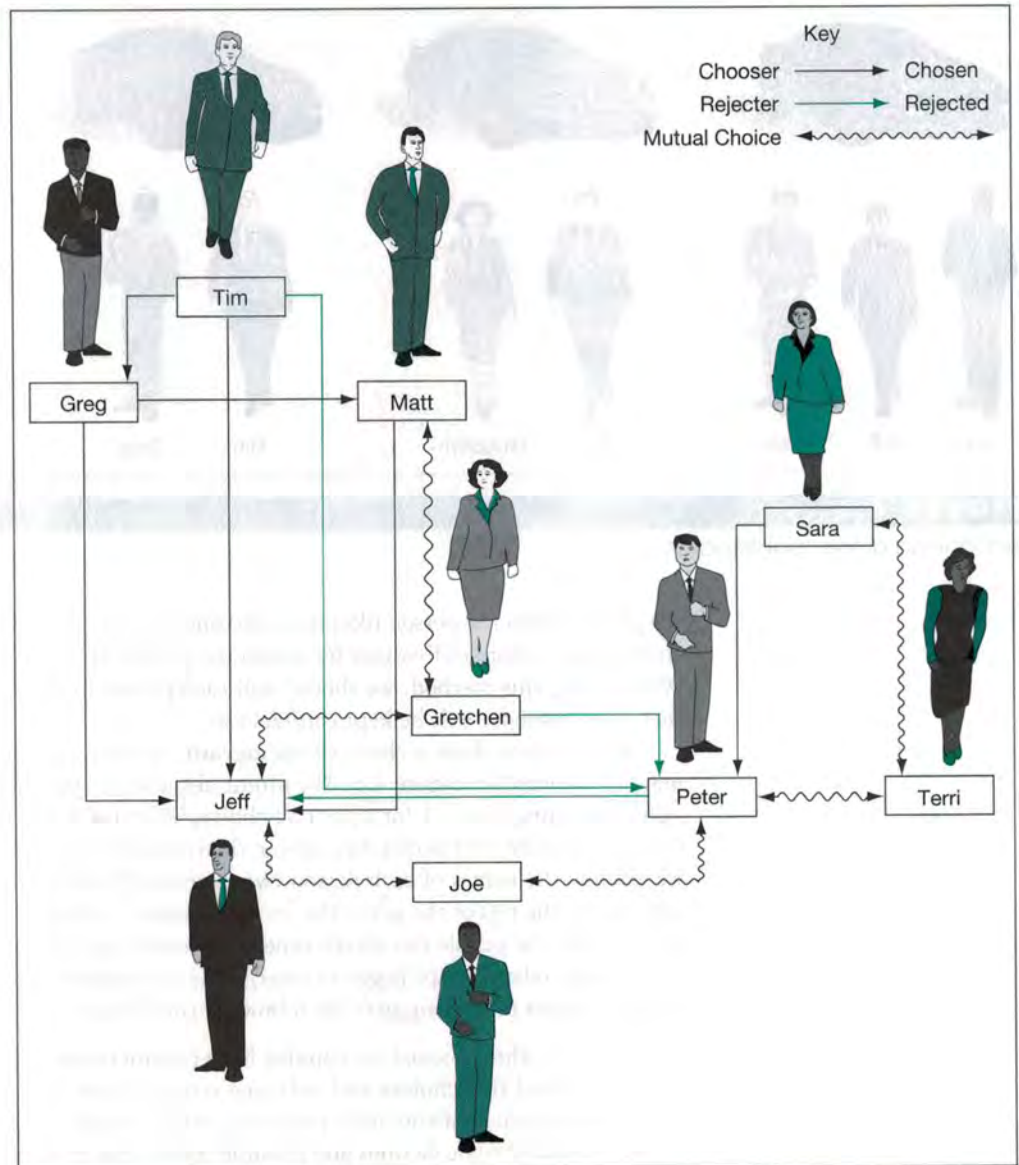


FIGURE 4.5

Sociogram of interpersonal dynamics



- Tim apparently is the isolate of the group. He received no choices; he is neither liked nor disliked. In such a position, he is probably the least influential member of the group.

With this example we have illustrated what it means to interpret data by measuring an insubstantial phenomenon and analyzing the resulting data. Notice that we did not just observe the behaviors of nine individuals at a social occasion, but we also looked below the surface to discover hidden social forces at play in their behaviors with one another. Our example is a simple one, to be sure. Measurement of interpersonal dynamics and social networks can certainly take more complex forms, including some that are especially helpful in studying social forces within large, extended groups (e.g., Chatterjee & Srivastava, 1982; Freeman, 2004; Wasserman & Faust, 1994).

Types of Measurement Scales

Virtually any form of measurement falls into one of four categories, or **scales**: (a) nominal, (b) ordinal, (c) interval, and (d) ratio (Stevens, 1946). The scale of measurement will ultimately dictate the statistical procedures (if any) that can be used in processing the data.

TABLE 4.1 Data from Figure 4.5 presented as a sociometric matrix

How Each Person Was Rated by the Others										
How Each Person Rated the Others		Gretchen	Joe	Greg	Sara	Peter	Jeff	Tim	Matt	Terri
	Gretchen	—	0	0	0	-1	+1	0	+1	0
	Joe	0	—	0	0	+1	+1	0	0	0
	Greg	0	0	—	0	0	+1	0	+1	0
	Sara	0	0	0	—	+1	0	0	0	+1
	Peter	0	+1	0	0	—	-1	0	0	+1
	Jeff	+1	+1	0	0	0	—	0	0	0
	Tim	0	0	+1	0	-1	+1	—	0	0
	Matt	+1	0	0	0	0	+1	0	—	0
	Terri	0	0	0	+1	+1	0	0	0	—
Totals	2	2	1	1	1	4	0	2	2	

Nominal Scales

The word *nominal* comes from the Latin *nomen*, meaning “name.” Hence we can “measure” data to some degree by assigning names to them. Recall that the definition of measurement presented earlier includes the phrase *limiting the data*. That is what a **nominal scale** does—it limits the data—and just about all that it does. Assign a specific name to anything, and you have restricted that thing to the meaning of its name. For example, we can measure a group of children by dividing it into two groups: girls and boys. Each subgroup is thereby measured—restricted—by virtue of gender as belonging to a particular category.

Things can be measured nominally in an infinite number of ways. We can further measure girls and boys according to the home site of each child. Imagine that the town in which the children live is divided into two sections by Main Street, which runs from east to west. Those children who live north of Main Street are “the Northerners”; those who live south of it are “the Southerners.” In one period of U.S. history, we measured the population of the entire nation in just such a manner.

Nominal measurement is quite simplistic, but it does divide data into discrete categories that can be compared with one another. Let’s take an example. Imagine that we have six children: Zahra, Paul, Kathy, Binh, Ginger, and Nicky. They can be divided into six units of one child each. They can also form two groups: Zahra, Kathy, and Ginger (the girls) in one group and Paul, Binh, and Nicky (the boys) in the other. Perhaps all six children are students in a class that meets in Room 12 at Thompson’s Corner School. By assigning a room number, we have provided the class with a name, even though that “name” is a number. In this case, the number has no quantitative meaning: Room 12 is not necessarily bigger or better than Room 11, nor is it inferior to Room 13.

Only a few statistical procedures are appropriate for analyzing nominal data. We can use the *mode* as an indicator of the most frequently occurring category within our data set; for example, we might determine that there are more boys than girls in Room 12 at Thompson’s Corner School. We can find the *percentage* of people in various subgroups within the total group; for example, we could calculate the percentage of boys in each classroom. We can use a *chi-square test* to compare the relative frequencies of people in various categories; for example, we might discover that more boys than girls live north of Main Street but that more girls than boys live south of Main Street. (We discuss these statistics, as well as the statistics listed in the following discussions of the other three scales, in Chapter 11.)

Ordinal Scales

With an **ordinal scale** of measurement, we can think in terms of the symbols $>$ (greater than) and $<$ (less than). We can compare various pieces of data in terms of one being greater or higher than another. In essence, this scale allows us to *rank-order* our data—hence its name *ordinal*.

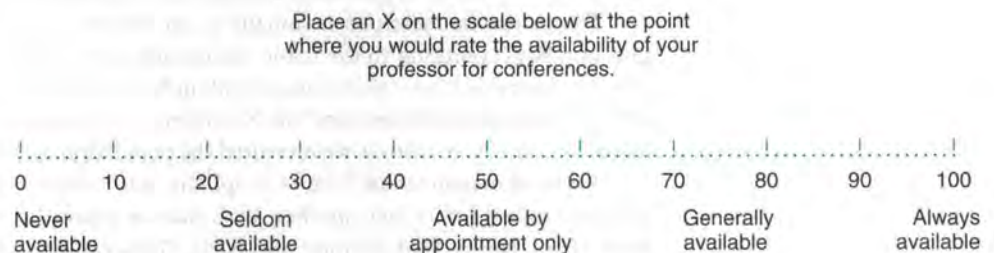
As an example, we can roughly measure level of education on an ordinal scale by classifying people as being unschooled or having completed an elementary, high school, college, or graduate education. Likewise, we can measure members of the workforce by grades of proficiency: unskilled, semiskilled, or skilled.

An ordinal scale expands the range of statistical techniques we can apply to our data. In addition to the statistics we can use with nominal data, we can also determine the *median*, or halfway point, in a set of data. We can use a *percentile rank* to identify the relative position of any item or individual in a group. We can determine the extent of the relationship between two characteristics by means of Spearman's *rank order correlation*.

Interval Scales

An **interval scale** of measurement is characterized by two features: (a) It has equal units of measurement and (b) its zero point has been established arbitrarily. The Fahrenheit (F) and Celsius (C) scales for measuring temperature are examples of interval scales: The intervals between any two successive numbers of degrees reflect equal changes in temperature, but the zero point is not equivalent to a total absence of heat. For instance, when Gabriel Fahrenheit was developing his Fahrenheit scale, he first took as his zero point the coldest temperature he observed in Iceland. Later, he made it the lowest temperature obtainable with a mixture of salt and ice. This was purely an arbitrary decision. It placed the freezing point of water at 32° and the boiling point at 212° above zero.

The rating scales employed by many survey groups, businesses, and professional organizations are often assumed to be on interval scales. For instance, many universities ask students to use rating scales to evaluate the teaching effectiveness of various professors. Following is an example of an item from one university's teaching evaluation form:



Notice that the scale has 11 equidistant points ranging from 0 to 100. The equidistance creates what is presumed to be an interval scale for the measure. At five points along the scale are descriptive labels that can help students determine how they should rate their professor's availability. We might place descriptors at more places along the scale (perhaps at 10-point distances), thus potentially making the scale more sensitive or more accurate. For indicating the availability of a professor, such fineness of discrimination may not be either possible or desirable, but one can conceive of situations in which such a degree of discrimination may be necessary and appropriate.

Interval scales of measurement allow statistical analyses that are not possible with nominal or ordinal data. Because an interval scale reflects equal distances among adjacent points, any statistics that are calculated using addition or subtraction—for instance, *means*, *standard deviations*, and *Pearson product moment correlations*—can now be used.

Ratio Scales

Two commonly used measurement instruments—a thermometer and a yardstick—might help you understand the difference between the interval and ratio scales. If we have a thermometer that measures temperature on the Fahrenheit scale, we *cannot* say that 80°F is twice as warm as

40°F. Why? Because this scale does not originate from a point of absolute zero; a substance may have some degree of heat even though its measured temperature falls *below* zero. With a yardstick, however, the beginning of linear measurement is absolutely the beginning. If we measure a desk from the left edge to the right edge, that's it. There is no more desk in either direction beyond those limits. A measurement of "zero" means there is no desk at all, and a "minus" desk width isn't even possible.

More generally, a **ratio scale** has two characteristics: (a) equal measurement units (similar to an interval scale) and (b) *an absolute zero point*, such that 0 on the scale reflects a total absence of the entity being measured.

Let's consider once again the "availability" scale presented earlier for measuring professor effectiveness. This scale could never be considered a ratio scale. Why? Because there is only one condition in which the professor would be absolutely unavailable—if the professor were dead!—in which case we would not be asking students to evaluate this individual.

What distinguishes the ratio scale from the other three scales is that *the ratio scale can express values in terms of multiples and fractional parts*, and the ratios are *true ratios*. A yardstick can do that: A yard is a *multiple* (by 36) of a 1-inch distance; an inch is one-twelfth (a *fractional part*) of a foot. The ratios are 36:1 and 1:12, respectively.

Ratio scales outside the physical sciences are relatively rare. And whenever we cannot measure a phenomenon in terms of a ratio scale, we must refrain from making comparisons such as "this thing is three times as great as that" or "we have only half as much of one thing as another." Only ratio scales allow us to make comparisons that involve multiplication or division.

We can summarize our description of the four scales this way:

If you can say that

- One object is different from another, you have a *nominal scale*;
- One object is bigger or better or more of anything than another, you have an *ordinal scale*;
- One object is so many units (degrees, inches) more than another, you have an *interval scale*;
- One object is so many times as big or bright or tall or heavy as another, you have a *ratio scale*.

(Senders, 1958, p. 51)

Table 4.2 provides a quick reference for the various types of scales, their distinguishing characteristics, and the statistical analysis possibilities for each scale. When we consider the statistical interpretation of data in later chapters (especially in Chapter 11), you may want to refer to this table to determine whether the type of measurement instrument you have used will support the statistical operation you are contemplating.

TABLE 4.2 A summary of measurement scales, their characteristics, and their statistical implications

	Measurement Scale	Characteristics of the Scale	Statistical Possibilities of the Scale
Non-Interval Scales	Nominal scale	A scale that "measures" only in terms of names or designations of discrete units or categories	Enables one to determine the mode, percentage values, or chi-square
	Ordinal scale	A scale that measures in terms of such values as "more" or "less," "larger" or "smaller," but without specifying the size of the intervals	Enables one also to determine the median, percentile rank, and rank correlation
Interval Scales	Interval scale	A scale that measures in terms of equal intervals or degrees of difference, but with an arbitrarily established zero point that does not represent "nothing" of something	Enables one also to determine the mean, standard deviation, and product moment correlation; allows one to conduct most inferential statistical analyses
	Ratio scale	A scale that measures in terms of equal intervals and an absolute zero point	Enables one also to determine the geometric mean and the percentage variation; allows one to conduct virtually any inferential statistical analysis

CONCEPTUAL ANALYSIS EXERCISE Identifying Scales of Measurement

Each of the following scenarios involves measuring one or more variables. Decide whether the various measurements reflect nominal, ordinal, interval, or ratio scales, and justify your choices. Be careful, as the answers are not always as obvious as they might initially appear. The answers are provided after the “For Further Reading” list at the end of the chapter.

1. An environmental scientist collects water samples from streams and rivers near large industrial plants and saves exactly one liter of water from each sample. Then, back at the lab, the researcher determines the amounts of certain health-jeopardizing bacteria in each sample. What measurement scale does the measurement of bacteria content reflect?
2. A tourism researcher is studying the relationship between (a) a country's average annual temperature and (b) the amount of tourist dollars that the country brings in every year. What scales underlie the two variables in this study?
3. A political science researcher wants to determine whether people's political party membership is correlated with the frequency with which they have voted in local elections in the past five years. The researcher can easily obtain information about people's party membership and voting records from the town clerks in several communities. To simplify data collection, the researcher uses the following coding scheme for party membership: 1 = Registered as Democrat, 2 = Registered as Republican, 3 = Registered as member of another party, 0 = No declared party affiliation. What measurement scale(s) underlie (a) political party membership and (b) voting frequency?
4. A marketing researcher in the United States wants to determine whether a certain product is more widely used in some parts of the country than others. The researcher separates the country into ten regions based on zip code; zip codes below 10000 are northeastern states, zip codes of 90000 and above are western states, and so on. What measurement scale does the researcher's coding scheme for the regions represent?
5. An economist is studying the home-buying behaviors of people of different income levels. The researcher puts people into four categories: Group A includes those earning up to \$20,000 per year, Group B includes those earning between \$20,001 and \$50,000 per year, Group C includes those earning between \$50,001 and \$100,000 per year, and Group D includes those earning more than \$100,000 per year. In this study, what kind of scale is income level?
6. A geographer is studying traffic patterns on four different types of roads that vary in quality: superhighways (i.e., roads accessible only by relatively infrequent on-off ramps), highways (i.e., roads that allow relatively high speeds for long distances but may have an occasional traffic light), secondary roads (i.e., well-paved two-lane roads), and tertiary roads (narrow, infrequently traveled roads; some may consist only of gravel). The type of road in this study reflects which type of measurement scale?
7. A psychologist is developing an instrument designed to measure college students' test anxiety. The instrument includes 25 statements—for example, “My heart starts to pound when I see the word *test* on a course syllabus” and “My palms get sweaty while I'm taking a multiple-choice test.” Students must rate each of these statements on a 5-point scale, as follows:
 - 0 This is never true for me.
 - 1 This is rarely true for me.
 - 2 This is sometimes true for me.
 - 3 This is often true for me.
 - 4 This is always true for me.

Students who answer “never” to each of the 25 questions get the lowest possible score of 0 on the instrument. Students who answer “always” to each of the 25 questions get the highest possible score of 100 on the instrument. Thus, scores on the instrument range from 0 to 100. What kind of scale do the scores represent?

Validity and Reliability in Measurement

Regardless of the type of scale a measurement instrument involves, the instrument must have both validity and reliability for its purpose. The validity and reliability of measurement instruments influence the extent to which a researcher can learn something about the phenomenon under investigation, the probability that the researcher will obtain statistical significance in any data analysis, and the extent to which the researcher can draw meaningful conclusions from the data.

In your research proposal and final research report, you should provide evidence that the instruments you use have a reasonable degree of validity and reliability for your purposes. But validity and reliability take different forms, depending on the nature of the research problem, the methodology being used to address the problem, and the nature of the data that are collected. In the following sections, we briefly discuss the kinds of validity and reliability you may need to consider and how you might determine and enhance them.

Validity of Measurement Instruments

The **validity** of a measurement instrument is the extent to which the instrument measures what it is intended to measure. Certainly no one would question the premise that a yardstick is a valid means of measuring length. Nor would most people doubt that a thermometer measures temperature; for instance, in a mercury thermometer, the level to which the mercury rises is a function of how much it expands, which is a function of the degree to which it is hot or cold.

But to what extent does an intelligence test actually measure a person's intelligence? How accurately do people's annual incomes reflect their social class? And how well does a sociogram capture the interpersonal dynamics in a group of nine people? Especially when we are measuring *insubstantial* phenomena—phenomena without a direct basis in the physical world—our measurement instruments may be somewhat suspect in terms of validity.

Let's return to the rating-scale item presented earlier to assess a professor's availability to students (see p. 86) and consider its validity as such a measure. Notice how fuzzy some of the labels are. The professor is "always available." What does *always* mean? Twenty-four hours a day? Could you call the professor at 3:00 a.m. any day of the week, or only whenever the professor is on campus? If the latter is the case, could you call your professor out of a faculty meeting or out of a conference with the college president? We might have similar problems in interpreting "generally available," "seldom available," and "never available." On careful inspection, what seems at first glance to be a scale that anyone could understand has limitations *as a measurement instrument* for research purposes.

A paper-and-pencil test may be intended to measure a certain characteristic, and it may be *called* a measure of that characteristic, but these facts don't necessarily mean that the test actually measures what its creator says it does. For example, consider a paper-and-pencil test of personality traits in which, with a series of check marks, a person indicates his or her most representative characteristics or behaviors in given situations. The person's responses on the test are presumed to reveal relatively stable personality traits. The question that validity asks is: Does such a test, in fact, measure the person's personality traits, or does it measure something else altogether? The answer depends, at least in part, on the extent to which the person is or *can* be truthful in responding. If the person responds in terms of characteristics and behaviors that he or she believes to be socially desirable, the test results may reveal not the person's actual personality, but rather an idealized portrait of how he or she would like to be perceived by others.

The validity of a measurement instrument can take several different forms, each of which is important in different situations:

- **Face validity** is the extent to which, on the surface, an instrument *looks like* it is measuring a particular characteristic. Face validity is often useful for ensuring the cooperation of people who are participating in a research study. But because it relies entirely on subjective judgment, it is not, in and of itself, a terribly dependable indicator that an instrument is truly measuring what the researcher wants to measure.
- **Content validity** is the extent to which a measurement instrument is a representative sample of the content area (domain) being measured. Content validity is often a consideration when

a researcher wants to assess people's *achievement* in some area—for instance, the knowledge students have learned during classroom instruction or the job skills workers have acquired in a training program. A measurement instrument has high content validity if its items or questions reflect the various parts of the content domain in appropriate proportions and if it requires the particular behaviors and skills that are central to that domain.

- **Criterion validity** is the extent to which the results of an assessment instrument correlate with another, presumably related measure (the latter measure is, in this case, the *criterion*). For example, a personality test designed to assess a person's shyness or outgoingness has criterion validity if its scores correlate with other measures of a person's general sociability. An instrument designed to measure a salesperson's effectiveness on the job should correlate with the number of sales the individual actually makes during the course of a business week.
- **Construct validity** is the extent to which an instrument measures a characteristic that cannot be directly observed but is assumed to exist based on patterns in people's behavior (such a characteristic is a *construct*). Motivation, creativity, racial prejudice, love—all of these are constructs, in that none of them can be *directly* observed and measured. When researchers ask questions, present tasks, or observe behaviors as a way of assessing an underlying construct, they should obtain some kind of evidence that their approach does, in fact, measure the construct in question.

Sometimes there is universal agreement that a particular instrument provides a valid instrument for measuring a particular characteristic; such is the case for yardsticks, thermometers, barometers, oscilloscopes, and so on. But whenever we do *not* have such widespread agreement, we must provide evidence that an instrument we are using has validity for our purpose.

It is critical to note that the validity of any measurement instrument can vary considerably depending on the purpose for which it is being used. In other words, *the validity of an instrument is specific to the situation*. For example, a tape measure wrapped horizontally around a person's head is a valid measure of the person's head circumference but *not* a valid measure of the person's intelligence. Likewise, a widely used intelligence test might provide a reasonable estimation of children's general cognitive development but is *not* suitable for determining how well the children can perform in, say, a geometry class or interpersonal conflict.

Determining the validity of a measurement instrument. An in-depth discussion of how to determine validity is beyond the scope of this book; measurement textbooks, such as those listed in the "For Further Reading" section at the end of the chapter, provide more detailed information. But here we offer three examples of what researchers sometimes do to demonstrate that their measurement instruments have validity for their purposes:

- **Table of specifications.** To construct a measurement instrument that provides a representative sample of a particular content domain—in other words, to establish content validity—the researcher often constructs a two-dimensional grid (*table of specifications*) listing the specific topics and behaviors that reflect achievement in the domain. In each cell of the grid, the researcher indicates the relative importance of each topic–behavior combination. He or she then develops a series of tasks or test items that reflects the various topics and behaviors in appropriate proportions.
- **Multitrait–multimethod approach.** Two or more different characteristics are each measured using two or more different approaches (Campbell & Fiske, 1959). The different measures of the same characteristic should be highly correlated. The same ways of measuring different characteristics should *not* be highly correlated. For example, in a classroom situation, the constructs *achievement motivation* and *social motivation* might each be measured by both self-report questionnaires and teacher observation checklists. Statistical analyses should reveal that the two measures of achievement motivation are highly correlated and that the two measures of social motivation are also highly correlated. Results from the two self-report questionnaires—because they are intended to assess different and presumably unrelated characteristics—should *not* be highly correlated, nor should results from the two teacher checklists.

- **Judgment by a panel of experts.** Several experts in a particular area are asked to scrutinize an instrument and give an informed opinion about its validity for measuring the characteristic in question.

Although none of the approaches just described guarantees the validity of a measurement instrument, each one increases the likelihood of such validity.

Reliability of Measurement Instruments

Imagine that you are concerned about your growing waistline and decide to go on a diet. Every day you put a tape measure around your waist and pull the two ends together snugly to get a measurement. But just how tight is “snug”? Quite possibly, the level of snugness might be different from one day to the next. In fact, you might even measure your waist with different degrees of snugness from one *minute* to the next. To the extent that you are not measuring your waist in a consistent fashion—even though you always use the same tape measure—you have a problem with reliability.

More generally, **reliability** is the consistency with which a measuring instrument yields a certain, consistent result when the entity being measured hasn’t changed. As we have just seen in our waist-measuring situation, instruments that measure physical phenomena aren’t necessarily completely reliable. As another example, think of a measuring cup that a baker might use while making a cake. When measuring a half-cup of flour, the baker won’t always measure *exactly* the same amount of flour each time.

Instruments designed to measure social and psychological characteristics (insubstantial phenomena) tend to be even less reliable than those designed to measure physical (substantial) phenomena. For example, a student using the rating-scale item presented earlier for measuring professor availability might easily rate the professor as “70” one day and “90” the next, not because the professor’s availability has changed overnight but because the student’s interpretations of the phrases “generally available” and “always available” *have* changed. Similarly, if we asked the nine people portrayed in Figure 4.3 (Gretchen, Joe, Greg, etc.) to indicate the people they liked best and least among their colleagues, they wouldn’t necessarily always give us the same answers they gave us previously, even if the interpersonal dynamics within the group have remained constant.

Determining the reliability of a measurement instrument. Like validity, reliability takes different forms in different situations. But in the case of reliability, its particular form is essentially equivalent to the procedure used to determine it. Following are four forms of reliability that are frequently of interest in research studies:

- **Interrater reliability** is the extent to which two or more individuals evaluating the same product or performance give identical judgments.
- **Test-retest reliability** is the extent to which a single instrument yields the same results for the same people on two different occasions.
- **Equivalent forms reliability** is the extent to which two different versions of the same instrument (e.g., “Form A” and “Form B” of a scholastic aptitude test) yield similar results.
- **Internal consistency reliability** is the extent to which all of the items within a single instrument yield similar results.

For each of these forms, determining reliability involves two steps:

1. Getting two measures for each individual in a reasonably large group of individuals—in particular by doing one of the following:
 - a. Having two different raters evaluate the same performance for each individual (interrater reliability)
 - b. Administering the same instrument to the individuals at two different points in time—perhaps a day, a week, or a month apart (test-retest reliability)

- c. Giving each individual two parallel versions of the same instrument (equivalent forms reliability)
 - d. Administering only one instrument but calculating two subscores for the instrument—for instance, calculating one score for odd-numbered items and another score for even-numbered items (internal consistency reliability)
2. Calculating a correlation coefficient that expresses the degree to which the two measures are similar (see Chapter 11 for a discussion of correlation coefficients)

You can find more in-depth discussions about determining reliability in almost any general measurement textbook.

Enhancing the Reliability and Validity of a Measurement Instrument

Both validity and reliability reflect the degree to which we may have *error* in our measurements. In many instances—and especially when we are measuring insubstantial phenomena—a measurement instrument may allow us to measure a characteristic only indirectly and so may be subject to a variety of biasing factors (e.g., people's responses on a rating scale are apt to be influenced by their interpretations, prejudices, memory lapses, etc.). In such cases, we have error due to the imperfect *validity* of the measurement instrument. Yet typically—even when we are measuring substantial phenomena—we may get slightly different measures from one time to the next simply because our measurement tool is imprecise (e.g., the waist or head size we measure may depend on how snugly we pull the tape measure). In such cases, we have error due to the imperfect *reliability* of the measure. Generally speaking, validity errors reflect biases in the instrument itself and are relatively constant sources of error. In contrast, reliability errors reflect *use* of the instrument and are apt to vary unpredictably from one occasion to the next.

We can measure something accurately only when we can also measure it consistently. Hence, by increasing the reliability of a measurement instrument, we might also increase its validity. A researcher can enhance the reliability of a measurement instrument in several ways. First, the instrument should always be administered in a consistent fashion. In other words, there should be **standardization** in use of the instrument from one situation or individual to the next. Second, to the extent that subjective judgments are required, specific *criteria* should be established that dictate the kinds of judgments the researcher makes. And third, any research assistants who are using the instrument should be *well trained* so that they obtain similar results for any single individual or situation being measured.

Yet even if we enhance the reliability of our measurements, we don't necessarily increase their accuracy. In other words, *reliability is a necessary but insufficient condition for validity*. For example, we could use a tape measure to measure a person's head circumference and claim that the result is a good reflection of intelligence. In this situation, we might have reasonable reliability—we are apt to get similar measures of an individual's head circumference on different occasions—but absolutely no validity. Head size is *not* a good indication of intelligence level.

Creative researchers use a wide variety of strategies to enhance the validity of their measurement instruments. One important strategy is to consult the literature in search of measurement techniques that other researchers have effectively used. Another is to show a first draft of an instrument to experienced colleagues and ask for their feedback and suggestions. Still another strategy is to conduct one or more *pilot studies* specifically to try out a particular instrument, carefully scrutinizing it for obvious or possible weaknesses, and then modify it in minor or major ways.

We cannot overemphasize the importance of determining and maximizing the validity and reliability of your measurement instruments. Without reasonably valid and reliable measures of the characteristics and phenomena under investigation, you cannot possibly obtain informative and useful data for addressing and solving your research problem.

As you plan your research project, you should clearly identify the nature of the measurement instruments you will use and carefully scrutinize them with respect to their potential validity and reliability. Furthermore, in your research proposal and final research report, you

should describe any instrument in explicit, concrete terms. For example, if you are using a particular piece of equipment to measure a certain physical characteristic or phenomenon, you should describe the equipment's specific nature (e.g., its manufacturer, model number, and level of precision). And if you are assessing some aspect of human thought or behavior, you should describe the questions asked or tasks administered, the overall length of the instrument (e.g., number of items, time required for administration), and the method of scoring responses.

CONCEPTUAL ANALYSIS EXERCISE Identifying Problems with Validity and Reliability in Measurement

In each of the scenarios in this exercise, a researcher encounters a measurement problem. Some of the scenarios reflect a problem with the validity of a measure. Others reflect a problem with a measure's reliability—a problem that indirectly also affects the measure's validity. For each scenario, choose the most obvious problem from among the following alternatives:

- | | |
|----------------------|------------------------------------|
| ■ Face validity | ■ Interrater reliability |
| ■ Content validity | ■ Test-retest reliability |
| ■ Criterion validity | ■ Equivalent forms reliability |
| ■ Construct validity | ■ Internal consistency reliability |

The answers appear after the "For Further Reading" list at the end of this chapter.

1. After using two different methods for teaching basic tennis skills to non-tennis-playing adults, a researcher assesses the effectiveness of the two methods by administering a true-false test regarding the rules of the game (e.g., faults and double-faults, scoring procedures).
2. A researcher writes 120 multiple-choice questions to assess middle school students' general knowledge of basic world geography (e.g., what the equator is, where Africa is located). To minimize the likelihood that students will cheat on the test by copying one another's answers, the researcher divides the questions into three different sets to create three 40-item tests. In collecting data, the researcher distributes the three tests randomly to students in any single classroom. After administering the tests to students at many different middle schools, the researcher computes the students' test scores and discovers that students who answered one particular set of 40 questions scored an average of 3 points higher than students who answered either of the other two 40-question sets.
3. In order to determine what kinds of situations provoke aggression in gorillas, two researchers observe mountain gorillas in the Virunga Mountains of northwestern Rwanda. As they watch a particular gorilla family and take notes about family members' behaviors, the researchers often disagree about whether certain behaviors constitute "aggression" or, instead, reflect more peaceful "assertiveness."
4. A researcher uses a blood test to determine people's overall energy level after drinking or not drinking a can of a high-caffeine cola drink. Unfortunately, when two research assistants independently rate people's behaviors for energy level for a 4-hour period after drinking the cola, their results don't seem to have any correlation with the blood-test results.
5. In a two-week period during the semester, a researcher gains entry into several college classrooms in order to administer a short survey regarding college students' beliefs about climate change. The survey consists of 20 statements about climate change (e.g., "Devastating floods in recent years are partly the result of the earth's gradually rising overall temperature"), to which students must respond "Strongly disagree," "Disagree," "Agree," or "Strongly agree." Many of the students voluntarily put their names on their surveys. Thanks to the names on many survey forms, the researcher discovers that a few

students were in two of the classes surveyed and thus completed the survey twice. Curiously, however, these students sometimes gave different responses to particular statements on the two different occasions, and hence their overall scores were also different.

6. In order to get a sense of how harmonious most long-term marriages are, a researcher administers a questionnaire to married couples who have been married for at least 20 years. The questionnaire consists of 60 statements to which both spouses must individually respond either "This describes my marriage" or "This doesn't describe my marriage." All 60 statements describe a possible characteristic of a nonharmonious marriage (e.g., "We fight all the time," "We rarely agree about how to spend our money"), and the researcher has sequenced them in a random order on the questionnaire. Even so, the researcher discovers that respondents more frequently agree with the first 30 items than with the last 30 items. If one were to look only at responses to the first 30 items, then, one would think that married couples fight a lot. But if one were to look only at responses to the last 30 items, one would conclude that most long-term couples live in relative peace and harmony. (*Note:* We recommend that questionnaires *not* be slanted in a one-way direction, as this one is; see the "Constructing a Questionnaire" guidelines in Chapter 8).
7. A researcher develops and uses a questionnaire intended to measure the extent to which college students display tolerance toward a particular religious group. However, several experts in the researcher's field of study suggest that the questionnaire measures not how tolerant students actually *are*, but what students would like to *believe* about their tolerance for people of a particular religion.
8. Students in an introductory college psychology course must satisfy their "research methods" requirement in one of several ways; one option is to participate in a research study called "Intelligence and Motor Skill Learning." When students choosing this option report to the laboratory, one of their tasks is to respond as quickly as possible to a series of simple computer-generated questions. Afterward, the researcher debriefs the students about the nature of the study and tells them that the reaction-time measure was designed to be a simple measure of intelligence. Some of the students object, saying, "That's not a measure of intelligence! Intelligence isn't how quickly you can do something, it's how *well* you can do it."

Linking Data and Research Methodology

Earlier in the chapter we described data as a link between Absolute Truth and the researcher's inquiring mind. Data are like ore: They contain pieces of the truth but are in a rather unrefined state. To extract meaning from the data, we employ what is commonly called *research methodology*.

Data and methodology are inextricably intertwined. For this reason, the methodology to be used for a particular research problem must always take into account the nature of the data that will be collected in the resolution of the problem.

An example may help clarify this point. Imagine that a man from a remote village decides to travel to the big city. While he is there, he takes his first ride on a commercial jetliner. No one in his village has ever flown before, so naturally when he returns home, they are curious about his trip. One afternoon two friends ask him about his experience, yet each one asks very different questions. The first friend asks, "How fast did you move?" "How far did you go?" and "How high did you fly?" The second friend asks, "How did you feel when you were moving so fast?" "What was it like being above the clouds?" and "What did the city look like from so high?" Both friends are asking questions that can help them learn more about the experience of flying in an airplane. Yet they differ considerably in what they want to know. Because they ask different kinds of questions, they obtain different kinds of information. If they were to recount what they had learned about "flying in a jetliner," they would most likely describe very

different things. Although neither of them has the “wrong” story, neither does each one have the whole story.

In research, too, different questions yield different types of information. Different research problems lead to different research designs and methods, which in turn result in the collection of different types of data and different interpretations of those data.

Furthermore, many kinds of data may be suitable only for a particular methodology. To some extent, *the data dictate the research method*. As an example, consider historical data, those pieces of information gleaned from written records of past events. You cannot extract much meaning from historical documents by using a laboratory experiment. An experiment is simply not suited to the nature of the data.

Over the years, numerous research methodologies have emerged to accommodate the many different forms that data are likely to take. Accordingly, we must take a broad view of the approaches the term *research methodology* encompasses. Above all, we must not limit ourselves to the belief that only a true experiment constitutes “research.” Such an attitude prohibits us from agreeing that we can better understand Coleridge’s poetry by reading the scholarly research of John Livingston Lowes (1927, 1955) or from appreciating Western civilization more because of the historiography of Arnold Toynbee (1939–1961).

No single highway leads us exclusively toward a better understanding of the unknown. Many highways can take us in that direction. They may traverse different terrain, but they all converge on the same destination: the enhancement of human knowledge.

In Chapters 6 through 10 of this book we zero in on various research methodologies. But many researchers tend to categorize research studies into two broad categories: quantitative research and qualitative research. We now look at this distinction.

Comparing Quantitative and Qualitative Approaches

As you might guess, **quantitative research** involves looking at amounts, or *quantities*, of one or more variables of interest. A quantitative researcher typically tries to measure variables in some numerical way, perhaps by using commonly accepted measures of the physical world (e.g., rulers, thermometers, oscilloscopes) or carefully designed measures of psychological characteristics or behaviors (e.g., tests, questionnaires, rating scales).

In contrast, **qualitative research** involves looking at characteristics, or *qualities*, that cannot be entirely reduced to numerical values. A qualitative researcher typically aims to examine the many nuances and complexities of a particular phenomenon. You are most likely to see qualitative research in studies of complex human situations (e.g., people’s in-depth perspectives about a particular issue, the behaviors and values of a particular cultural group) or complex human creations (e.g., television commercials, works of art). Qualitative research is not limited to research problems involving human beings, however. For instance, some biologists study, in a distinctly qualitative manner, the complex social behaviors of other animal species; Dian Fossey’s work with gorillas and Jane Goodall’s studies of chimpanzees are two well-known examples (e.g., see Fossey, 1983; Goodall, 1986).

Quantitative and qualitative approaches involve similar processes—for instance, they both entail identifying a research problem, reviewing related literature, and collecting and analyzing data. Yet these processes are often combined and carried out in different ways, leading to distinctly different research methods. For instance, quantitative researchers often start with one or more specific hypotheses to be tested. They isolate the variables they want to study, use a standardized procedure to collect some form of numerical data, and use statistical procedures to analyze and draw conclusions from the data. In contrast, qualitative researchers often start with general research questions rather than specific hypotheses, collect an extensive amount of verbal data and/or nonverbal artifacts, organize those data and artifacts into some form that gives them coherence, and use verbal descriptions to portray the situation they have studied.

To some extent, quantitative and qualitative research designs are appropriate for answering different kinds of questions. As a result, we learn more about the world when we have both

quantitative and qualitative methodologies at our disposal than when we are limited to only one approach or the other.

Let's consider how the two approaches might look in practice. Suppose two researchers are interested in investigating the "effectiveness of the case-based method of teaching business management practices." The first researcher asks the question, "How effective is case-based instruction in comparison with lecture-based instruction?" She finds five instructors who are teaching case-based business management classes; she finds five other instructors who are teaching the same content using lectures. At the end of the semester, the researcher administers an achievement test to students in all 10 classes. Using statistical analyses, she compares the scores of students in case-based and lecture-based courses to determine whether the achievement of one group is significantly higher than that of the other group. When reporting her findings, she summarizes the results of her statistical analyses. This researcher has conducted a *quantitative* study.

The second researcher is also interested in the effectiveness of the case method but asks the question, "What makes case-based instruction effective or ineffective?" To answer this question, he sits in on a case-based business management course for an entire semester. He spends an extensive amount of time talking with the instructor and some of the students in an effort to try to understand the participants' perspectives on case-based instruction. He carefully scrutinizes his data for patterns and themes in the responses. He then writes an in-depth description and interpretation of what he has observed in the classroom setting. This researcher has conducted a *qualitative* study.

Table 4.3 presents a summary of differences between quantitative and qualitative approaches. We briefly discuss these differences in the next few paragraphs—not to persuade you that one approach is better than the other, but to help you make a more informed decision about which approach might be better for your own research question.

Purpose. Quantitative researchers seek explanations and predictions that will generalize to other persons and places. The intent is to establish, confirm, or validate relationships and to develop generalizations that contribute to existing theories.

Qualitative researchers seek a better understanding of complex situations. Their work is sometimes (although not always) exploratory in nature, and they may use their observations to build theory from the ground up.

TABLE 4.3

Distinguishing characteristics of quantitative and qualitative approaches

Question	Quantitative	Qualitative
What is the purpose of the research?	• To explain and predict • To confirm and validate • To test theory	• To describe and explain • To explore and interpret • To build theory
What is the nature of the research process?	• Focused • Known variables • Established guidelines • Predetermined methods • Somewhat context-free • Detached view	• Holistic • Unknown variables • Flexible guidelines • Emergent methods • Context-bound • Personal view
What are the data like, and how are they collected?	• Numeric data • Representative, large sample • Standardized instruments	• Textual and/or image-based data • Informative, small sample • Loosely structured or nonstandardized observations and interviews
How are data analyzed to determine their meaning?	• Statistical analysis • Stress on objectivity • Deductive reasoning	• Search for themes and categories • Acknowledgment that analysis is subjective and potentially biased • Inductive reasoning
How are the findings communicated?	• Numbers • Statistics, aggregated data • Formal voice, scientific style	• Words • Narratives, individual quotes • Personal voice, literary style (in some disciplines)

Process. Because quantitative studies represent the mainstream approach to research, carefully structured guidelines exist for conducting them. Concepts, variables, hypotheses, and methods of measurement tend to be defined before the study begins and remain the same throughout. Quantitative researchers choose methods that allow them to objectively measure the variable(s) of interest. They also try to remain detached from the phenomena and participants they are observing so that they can draw unbiased conclusions.

The qualitative research process is often more holistic and “emergent,” with the specific focus, design, measurement instruments (e.g., observations, interviews), and interpretations developing and possibly changing along the way. Researchers enter the setting with open minds, prepared to immerse themselves in the complexity of the situation and interact with their participants. Categories (variables) emerge from the data, leading to “context-bound” information, patterns, and/or theories that help explain the particular phenomenon under study.

Data collection. Quantitative researchers identify one or a few variables that they intend to study and then collect data specifically related to those variables. Methods of measuring each variable are identified, developed, and standardized, with considerable attention to the validity and reliability of the measurement instruments. Data are collected from a population—or from one or more large samples that represent the population—in forms that are easily converted to numbers.

Qualitative researchers operate under the assumption that reality is not easily divided into discrete, measurable variables. Some qualitative researchers describe themselves as *being* the research instrument because the bulk of their data collection is dependent on their personal involvement in the setting. Rather than sample a large number of participants with the intent of making generalizations, qualitative researchers tend to select a few participants who might best shed light on the phenomenon under investigation. Both verbal data (interview responses, documents, field notes) and nonverbal data (drawings, photographs, videotapes) may be collected.

Data analysis. All research requires logical reasoning. Quantitative researchers tend to rely more heavily on deductive reasoning, beginning with certain premises (e.g., hypotheses, theories) and then drawing logical conclusions from them. They also try to maintain objectivity in their data analysis, conducting predetermined statistical procedures and using objective criteria to evaluate the outcomes of those procedures.

In contrast, qualitative researchers make considerable use of inductive reasoning: They make many specific observations and then draw inferences about larger and more general phenomena. Furthermore, their data analysis is more subjective in nature: They scrutinize the body of data in search of patterns—subjectively identified—that the data reflect.

It is important to note, however, that quantitative research is not exclusively deductive, nor is qualitative research exclusively inductive. Researchers of all methodological persuasions typically use both types of reasoning in a continual, cyclical fashion. Quantitative researchers often formulate a theory by inductive reasoning (e.g., by observing a few situations), engage in the theory-building process described in Chapter 1, and then try to support their theory by drawing and testing the conclusions that follow logically from it. Similarly, after qualitative researchers have identified a theme in their data using an inductive process, they typically move into a more deductive mode to verify or modify it with additional data.

Reporting findings. Quantitative researchers typically reduce their data to means, medians, correlations, and other summarizing statistics. It is not necessary to look at individual performances; rather, the *averages* of those performances are of greater interest. The results are usually presented in a report that employs a formal, scientific style using impersonal language.

Qualitative researchers often construct interpretive narratives from their data and try to capture the complexity of a particular phenomenon. Especially in certain disciplines (e.g., anthropology), qualitative researchers may use a more personal, literary style than quantitative researchers do, and they often include the participants’ own language and perspectives. Although all researchers must be able to write clearly, effective qualitative researchers must be especially skillful writers.

Combining Quantitative and Qualitative Designs

We have drawn the preceding quantitative–qualitative distinctions, in part, as a way of pointing out the relative strengths and weaknesses of the two approaches. For example, a common weakness of quantitative research with human beings is that it is sometimes conducted in a laboratory—and thus in a somewhat artificial setting. Although contrived circumstances can give the researcher considerable control over the events that occur, the results obtained may in some cases not generalize to more naturalistic settings. In contrast, qualitative research occurs within natural contexts and so, in this respect, is more “true to life.” Yet the findings of qualitative studies may be so specific to a particular context that they do not apply (generalize) to other contexts.

Fortunately, quantitative and qualitative research designs are not necessarily mutually exclusive. Many researchers successfully combine elements of both approaches in what is sometimes called a **mixed-methods design**. For example, it is not unusual for researchers to *count* (and therefore quantify) certain kinds of data in what is, for all intents and purposes, a qualitative investigation. Nor is it unusual for quantitative researchers to report participants’ perceptions of or emotional reactions to various experimental treatments. Especially in studies of human behavior, mixed-methods designs with both quantitative and qualitative elements often provide a more complete picture of a particular phenomenon than either approach could do alone. We explore mixed-methods designs in more detail in Chapter 10.

PRACTICAL APPLICATION Choosing a General Research Approach

Although we believe that research studies are sometimes enhanced by combining both quantitative and qualitative methods, we also realize that many novice researchers may not have the time, resources, or expertise to effectively combine approaches for their initial forays into research. Furthermore, good research does not necessarily have to involve a complex, multifaceted design. For example, in an article reviewing classic studies in his own discipline, psychologist Christopher Peterson had this to say in his abstract:

Psychology would be improved if researchers stopped using complicated designs, procedures, and statistical analyses for the sole reason that they are able to do so. . . . [S]ome of the classic studies in psychology [are] breathtakingly simple. . . . More generally, questions should dictate research methods and statistical analyses, not vice versa. (Peterson, 2009, p. 7)

As you choose your own general approach to addressing your research problem—whether to use a quantitative approach, a qualitative approach, or a combination of the two—you should base your decision on the research problem you want to address and the skills you have as a researcher, *not* on what tasks you wish to avoid. For example, disliking mathematics and wanting to avoid conducting statistical analyses are not good reasons for choosing a qualitative study over a quantitative one. The guidelines we offer here can help you make a reasonable decision.

GUIDELINES Deciding Whether to Use a Quantitative or Qualitative Approach

Table 4.4 can help guide you in your choice between quantitative and qualitative approaches. Keep in mind that the items in the table are not necessarily ordered from most to least important. Each item should factor into your decision. Consider each component carefully before making your final selection.

Qualitative studies have become increasingly popular in recent years, even in some disciplines that have historically placed heavy emphasis on quantitative approaches. Yet we have met many students who have naively assumed that qualitative studies are easier or in some other way more “comfortable” than quantitative designs. Be forewarned: Qualitative studies require as much effort and rigor as quantitative studies, and data collection alone often stretches over the

TABLE 4.4

Which approach should I use?

Use this approach if:	Quantitative	Qualitative
1. You believe that:	There is an objective reality that can be measured	There are multiple possible realities constructed by different individuals
2. Your audience is:	Familiar with/supportive of quantitative studies	Familiar with/supportive of qualitative studies
3. Your research question is:	Confirmatory, predictive	Exploratory, interpretive
4. The available literature is:	Relatively large	Limited
5. Your research focus:	Covers a lot of breadth	Involves in-depth study
6. Your time available is:	Relatively short	Relatively long
7. Your ability/desire to work with people is:	Medium to low	High
8. Your desire for structure is:	High	Low
9. You have skills in the areas of:	Deductive reasoning and statistics	Inductive reasoning and attention to detail
10. Your writing skills are strong in the areas of:	Technical, scientific writing	Literary, narrative writing

course of many months. In the following paragraphs, we briefly discuss each of the components in Table 4.4 from the perspective of someone who might be inclined to “go qualitative.”

1. *Consider your own comfort with the assumptions of the qualitative tradition.* If you believe that no single reality underlies your research problem but that, instead, different individuals may have constructed different, possibly equally valid realities relevant to your problem, then qualitative research is more appropriate.

2. *Consider the audience for your study.* If your intended audience (e.g., a dissertation committee, a specific journal editor, or colleagues in your field) is not accustomed to or supportive of qualitative research, it makes little sense to spend the time and effort needed to do a good qualitative study (e.g., see Miller, Nelson, & Moore, 1998).

3. *Consider the nature of the research question.* Qualitative designs can be quite helpful for addressing exploratory or interpretive research questions. But they may be of little use in testing specific hypotheses about cause-and-effect relationships.

4. *Consider the extensiveness of the related literature.* If the literature base is weak, underdeveloped, or altogether missing, a qualitative design can provide the researcher with the freedom and flexibility needed to explore a specific phenomenon so that important variables might be identified.

5. *Consider the depth of what you wish to discover.* If you want to examine a phenomenon in depth with a relatively small number of participants, a qualitative approach is ideal. But if you are skimming the surface of a phenomenon and wish to do so using a large number of participants, a quantitative study will be more efficient.

6. *Consider the amount of time you have available for conducting the study.* Qualitative studies typically involve an extensive amount of time both on site and off site. If your time is limited, you may not be able to complete a qualitative study satisfactorily.

7. *Consider the extent to which you are willing to interact with the people in your study.* Qualitative researchers who are working with human beings must be able to establish rapport and trust with their participants and interact with them on a fairly personal level. Also, gaining initial entry into a research site involves much advance planning and numerous preliminary contacts.

8. *Consider the extent to which you feel comfortable working without much structure.* Qualitative researchers typically work with few specific rules and procedures; their work is exploratory in

many respects. Thus, they must think creatively about how best to address various aspects of a research problem, and they need a high tolerance for ambiguity.

9. *Consider your ability to organize and draw inferences from a large body of information.* Qualitative research often involves the collection of a great many field notes, interview responses, and so on, that are not clearly organized at the beginning of the process. Working with extensive amounts of data and reasoning inductively about them require considerable self-discipline and organizational ability. In comparison, conducting a few statistical analyses—even for those who have little affection for mathematics—is a much easier task.

10. *Consider your writing skills.* Qualitative researchers must have excellent writing skills. Communicating findings is the final step in all research projects; the success of your research ultimately will be judged by how well you accomplish this final component of the research process.

Once you have decided whether to take a quantitative or qualitative approach, you need to pin down your research method more precisely. Table 4.5 lists some common research methodologies and the types of problems for which each is appropriate.

TABLE 4.5 Methodology and concomitant research goals

Method	Characteristics of the Method and the Research Goals the Method Attempts to Achieve
Action research	A type of applied research that focuses on finding a solution to a local problem in a local setting. For example, a teacher might investigate whether a new spelling program she has adopted leads to improvement in her students' achievement scores. (See Mertler [2009] or Mills [2011] for guidance on conducting action research.)
Case study	A type of qualitative research in which in-depth data are gathered relative to a single individual, program, or event, for the purpose of learning more about an unknown or poorly understood situation. (See Chapter 7.)
Content analysis	A detailed and systematic examination of the contents of a particular body of material (e.g., television shows, magazine advertisements, Internet websites, works of art) for the purpose of identifying patterns, themes, or biases within that material. (See Chapter 7.)
Correlational research	A statistical investigation of the relationship between two or more variables. Correlational research looks at surface relationships but does not necessarily probe for causal reasons underlying them. For example, a researcher might investigate the relationships among high school seniors' achievement test scores and their grade point averages a year later when they are first-year college students. (See Chapter 9.)
Developmental research	An observational-descriptive type of research that either compares people in different age groups (a <i>cross-sectional study</i>) or follows a particular group over a lengthy period of time (a <i>longitudinal study</i>). Such studies are particularly appropriate for looking at developmental trends. (See Chapter 9.)
Ethnography	A type of qualitative inquiry that involves an in-depth study of an intact cultural group in a natural setting. (See Chapter 7.)
Experimental research	A study in which participants are randomly assigned to groups that undergo various researcher-imposed treatments or interventions, followed by observations or measurements to assess the effects of the treatments. (See Chapter 10.)
Ex post facto research	An approach in which one looks at conditions that have already occurred and then collects data to investigate a possible relationship between these conditions and subsequent characteristics or behaviors. (See Chapter 10.)
Grounded theory research	A type of qualitative research aimed at deriving theory through the use of multiple stages of data collection and interpretation. (See Chapter 7.)
Historical research	An effort to reconstruct or interpret historical events through the gathering and interpretation of relevant historical documents and/or oral histories. (See Chapter 8.)
Observation study	A type of quantitative research in which a particular aspect of behavior is observed systematically and with as much objectivity as possible. (See Chapter 9.)
Phenomenological research	A qualitative method that attempts to understand participants' perspectives and views of social realities. (See Chapter 7.)
Quasi-experimental research	A method similar to experimental research but without random assignment to groups. (See Chapter 10.)
Survey research	A study designed to determine the incidence, frequency, and distribution of certain characteristics in a population; especially common in business, sociology, and government research. (See Chapter 9.)

Considering the Validity of Your Method

No matter what research methodology you choose, you must think about the *validity* of your approach. We have already described the importance of validity in measurement instruments. But here we are talking about the validity of *the research project as a whole*—its accuracy, meaningfulness, and credibility. Your research effort will be worth your time and effort only to the extent that it allows you to draw meaningful and defensible conclusions from your data.

When researchers consider the validity of a research study, they need to ask themselves two basic questions. First, does the study have sufficient controls to ensure that the conclusions drawn are truly warranted by the data? And second, can the results obtained reasonably be used to make generalizations about the world beyond that specific research context? The answers to these two questions address the issues of *internal validity* and *external validity*, respectively.

Internal Validity

The **internal validity** of a research study is the extent to which its design and the data it yields allow the researcher to draw accurate conclusions about cause-and-effect and other relationships within the data. To illustrate, we present three situations in which the internal validity of a study is suspect:

1. A marketing researcher wants to study how humor in television commercials affects sales in the United States and Canada. To do so, the researcher studies the effectiveness of two commercials that have been developed for a new soft drink called Zowie. One commercial, in which a well-known but humorless television actor describes how Zowie has a zingy and refreshing taste, airs during the months of March, April, and May. The other commercial, a humorous scenario in which several teenagers spray one another with Zowie on a hot summer day, airs during the months of June, July, and August. The researcher finds that in June through August, Zowie sales are almost double what they were in the preceding three months. "Humor boosts sales," the researcher concludes.
2. An industrial psychologist wants to study the effects of soft classical music on the productivity of a group of typists in a typing pool. At the beginning of the month, the psychologist meets with the typists to explain the rationale for the study, gets their consent to play the music during the working day, and then begins to have music piped into the office where the typists work. At the end of the month, the typists' supervisor reports a 30% increase in the number of documents completed by the typing pool that month. "Classical music increases productivity," the psychologist concludes.
3. An educational researcher wants to study the effectiveness of a new method of teaching reading to first graders. The researcher asks all 30 of the first-grade teachers in a particular school district whether they would like to receive training in the new method and then use it during the coming school year. Fourteen teachers volunteer to learn and use the new method; 16 teachers say that they would prefer to use their current approach. At the end of the school year, students who have been instructed with the new method have, on average, significantly higher average scores on a reading achievement test than students who have received more traditional reading instruction. "The new method is definitely better than the old one," the researcher concludes.

Did you detect something wrong with the conclusions these researchers drew? If not, go back and read the three descriptions again. *None of the conclusions is warranted from the study conducted.*

In the first research study, the two commercials apparently differed from each other in many ways (e.g., the presence of teenagers, the amount of action) in addition to humor. And, of course, we should not overlook the fact that the humorous commercial aired during the summer months. People are more likely to drink soft drinks (including Zowie) when they are hot!

In the second study, the typists *knew* they were participating in a research study; they also knew the nature of the researcher's hypothesis. Sometimes the participants in a research study

change their behavior simply because they know they are in a research study and are getting extra attention as a result. This effect, known as the **Hawthorne effect**,³ is an example of **reactivity**, a more general phenomenon in which people change their behavior when they are aware that they are being observed. But other explanations for the second study's results are possible as well. Perhaps the typists typed more because they liked the researcher and wanted to help him support his hypothesis. Perhaps the music perked up the typists for a few weeks simply because it was a change in their environment—a phenomenon known as the **novelty effect**. (In such a situation, reverting back to *no* music after a month or two might *also* lead to an increase in productivity.) Furthermore, the researcher did not consider the number of people who were working before and after the music started. Perhaps productivity increased simply because two people in the typing pool had just returned from vacation!

In the third study, notice that the researcher looked for *volunteers* to use the new method for teaching reading. Were the volunteer teachers different in some way from the nonvolunteers? Were they better educated or more motivated? Did they teach with more enthusiasm and energy because they *expected* the new method to be more effective? Or did the volunteer teachers happen to teach in areas of the school district where children had a better head start in reading skills before coming to school? Perhaps the children in the volunteers' classrooms performed better on the achievement test not because the instructional method was more effective, but because, as a group, they had been read to more frequently by their parents, gone to better preschools, and so on.

To ensure the internal validity of a research study, researchers need to take whatever precautions they can to *eliminate other possible explanations for the results observed*. Following are several strategies researchers sometimes use to increase the probability that *their explanations are the most likely ones* for the observations they have made:

- **A controlled laboratory study.** An experiment is conducted in a laboratory setting so that environmental conditions can be carefully regulated.
- **A double-blind experiment.** In a situation where two or more different interventions are being compared, neither the participants in the study nor the people administering the interventions (e.g., teachers, research assistants) know which intervention each participant is receiving. Such lack of knowledge ("blindness") decreases the likelihood that people's expectations for outcomes might influence the *actual* outcomes.
- **Unobtrusive measures.** People are observed in such a way that they do not know their actions are being recorded. We offer two real-life examples to illustrate. In one case, a university library measured student and faculty use of different parts of the library by looking at wear-and-tear patterns on the carpet. In another situation, researchers for the U.S. National Park Service looked at hikers' frequency of using different hiking trails by installing electronic counters in hard-to-notice locations beside the trails (R. K. Ormrod & Trahan, 1982). (Note that ethical issues sometimes arise when we observe people without their permission; we discuss ethics a bit later in this chapter.)
- **Triangulation.** Multiple sources of data are collected with the hope that they will all converge to support a particular hypothesis or theory. This approach is especially common in qualitative research; for instance, a researcher might engage in many informal observations in the field *and* conduct in-depth interviews, then look for common themes that appear in the data gleaned from both methods. Triangulation is also common in mixed-methods designs, in which both quantitative and qualitative data are collected to answer a single research question.

Internal validity is especially of concern in experimental designs, where the specific intent is to identify cause-and-effect relationships; accordingly, we will revisit this issue in Chapter 9. But to some degree, internal validity is important in *any* research study. The researcher must have confidence that the conclusions drawn are warranted from the data collected.

³The effect owes its name to the Hawthorne Works, an industrial complex in Illinois where the effect was first observed.

External Validity

The **external validity** of a research study is the extent to which its results apply to situations beyond the study itself—in other words, the extent to which the conclusions drawn can be *generalized* to other contexts. As a general rule, researchers contribute more to humanity's knowledge about the world when they conduct research that has implications that extend far beyond the specific situation studied.

Following are three commonly used strategies that enhance the external validity of a research project:

- **A real-life setting.** Earlier we mentioned that researchers sometimes use laboratory experiments to help them control the environmental conditions in which a study takes place. Laboratory studies have a downside, however: They provide an artificial setting that may be quite different from real-life circumstances. Research that is conducted in the outside world, although it may not have the tight controls of a laboratory project, may be more valid in the sense that it yields results with broader applicability to other real-world contexts.⁴
- **A representative sample.** Whenever researchers seek to learn more about a particular category of objects or creatures—whether they are studying rocks, salamanders, or human beings—they often study a *sample* from that category and then draw conclusions about the category as a whole. (Here is a classic example of inductive reasoning.) For example, to study the properties of granite, researchers might take pieces of granite from anywhere in the world and assume that their findings based on those pieces might be generalizable to granite found in other locations. The same might hold true for salamanders if researchers limit their conclusions to the particular species of salamander they have studied.
Human beings are another matter. The human race is incredibly diverse in terms of culture, childrearing practices, educational opportunities, personality characteristics, and so on. To the extent that researchers restrict their research to people with a particular set of characteristics, they may not be able to generalize their findings to people with a very different set of characteristics. Ideally, then, researchers want participants in a research study to be a *representative sample* of the population about which they wish to draw conclusions. In Chapter 8, we consider a number of strategies for obtaining representative samples.
- **Replication in a different context.** Imagine that one researcher draws a conclusion from a particular study in a specific context, and another researcher who conducts a similar study in a very different context reaches the same conclusion, and perhaps additional researchers also conduct similar studies in dissimilar contexts and, again, draw the same conclusion. Under such circumstances, these studies, taken together, provide evidence that the conclusion has validity and applicability across diverse contexts and situations.

Any researcher should consider both internal validity and external validity when designing a research project. One's conclusions are valid and meaningful only to the extent that they are warranted based on the data collected *and* have applicability beyond the specific research situation itself.

In Chapter 2 we introduced you to the distinction between basic and applied research. Well-designed basic research—research conducted under tightly controlled (and possibly artificial) conditions—ensures internal validity; that is, it allows the researcher to rule out other possible explanations for the results obtained. Applied research—research conducted in more naturalistic but invariably more complex environments—is more useful for external validity; that is, it increases the chances that a study's findings are generalizable to other real-life situations and

⁴The artificial nature of laboratory research has been a concern in psychology for many years. Anderson, Lindsay, and Bushman (1999) compared the results of laboratory and field studies related to a wide variety of psychological phenomena. They discovered that the two kinds of studies typically yield similar results—that studies conducted in the lab and those conducted in natural settings lead to the same conclusions about human nature. In their own words, "The psychological laboratory has generally produced psychological truths, rather than trivialities" (p. 3).

problems. Keep in mind, however, that the basic-versus-applied distinction is really a continuum rather than a dichotomy: Research studies can have varying degrees of artificiality versus real-world authenticity.

Validity in Qualitative Research

The concepts of internal and external validity originated in discussions of quantitative research (Campbell & Stanley, 1963), and some qualitative researchers have questioned their relevance to qualitative designs. For instance, Lincoln and Guba (1985) and Creswell (2007) have suggested that, for qualitative research, such words as *credibility*, *trustworthiness*, *confirmability*, and *validation* be used instead of the term *validity*.

As noted earlier, qualitative researchers frequently use triangulation—comparing multiple data sources in search of common themes—to support the validity of their findings. Following are several additional strategies they employ:

- **Extensive time in the field.** The researcher may spend several months, perhaps even a year or more, studying a particular phenomenon, forming tentative hypotheses, and continually looking for evidence that either supports or disconfirms those hypotheses.
- **Negative case analysis.** The researcher actively looks for cases that contradict existing hypotheses, then continually revises his or her explanation or theory until all cases have been accounted for.
- **Thick description.** The situation is described in sufficiently rich, “thick” detail that readers can draw their own conclusions from the data presented.
- **Feedback from others.** The researcher seeks the opinion of colleagues in the field to determine whether they agree or disagree that the researcher has made appropriate interpretations and drawn valid conclusions from the data.
- **Respondent validation.** The researcher takes his or her conclusions back to the participants in the study and asks quite simply, Do you agree with my conclusions? Do they make sense based on your own experiences?

As you can see, then, researchers use a wide variety of approaches to support the validity of their findings. Different approaches are appropriate in different situations, depending on the nature of the data and the specific methodologies used.

Regardless of the kind of study you decide to conduct, you must address the validity of your study at the very beginning of your project—that is, *at the planning stage*. If you put off validity issues until later in the game, you may end up conducting a study that has little apparent validity, either in terms of minimizing alternative explanations for the results obtained (internal validity) or in terms of being generalizable to the world “out there” (external validity). As a result, you are almost certainly wasting your time and effort on what is, for all intents and purposes, a trivial enterprise.

Ethical Issues in Research

Within certain disciplines—the social sciences, education, medicine, and similar areas of study—the use of human beings in research is, of course, quite common. And in biology the subjects of investigation are often nonhuman animals. Whenever human beings or other creatures with the potential to think, feel, and experience physical or psychological distress are the focus of investigation, researchers must look closely at the ethical implications of what they are proposing to do.

Most ethical issues in research fall into one of four categories: protection from harm, voluntary and informed participation, right to privacy, and honesty with professional colleagues. In the following sections we raise issues related to each of these categories. We then describe the internal review boards and professional codes of ethics that provide guidance for researchers.

Protection from Harm

Researchers should not expose research participants—whether they be human beings or non-human animals—to unnecessary physical or psychological harm. When a study involves human beings, the general rule of thumb is that the risk involved in participating in a study should not be appreciably greater than the normal risks of day-to-day living. Participants should not risk losing life or limb, nor should they be subjected to unusual stress, embarrassment, or loss of self-esteem.

In thinking about this issue, researchers must be particularly sensitive to and thoughtful about potential harm they might cause participants from potentially vulnerable populations (Sieber, 2000). For example, some participants may have allergies or health conditions that place them at greater-than-average risk in certain environments or with certain foods or medications. Participants of a particular gender, cultural background, or sexual orientation might feel embarrassed or otherwise uncomfortable when asked to answer some kinds of questions or engage in some kinds of activities. Special care must be taken with participants who cannot easily advocate for their own needs and desires—children, the elderly, people with disabilities, and the like.

Especially when working with human participants, a researcher should ideally also think about potential *benefits* that participation in a study might offer. At a minimum, the researcher should treat all participants in a courteous and respectful manner. A researcher can also consider how people might gain something useful from participating in a study—perhaps unique insights about a topic of personal interest or perhaps simply a sense of satisfaction about contributing in a small way to advancements in society's collective knowledge about the world. In some cases a researcher can offer an incentive for participating (e.g., money or course credit), provided that it isn't so excessive that it's essentially a form of disguised coercion (Scott-Jones, 2000).⁵

In cases where the nature of a study involves creating a small amount of psychological discomfort, participants should know this ahead of time, and any necessary debriefing or counseling should follow immediately after their participation. A debriefing can simultaneously accomplish several things (Sales & Folkman, 2000):

- It can help alleviate any uncomfortable reactions—either anticipated or unanticipated—to certain questions, tasks, or activities.
- It can alert the researcher to necessary follow-up interventions for any participants experiencing extreme reactions.
- It provides an opportunity for the researcher to correct any misinformation participants might have gotten during the study.
- It provides a time during which participants can learn more about the nature and goals of the study, about how its results may fit in with what is already known about a topic, and about the nature of research more generally.

Voluntary and Informed Participation

When research involves public documents or records that human beings have previously created—birth certificates, newspaper articles, and so on—such documents and records are generally considered to be fair game for research investigation. But when people are specifically recruited for participation in a research study, they should be told the nature of the study to be conducted and given the choice of either participating or not participating. Furthermore, they should be told that, if they agree to participate, they have the right to withdraw from the study at any time. And under no circumstances should people feel pressure to participate from employers or other more powerful individuals. *Any participation in a study should be strictly voluntary.*

In general, research with human beings requires **informed consent**. That is, participants—or legal guardians in the case of children and certain other populations—must know the nature

⁵ Two qualifications should be noted here. When working with children, enticing incentives should be offered *only after* parents have already given permission for their participation. And when offering course credit to college students, alternative ways to earn the same credit must be provided as well—for instance, reading and writing a review of a research article (Scott-Jones, 2000).

of the study and grant written permission. One common practice—and one that is *required* for certain kinds of studies at most research institutions—is to present an **informed consent form** that describes the nature of the research project, as well as the nature of one's participation in it. Such a form should contain the following information:

- A brief description of the nature and goal(s) of the study, written in language that its readers can readily understand
- A description of what participation will involve, in terms of activities and duration
- A statement indicating that participation is voluntary and can be terminated at any time without penalty
- A description of any potential risk and/or discomfort that participants may encounter
- A description of potential benefits of the study, including those for participants, science, and/or human society as a whole
- A guarantee that all responses will remain confidential and anonymous
- The researcher's name, plus information about how the researcher can be contacted
- An individual or office that participants can contact if they have questions or concerns about the study
- An offer to provide detailed information about the study (e.g., a summary of findings) upon its completion
- A place for the participant to sign and date the letter, indicating agreement to participate (when children are asked to participate, their parents must read and sign the letter)

An example of such a form, used by Rose McCallin in a research project for her doctoral dissertation, is presented in Figure 4.6. The form was used to recruit college students who were enrolled in a class in a teacher preparation program. It is missing one important ingredient: an offer to provide information about the study after its completion. Instead, McCallin appeared in class a few weeks after she had collected data to give a summary of the study and its implications for teachers.

A dilemma sometimes arises as to *how informed* potential participants should be. If people are given too much information—for instance, if they are told the specific research hypothesis being tested—they may behave differently than they would under more normal circumstances (recall the earlier description of a study involving classical music and typists' productivity). A reasonable compromise is to give potential participants a general idea of what the study is about (e.g., "This study is investigating the effects of a physical exercise program on people's overall mental health") and to describe what specific activities their participation will involve—in other words, to give them sufficient information to make a reasonable, informed judgment about whether they wish to participate.

On rare occasions (e.g., in some studies of social behavior), telling participants the true nature of a study might lead them to behave in ways that would defeat the purpose of the study. In general, deception of any kind is frowned on and should be used only when the study cannot meaningfully be conducted without it. Even then, the degree of deception should be as minimal as possible, and participants should be told the true nature of the research as soon as their involvement is over. (An internal review board, to be described shortly, can give you guidance regarding this matter.)

Earlier in the chapter we mentioned the use of unobtrusive measures as a strategy for measuring behavior. Strictly speaking, unobtrusive measures violate the principle of informed consent. But if people's behaviors are merely being recorded in some way during their normal daily activities—if people are not being asked to do something they ordinarily would not do—and if they are not being scrutinized in any way that might be potentially invasive or embarrassing, then unobtrusive measures are quite appropriate. Recall our two earlier examples: examining the frequency with which people use different parts of the library and the frequency with which they hike along certain trails in a national park. Both of these examples involved behaviors within the scope of participants' normal activities.

Understanding How Students Organize Knowledge

You are being asked to participate in a study investigating ways in which students organize their knowledge.

We are interested in determining how students organize their knowledge in memory and use that knowledge. It is hoped that the results of this study can be useful in helping teachers understand why students perform differently from one another in the classroom.

As a future teacher, you will most likely have to use your knowledge in a variety of situations. However, relatively little is known about relationships among factors involved in knowledge application. Your participation may help to clarify some of these relationships so that we can better identify why students perform differently. And, although you may not directly benefit from this research, results from the study may be useful for future students, both those you teach and those who, like yourself, plan to be teachers.

If you agree to participate, you will complete two activities. In addition, we need to use your anonymous grade point average (GPA) as a control variable in order to account for initial differences among students. To ensure anonymity, we will submit only your social security number to the UNC Registrar, who will use this number to locate your GPA. The Registrar will black out the first three digits of your social security number before giving us this information, and the remaining six-digit number will be used only to keep track of your performance on the other activities. You will not be putting your name on anything except this form. And, there will be no attempt to link your name with the last six digits of your social security number because individual performance is not of interest in this study. Only group results will be reported.

In the first activity, you will be asked to complete a 15-minute Self-Rating Checklist. This checklist consists of statements about knowledge application that you will judge to be true or false according to how each statement applies to you. In the second activity (which will be administered two days later), you will be given a list of concepts and asked to organize them on a sheet of paper, connect concepts you believe to be related, and describe the type of relationship between each connected pair of concepts. This activity should take about 30 minutes.

Although all studies have some degree of risk, the potential in this investigation is quite minimal. All activities are similar to normal classroom procedures, and all performance is anonymous. You will not incur any costs as a result of your participation in this study.

Your participation is voluntary. If at any time during this study you wish to withdraw your participation, you are free to do so without prejudice.

If you have any questions prior to your participation or at any time during the study, please do not hesitate to contact us.

AUTHORIZATION: I have read the above and understand the nature of this study. I understand that by agreeing to participate in this study I have not waived any legal or human right and that I may contact the researchers at the University of Northern Colorado (Dr. Jeanne Ormrod or Rose McCallin, 303-555-2807) at any time. I agree to participate in this study. I understand that I may refuse to participate or I may withdraw from the study at any time without prejudice. I also grant permission to the researchers to obtain my anonymous grade point average from the UNC Registrar for use as a control variable in the study. In addition, I understand that if I have any concerns about my treatment during the study, I can contact the Chair of the Internal Review Board at the University of Northern Colorado (303-555-2392) at any time.

Participant's signature: _____ Date: _____

Researcher's signature: _____ Date: _____

FIGURE 4.6

Example of an informed consent form

Adapted from *Knowledge Application Orientation, Cognitive Structure, and Achievement* (pp. 109–110) by R. C. McCallin, 1988, unpublished doctoral dissertation, University of Northern Colorado, Greeley. Adapted with permission.

Right to Privacy

Any research study involving human beings must respect participants' right to privacy. Under no circumstances should a research report, either oral or written, be presented in such a way that other people become aware of how a particular participant has responded or behaved—unless, of course, the participant has specifically granted permission *in writing* for this to happen.

In general, a researcher must keep the nature and quality of individual participants' performance strictly confidential. For instance, the researcher might give each participant a unique, arbitrary code number and then label any written documents with that number rather than with the person's name. And if a particular person's behavior is described in depth in the research

report, he or she should be given a pseudonym—and other trivial, irrelevant details that might give away the person's identity should be changed—to ensure anonymity.

In this age of the Internet, researchers must also take precautions that computer hackers cannot access participants' individual data. Our advice here is simple: Don't post raw data or easily decodable data about individual participants online in any form. If you use the Internet to share your data with co-researchers living elsewhere, use e-mail and well-encoded attachments to transmit your data set; send your coding scheme in a separate e-mail message at another time.

Occasionally employers or other powerful individuals in a research setting might put considerable pressure on a researcher to reveal participants' individual responses. The researcher must not give in to such pressure. In general, knowledge about participants' individual performances should be revealed *only* to any co-researchers who have a significant role in the research investigation—unless, of course, participants have specifically granted permission *in writing* that it be shared with certain other individuals. There is one important exception to this rule: Researchers are legally obligated to report to the proper authorities any information that suggests present or imminent danger to someone (e.g., child abuse, a planned terrorist act).

Honesty with Professional Colleagues

Researchers must report their findings in a complete and honest fashion, without misrepresenting what they have done or intentionally misleading others about the nature of their findings. And under no circumstances should a researcher fabricate data to support a particular conclusion, no matter how seemingly “noble” that conclusion might be. Such an action constitutes scientific fraud, plain and simple.

Within this context, we ask you to recall our discussion in Chapter 3 about giving appropriate credit where credit is due. Any use of another person's ideas or words demands full acknowledgment; otherwise, it constitutes plagiarism and documentary theft. Full acknowledgment of all material belonging to another person is mandatory. To appropriate the thoughts, ideas, or words of another without acknowledgment—even if you paraphrase the borrowed ideas in your own language—is dishonest, unethical, and highly circumspect. Honest researchers do not hesitate to acknowledge their indebtedness to others.

Internal Review Boards

In the United States, any college, university, or research institution must have an **internal review board (IRB)**⁶ that scrutinizes all proposals for conducting human research under the auspices of the institution. This board, which is made up of scholars and researchers across a broad range of disciplines, checks proposed research studies to be sure that the procedures are not unduly harmful to participants, that appropriate procedures will be followed to obtain participants' informed consent, and that participants' privacy and anonymity are ensured.

It is important to note that the research is reviewed at the proposal stage. *A proposal must be submitted to and approved by the IRB before a single datum is collected.* Depending on the extent to which the study intrudes in some way on people's lives and imposes risk to participants, the board's chairperson may quickly declare it *exempt from review*, give it an *expedited review*, or bring it before the board for a *full review*. In any case, the researcher cannot begin the study until (a) the board has given its seal of approval to the study as originally designed or (b) the researcher has made modifications that the board requests.

The criteria and procedures of an IRB vary slightly from one institution to another. For examples of institutional policies and procedures, you might want to visit the websites of Duke University (www.ors.duke.edu), Tufts University (tnemcirb.tufts.edu), or University of New Hampshire (www.unh.edu/osr). You can find other helpful sites on the Internet by using a search engine (e.g., Google or Yahoo!) and such keywords as “IRB,” “human participants,” and “human subjects.”

⁶ You might also see this committee called something along the lines of “Committee for Protection of Human Subjects.”

Universities and other research institutions have review boards for animal research as well. Any research that may potentially cause suffering, distress, or death to animals must be described and adequately justified to an **institutional animal care and use committee (IACUC)**. Furthermore, the researcher must minimize or prevent such suffering and death to the extent it is possible to do so. For examples of research institutions' IACUC policies and procedures, we refer you to the University of Maryland (www.umresearch.umd.edu/IACUC) and the University of Arizona (www.uac.arizona.edu).

Professional Codes of Ethics

Many disciplines have their own codes of ethical standards governing research that involves human subjects and, when applicable, research involving animal subjects as well. One good source of discipline-specific ethical codes is, of course, the Internet. Following are examples of organizational websites with ethical codes related to research in their disciplines:

- American Anthropological Association (www.aaanet.org)
- American Association for Public Opinion Research (www.aapor.org)
- American Educational Research Association (www.aera.net)
- American Psychological Association (www.apa.org)
- American Society of Criminology (www.asc41.com)
- American Sociological Association (www.asanet.org)
- Society for Conservation Biology (www.conbio.org)

PRACTICAL APPLICATION Planning an Ethical Research Study

Ethical practices in research begin *at the planning* stage. The following checklist can help you scrutinize your own project for its potential ethical implications.

CHECKLIST

Determining Whether Your Proposed Study Is Ethically Defensible

- _____ 1. Might your study present any physical risks or hazards to participants? If so, list them here.

- _____ 2. Might your study incur any psychological harm to all or some participants (e.g., offensive stimulus materials, threats to self-esteem)? If so, identify the specific forms of harm that might result.

- _____ 3. Will participants incur any significant financial costs (e.g., transportation costs, mailing expenses)? If so, how might you minimize or eliminate those costs?

- _____ 4. What benefits might your study have for (a) participants, (b) your discipline, and (c) society at large?

_____ 5. Do you need to seek informed consent from participants? Why or why not?

_____ 6. If you need to seek informed consent, how might you explain the nature and goals of your study to potential participants in a way that they can understand? Write a potential explanation here.

_____ 7. What specific steps will you take to ensure participants' privacy? List them here.

_____ 8. If applicable, what format might a post-participation debriefing take? What information should you include in your debriefing?

Critically Scrutinizing Your Overall Plan

At this point, you have presumably attended to the nature and availability of the data you need in order to address your research problem, identified potentially appropriate ways of measuring your observations, chosen an overall approach to your research methodology, and examined the ethical implications of what you intend to do. But ultimately, you must step back a bit and look at the overall forest—the big picture—rather than at the specific, nitty-gritty trees. And, of course, you must be realistic and practical regarding what you can reasonably accomplish. Remember the title of this book: *Practical Research*.

PRACTICAL APPLICATION Judging the Feasibility of a Research Project

Many beginning researchers avoid looking closely at the practical aspects of a research endeavor. Envisioning an exotic investigation or a solve-the-problems-of-the-world study sometimes keeps a researcher from making an impartial judgment about practicability. Completing the following checklist can help you wisely plan and accurately evaluate the research you have in mind. After you have finished, review your responses. Then answer this question: Can you reasonably accomplish this study? If your answer is *no*, determine which parts of the project are not terribly practical, and identify things you might do to make it more realistically accomplishable.

✓ CHECKLIST

Determining Whether a Proposed Research Project Is Realistic and Practical

THE PROBLEM

- _____ 1. With what area(s) will the problem deal?
- _____ People
- _____ Things
- _____ Records

- ☐ _____ Thoughts and ideas
☐ _____ Dynamics and energy
☐ 2. Are data that relate directly to the problem available for each of the categories you've just checked? ☐ Yes ☐ No
☐ 3. What academic discipline is primarily concerned with the problem?

☐ 4. What other academic disciplines are possibly also related to the problem?

☐ 5. What special qualifications do you have as a researcher for this problem?
☐ Interest in the problem
☐ Experience in the problem area
☐ Education and/or training
☐ Other(specify): _____

THE DATA

- ☐ 6. How available are the data to you?
☐ Readily available
☐ Available with permission
☐ Available with great difficulty or rarely available
☐ Unavailable
☐ 7. How often are you personally in contact with the source of the data?
☐ Once a day ☐ Once a week ☐ Once a month
☐ Once a year ☐ Never
☐ 8. Will the data arise directly out of the problem situation?
☐ Yes ☐ No
 If your answer is no, where or how will you obtain the data?

☐ 9. How do you plan to gather the data?
☐ Observation ☐ Questionnaire ☐ Test ☐ Rating scale
☐ Photocopying of records ☐ Interview and audio recording
☐ Specialized machine/device ☐ Computer technology
☐ Other (explain): _____

☐ 10. Is special equipment or are special conditions necessary for gathering or processing the data?
☐ Yes ☐ No
 If your answer is yes, specify: _____

☐ 11. If you will need special equipment, do you have access to such equipment and the skill to use it?
☐ Yes ☐ No

If your answer is no, how do you intend to overcome this difficulty?

_____ 12. What is the estimated cost in time and money to gather the data?

_____ 13. What evidence do you have that the data you gather will be valid and reliable indicators of the phenomena you wish to study?

CRITERIA-BASED EVALUATION

_____ 14. Does your research project meet the four criteria applicable to all research? (For a refresher on these criteria, see the section “General Criteria for a Research Project.”)

Universality _____ Yes _____ No

Replication _____ Yes _____ No

Control _____ Yes _____ No

Measurement _____ Yes _____ No

_____ 15. As you review your responses to this checklist, might any of the factors you’ve just considered, or perhaps any other factors, hinder a successful completion of your research project?

_____ Yes _____ No

If your answer is yes, list those factors.

When You Can’t Anticipate Everything in Advance: The Value of a Pilot Study

Did you have trouble answering some of the questions in the checklist? For instance, did you have difficulty estimating how much time it would take you to gather or process your data? Did you realize that you might need to develop your own questionnaire, test, or other measurement instrument but then wonder how valid and reliable the instrument might be for your purpose?

Up to this point, we have been talking about planning a research project as something that occurs all in one fell swoop. In reality, a researcher may sometimes need to do a brief exploratory investigation, or **pilot study**, to try out particular procedures, measurement instruments, or methods of analysis. *A brief pilot study is an excellent way to determine the feasibility of your study.* Furthermore, although it may take some time initially, it may ultimately save you time by letting you know—after only a small investment on your part—which approaches will and will not be effective in helping you solve your overall research problem.

PRACTICAL APPLICATION Establishing Research Criteria and Justifying Your Research Methodology

In this exercise, you will continue developing your research proposal, for which you have already conducted a review of the related literature (see Chapter 3). In writing proposals, we leave nothing unspecified. Here we explore how to treat each subproblem in terms of its data, the criteria

for the admission of those data into your study, and a justification of the methodology you propose to employ. In the form presented in Figure 4.7 (extend it as needed on either a sheet of paper or word processing document), we suggest the following approach to justify your research methods:

1. Write your principal problem at the top of the page, and divide the rest of the page into the two columns.
2. In the left-hand column, write the first subproblem.
3. Immediately below the subproblem, write a description of the data you will need to resolve that subproblem.
4. In the right-hand column, write the criteria you will establish for the admissibility of those data into your research design. In other words, what kinds of data will be acceptable and what kinds will *not* be acceptable for inclusion in your study? Be very specific. For instance, if you are planning to administer a questionnaire to determine people's attitudes about a controversial issue, will you include data from questionnaires that are only partially completed, with some items left unanswered?
5. Repeat steps 2–4 for the remaining subproblems.

Now refer to the earlier sections “Comparing Quantitative and Qualitative Approaches” and “Guidelines: Deciding Whether to Use a Quantitative or Qualitative Approach.” Then do the following:

1. Describe the characteristics that the data in your research project will exhibit.
2. For the data you have just described, identify the approach—quantitative, qualitative, or some combination of the two—that would be most appropriate for collecting, analyzing, and interpreting those data. Justify your choice.

PRACTICAL APPLICATION Developing a Plan of Attack

Once you have determined that your research project is feasible, you can move ahead. Yet especially for a novice researcher, all the things that need to be done—writing and submitting the proposal, getting IRB or IACUC approval, arranging for access to one or more research sites, setting up any experimental interventions you have planned, collecting the data, analyzing and

FIGURE 4.7
Justifying your research approaches

The problem: (Write it as you stated it in Chapter 3.)	
A statement of each subproblem together with a description of those data needed for its solution. Subproblem 1: The data needed for its solution:	The criteria that will govern the admissibility of the data into the research design. Criterion 1: Criterion 2: Criterion 3:

interpreting it, and writing the final research report (almost always in multiple drafts)—may, in combination, seem like a gigantic undertaking. We authors recall, with considerable disappointment and sadness, the many promising doctoral students we have known who took all required courses, passed their comprehensive exams with flying colors, and then never earned their doctoral degrees because they couldn't persevere through the process of completing a dissertation.

You must accept the fact that *your project will take time—lots of time*. All too often, we have had students tell us that they anticipate completing a major research project (e.g., a thesis or dissertation) in a semester or less. In the vast majority of cases, such a belief is unrealistic. Consider the many steps we listed in the preceding paragraph. If you think you can accomplish all of these things within two or three months, you are almost certainly setting yourself up for failure and great disappointment. We would much rather you think of any research project—and especially your first project—as something that is a valuable learning experience in its own right. As such, it is worth however much of your time and effort it takes to do the job well.

The most effective strategy we can suggest here is to *develop a research and writing schedule and try to stick to it*. Figure 4.8 provides a workable format for your schedule. In the left-hand column, list all of the specific tasks you need to accomplish for your research project (writing the proposal, getting approval from the IRB and any other relevant faculty committees, conducting any needed pilot studies, etc.) in the order in which you need to accomplish them. In the second column, estimate the number of weeks or months it will take you to complete each task, always giving yourself a little more time than you think you will need. In the third column, establish appropriate target dates for accomplishing each task, taking into account any holidays, vacations, business trips, and other breaks in your schedule that you anticipate. Also include a little bit of slack time for unanticipated illnesses or family emergencies. Use the right-hand column to check off each step as you complete it.

Using Project Management Software and Electronic Planners



Project management software is commonplace in the business world. Such software is available both commercially (e.g., FastTrack Schedule, ManyMoon, Milestones, ToDoList) and in freeware available for download from the Internet (e.g., go to www.ganttproject.biz or download.cnet.com). You can use such software to organize and coordinate the various aspects of a research project. For example, it lets you outline the different phases of the project, the dates by which those phases need to be completed, the ways in which they are interconnected, and the person who is responsible for completing each task. This information can be displayed in graphic form with due dates and milestones highlighted.

Project management software is especially helpful when a research project has many separate parts that all need to be carefully organized and coordinated. For example, suppose a large research effort is being conducted in a local school district. The effort requires a team of observers and interviewers to go into various schools and observe teachers in class, interview students during study halls, and discuss administrative issues with the principals. Coordinating the efforts of the many observers, teachers, students, and administrators is a complex task that can be easily laid out and scheduled by project management software.

You might consider computerizing your schedule even if you don't expect your research project to be as multifaceted as the one just described. For instance, computerized versions of day planners (e.g., Daily Planner, DevPlanner, Now Up-to-Date) are widely available. These customized databases can store and organize the various tasks you need to accomplish. You can insert electronic reminders that you need to do certain things on such-and-such a date. And you can easily revise your long-term schedule if unforeseen circumstances occur.

Keeping an Optimistic and Task-Oriented Outlook

In our own experiences, we have found that a schedule goes a long way in helping us complete a seemingly humongous task. In fact, this is exactly the approach we took when we wrote various editions of this book. Make no mistake about it, writing a book such as this one can be even more overwhelming than conducting a research project!

[illegible]

FIGURE 4.8

Establishing a schedule for your project

A schedule in which you break your project into small, easily doable steps accomplishes several things for you simultaneously. First, it gives you the confidence that you *can* complete your project if you simply focus on one piece at a time. Second, it helps you persevere by giving you a series of target dates that you strive to meet. And last (but certainly not least!), checking off each task as you complete it provides a regular reminder that you are making progress toward your final goal of solving the research problem.

For Further Reading

Planning Your Research Design

- Bordens, K. S., & Abbott, B. B. (2010). *Research design and methods: A process approach* (8th ed.). New York: McGraw-Hill.
- Butler, D. L. (2006). Frames of inquiry in educational psychology: Beyond the quantitative-qualitative divide. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 903–927). Mahwah, NJ: Erlbaum.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.
- Ercikan, K., & Roth, W.-M. (2006). What good is polarizing research into qualitative and quantitative? *Educational Researcher*, 35(5), 14–23.
- Ethridge, D. (2004). *Research methodology in applied economics: Organizing, planning, and conducting economic research* (2nd ed.). New York: Wiley.
- Firestone, W. A. (1987). Meaning in method: The rhetoric of quantitative and qualitative research. *Educational Researcher*, 16(7), 16–21.
- Hedrick, T. E., Bickman, L., & Rog, D. J. (1993). *Applied research design: A practical guide*. Thousand Oaks, CA: Sage.
- Herr, K., & Anderson, G. L. (2005). *The action research dissertation: A guide for students and faculty*. Thousand Oaks, CA: Sage.
- Jacob, H. (1984). *Using published data: Errors and remedies*. Thousand Oaks, CA: Sage.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14–26.
- Malhotra, N. K. (2010). *Marketing research: An applied orientation* (6th ed.). Upper Saddle River, NJ: Prentice Hall.
- Maxfield, M. G., & Babbie, E. R. (2011). *Research methods for criminal justice and criminology* (6th ed.). Belmont, CA: Wadsworth/Cengage Learning.
- Neuman, W. L. (2011). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Upper Saddle River, NJ: Pearson.
- Singleton, R. A., Jr., & Straits, B. C. (2009). *Approaches to social research* (5th ed.). New York: Oxford University Press.
- Tashakkori, A., & Teddlie, C. (Eds.) (2010). *SAGE handbook of mixed methods in social and behavioral research* (2nd ed.). Thousand Oaks, CA: Sage.
- Wood, M. J., & Ross-Kerr, J. C. (2011). *Basic steps in planning nursing research: From question to proposal* (7th ed.). Sudbury, MA: Jones & Bartlett.

Fried, H. O., Knox Lovell, C. A., & Schmidt, S. S. (Eds.). (2008). *The measurement of productive efficiency and productivity growth*. New York: Oxford University Press.

Miller, D. C., & Salkind, N. J. (2002). *Handbook of research design and social measurement* (6th ed.). Thousand Oaks, CA: Sage.

Thorndike, R. M., & Thorndike-Christ, T. (2010). *Measurement and evaluation in psychology and education* (8th ed.). Upper Saddle River, NJ: Merrill/Pearson Education.

Ethics

American Educational Research Association. (1992). Ethical standards of the American Educational Research Association. *Educational Researcher*, 21(7), 23–36.

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist*, 57, 1060–1073.

Bankowski, Z., & Levine, R. J. (Eds.). (1993). *Ethics and research on human subjects: International guidelines*. Albany, NY: World Health Organization.

Brickhouse, N. W. (1992). Ethics in field-based research: Ethical principles and relational considerations. *Science Education*, 76(1), 93–103.

Cheney, D. (Ed.). (1993). *Ethical issues in research*. Frederick, MD: University Publishing Group.

Christians, C. G. (2000). Ethics and politics in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 133–155). Thousand Oaks, CA: Sage.

Eiserman, W. C., & Behl, D. (1992). Research participation: Benefits and considerations for the special educator. *Teaching Exceptional Children*, 24, 12–15.

Elliott, D., & Stern, J. E. (Eds.). (1997). *Research ethics: A reader*. Hanover, NH: University Press of New England.

Erwin, E., Gendin, S., & Kleiman, L. (Eds.). (1994). *Ethical issues in scientific research: An anthology*. New York: Garland.

Hemmings, A. (2009). Ethnographic research with adolescent students: Situated fieldwork ethics and ethical principles governing human research. *Journal of Empirical Research on Human Research Ethics*, 4(4), 27–38.

King, N. M. P., & Churchill, L. R. (2000). Ethical principles guiding research on child and adolescent subjects. *Journal of Interpersonal Violence (Special Issue: The Ethical, Legal, and Methodological Implications of Directly Asking Children about Abuse)*, 15, 710–724.

Loue, S., & Case, S. L. (2000). *Textbook of research ethics: Theory and practice*. New York: Plenum Press.

Macrina, F. L. (2005). *Scientific integrity: Text and cases in responsible conduct of research* (3rd ed.). Washington, DC: American Society for Microbiology.

Measurement

Aft, L. (1992). *Productivity measurement and improvement* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Earickson, R., & Harlin, J. (1994). *Geographic measurement and quantitative analysis*. Upper Saddle River, NJ: Prentice Hall.

- Mertens, D. M., & Ginsberg, P. (Eds.). (2008). *The handbook of social research ethics*. Thousand Oaks, CA: Sage.
- Neuman, W. L. (2011). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Upper Saddle River, NJ: Pearson. [Provides an excellent discussion of ethical issues.]
- Panther, A. T., & Sterba, S. K. (Eds.). *Handbook of ethics in quantitative methodology*. New York: Routledge.
- Pimple, K. D. (2008). *Research ethics*. Aldershot, England: Ashgate.
- Pimple, K. D., Orlans, F. B., & Gluck, J. P. (1997). *Ethical issues in the use of animals in research*. Mahwah, NJ: Erlbaum.
- Roberts, L. W. (2006). Ethical principles and practices for research involving human participants with mental illness. *Psychiatric Services*, 57, 552–557.
- Sales, B. D., & Folkman, S. (Eds.). (2000). *Ethics in research with human participants*. Washington, DC: American Psychological Association.
- Siebert, J. E. (1992). *Planning ethically responsible research: A guide for students and internal review boards*. Thousand Oaks, CA: Sage.
- Yan, E. G., & Munir, K. M. (2004). Regulatory and ethical principles in research involving children and individuals with developmental disabilities. *Ethics & Behavior*, 14(1), 31–49.

ANSWERS TO THE CONCEPTUAL ANALYSIS EXERCISE "Identifying Scales of Measurement" on page 88:

1. This is a *ratio* scale, with an absolute zero point (i.e., no bacteria at all).
2. A country's average temperature is an *interval* scale, because an average temperature of 0—regardless of whether the temperature is reported in Fahrenheit or Celsius—still means *some* heat. (In the Kelvin scale, a temperature of 0 means no heat at all, but people typically don't use this scale in reporting climatic temperature.) Amount of tourist dollars is, of course, a *ratio* scale.
3. The party membership coding scheme is a *nominal* scale, because the numbers assigned indicate only category membership, not quantity or order. For example, a Republican (who is coded "2") does not have "twice as much" party membership as a Democrat (who is coded "1"). Meanwhile, voting frequency—how many times each person has voted in the past five years—is a *ratio* scale, with equal units of measurement (every trip to the polls is counted once) and an absolute zero point (a score of 0 means that a person hasn't voted even once in the past five years).
4. The zip code strategy for creating regions is a *nominal* scale, reflecting only category membership. Regions with higher zip codes don't necessarily have "more" of anything, nor are they necessarily "better" in some respect.
5. Don't be distracted by the absolute zero point here (an income of \$0 means no money at all). The ranges of income are different in each group: Group A has a \$20,000 range, Group B has a \$30,000 range, Group C has a \$50,000 range, and Group D—well, who knows how much the richest person in the study makes each year? Because of the unequal measurement units, this is an *ordinal* scale.
6. This is an *ordinal* scale that reflects varying levels of quality. There is no indication that the four categories each reflect the same range of quality, and a true zero point (no road at all!) is not represented by the categorization scheme.
7. This is a tricky one. Despite the 0, this is *not* a ratio scale because virtually all students have at least a tiny amount of anxiety about tests, even if they respond "never" to all 25 questions. But the scale does involve an amount of something, so this must be either an ordinal or interval scale. Many psychologists would argue that the scores reflect an *interval* scale and would treat it as such in their statistical analyses. We authors don't necessarily agree, for two reasons. First, some of the statements on the instruments might reflect higher levels of test anxiety than others, so a "4" response to one item isn't necessarily the equivalent of a "4" response to another. Second, the 5-point rating scale embedded within the instrument ("never" to "always") doesn't necessarily reflect equal intervals of frequency; for instance, perhaps a student thinks of "sometimes" as being a broad range of frequencies of test-anxiety occurrence but thinks of "often" as being a more limited range. Thus, we argue that, in reality, the scores on the test anxiety instrument reflect an *ordinal* scale.

ANSWERS TO THE CONCEPTUAL ANALYSIS EXERCISE

"Identifying Problems with Validity and Reliability in Measurement" on page 93:

1. The test lacks *content validity*: It doesn't reflect the content domain that the instruction covered—actual tennis *skills*.
2. This is a problem of *equivalent forms reliability*: The different versions of the same instrument yield different results.
3. The issue here is *interrater reliability*: The researchers are evaluating the same behaviors differently.
4. The problem in this case is *criterion validity*: Two measures of energy level—blood-test results and observer ratings—yield very different results.
5. This is a problem of *test-retest reliability*: The students responded differently within a very short time interval (two weeks at most), even though their general opinions about climate change probably changed very little, if at all, during that interval.
6. The questionnaire lacks *internal consistency reliability*: Different items in the instrument yield different results, even though all items are intended to measure a single characteristic: matrimonial harmony.
7. In this case the instrument's *construct validity* is suspect. Religious tolerance is a hypothesized internal characteristic that can be inferred and measured only indirectly through observable patterns in people's behaviors. Here the behaviors being observed are simply responses to a questionnaire.
8. *Face validity* is at stake here: Although many psychologists contend that intelligence involves reaction time to some degree—and thus the reaction-time task might be a valid measure—on the surface the task doesn't *appear* to be a good measure. Although face validity is not "true" validity, a lack of it can sometimes negatively impact participants' cooperation in a research project.

Now go to MyEducationLab at www.myeducationlab.com to take assessments to evaluate your mastery of chapter concepts. Assignments and Activities exercises and Building Skills in Reading Research exercises are also available to help you master concepts and skills. Feedback for many of these exercises is provided so that you can see why your answers are correct or incorrect.

5

Writing the Research Proposal

Research is never a solo flight, an individual excursion. In today's world, researchers typically begin the research process by communicating their thoughts, objectives, plans, and methods for others to read, discuss, and act on. The mechanism that initiates such a dialogue is the research proposal. As a point of departure, it must be a precision instrument from the first word to the last.

Research is never a solitary activity. It involves many people and requires access to and use of resources far beyond one's own. For that reason, it must be carefully planned, laid out, inspected, and, in nearly every instance, approved by others. The graduate student conducting research for a thesis or dissertation must get the approval of an academic committee. A researcher seeking grant funding must get approval from the university or the organization for which he or she works, and the project must be deemed worthy of funding by the grant-awarding agency. Any researcher who plans to work with human subjects must get the approval of an internal review board, and one who plans to work with animals must get that of an institutional animal care and use committee (see Chapter 4). Such approvals are usually secured through the submission of a document known as a *research proposal*. The proposal lays out the problem for research, describes exactly how the research will be conducted, and outlines in precise detail the resources the researcher will use to achieve the desired results.

A proposal is as essential to successful research as an architect's plans are to the construction of a building. No one would start building a structure by rushing out to dig a hole for the foundation without knowing in detail how the building will look when it is finished. Before one turns a shovelful of earth, many questions must be addressed, many decisions made. What kind of building is desired? Is it a personal residence? If so, will it be a two-story colonial, a split-level, or a one-story ranch? How will the building be placed on the lot? Where will the doors be located? How will the roofs be pitched and shingled? What kind of heating system will be used? Each of these questions is important, and each should be addressed *before* a single pound of dirt is excavated, a single nail driven.

Even after all these decisions have been made, does the digging begin? Not at all! Further planning is necessary. The architect draws a floor plan of the entire structure, floor by floor, showing to the inch precisely where each room, closet, door, and window will be located. The architect also draws a series of elevations of the proposed building, showing each wall to scale as it will appear when completed. Finally, the architect draws up a set of specifications for the building, indicating exactly what lumber is to be used, how far apart certain beams will be placed, what kinds of doors and windows will be put in what locations, and all other details. Nothing is left to chance.

So is it now time to stake off the building's location and start digging for the foundation? Not so fast! Before the construction crew can do anything, they need *permission*. The contractor must get a building permit. Most communities have building codes that govern the kinds of buildings that can be constructed—codes regarding plumbing, wiring, distance from the street, and so on. A permit provides a means of ensuring that new buildings meet these codes. The point is this: *Permission is essential to the plan.*

Like the architect who presents detailed plans for a building, the researcher develops a written proposal for a research project. In this proposal, the problem and its attendant subproblems are clearly stated, hypotheses or questions are articulated, all necessary terms are defined, delimitations

are carefully spelled out, and the reason for conducting the study—why it is important—is explained. The researcher then specifies every anticipated detail of acquiring, organizing, analyzing, and interpreting the data. Parallel to the architect's specifications, the researcher sets forth the resources at hand for carrying out the research: his or her qualifications (and those of any assistants), the availability of the data, the means by which the data will be obtained, any needed equipment and facilities, and any other aspects of the total research process that merit explanation. Nothing is overlooked. All questions that may arise in the minds of those who review the proposal are anticipated and answered. Any unresolved matter is a weakness in the proposal and may seriously affect its approval.

We cannot overemphasize the importance of the proposal. It is the key that unlocks the door to the research endeavor. If it is not clearly and explicitly delineated, it may cause the whole project to be turned down.

Sometimes novice researchers think the proposal is merely a necessary formality and thus do not give it the serious consideration it deserves. They try to set forth the project in a few hastily written pages. Such an approach often fails. Those sponsoring a project, whether a graduate committee or a funding agency, realize that a researcher invests considerable time, effort, and (sometimes) money when doing research. Accordingly, no one should rush into a project without a clearly conceived goal and a thorough, objective evaluation of all aspects of the research endeavor.

No matter whether you are seeking funding for a project from a grant foundation or seeking approval for a thesis or dissertation from a university faculty committee, a clear, well-written proposal is essential. Nothing is a substitute for an explicit setting forth of both problem and procedure. Other names for a proposal are *prospectus*, *plan*, *outline*, *statement*, and *draft*. If you are asked to present any of these, you are being asked to present a research proposal.

Characteristics of a Proposal

Research demands that those who undertake it be able to think clearly, without confusion. The proposal will demonstrate—for better or for worse—whether the author possesses that ability. When one reads a proposal that is unfocused, poorly organized, and lacking in essential details, one gets the impression that the mind producing such a document is incapable of thinking precisely, logically, systematically, and thoroughly about the job that needs to be done. Unwelcome as the fact may be, the perceived qualifications of a researcher rest, more often than not, squarely on the quality of the proposal submitted.

Therefore, as you embark on the task of writing a research proposal, you should understand exactly what characteristics a proposal should have.

A Proposal Is a Straightforward Document

A proposal should not be cluttered with extraneous and irrelevant material. Right off the bat, it opens with a straightforward statement of the problem to be researched. It needs no explanatory props—no introduction, prologue, or statement telling why the researcher became interested in the problem or feels a burning desire to research it. Nor does it need explanations of why the researcher decided *not* to study certain other topics. Such information may be interesting, but none of it is necessary or appropriate. Those who will review your proposal are not interested in how-I-became-interested-in-my-topic excursions. Indeed, autobiographical excursions suggest that you cannot separate essentials from irrelevancies and so will neither enhance your stature as a researcher nor recommend you as one who can think in a clear and focused manner.

Whatever does not contribute directly to the delineation of the problem and its solution must be eliminated. Anything else is a distraction. Remember the architect's drawing: clean, clear, and economical. It contains all that is necessary, not one detail more.

Journalists are taught—or if not taught, soon learn—that the first words they write are the most important ones. You capture or lose your reader with the first sentence. It is impossible to overemphasize the importance of the beginning, whether it be of a speech, a news story, a journal article, or a proposal.

Imagine a proposal that begins with these words: “Five decades ago, the social and economic status of minority groups in the United States was . . .” A reader’s reaction might easily be: “Who cares, at this moment, what the social and economic status of minorities was 50 years ago? That’s history. What does the researcher *propose* to do *in the near future*? C’mon, what’s the problem? *State it!*”

You see, if your first sentence irritates your reader, you are immediately put at a disadvantage, and you have possibly sacrificed your reader’s interest. More significantly, the reader may infer that you cannot distinguish between history and future planning, and thus the reader may wonder about your ability as a researcher to think clearly and critically.

Keep in mind the meaning of *proposal*. The word suggests looking forward, to what the researcher *plans to do in the future*. If a writer intends to make an analytical comparison of the past and present social and economic conditions of minority groups, he or she might begin, “This study *will* analyze the social and economic status of certain minority groups today in comparison with their similar status five decades ago for the purpose of . . .” This is a no-nonsense beginning, and it indicates that the writer knows what a proposal should be.

A Proposal Is Not a Literary Production

An architect’s drawing is not a work of art. Similarly, a proposal is not a creative, “literary” production. The purpose of both is simply to communicate clearly. Just as an architect’s drawings present a plan for construction with economy of line and precision of measurement, a proposal describes a future research project with an economy of words and precision of expression.

The language must be clear, sharp, and precise. The proposal provides a chance to show with what clarity and detail the researcher can state a problem, delineate the collection of relevant data, and explain how those data will be interpreted and brought to bear on the research problem.

To those who have been taught that writing should be stylistically interesting and artistically creative, the preceding statements may come as a bit of a shock. Yet writing a superb proposal calls for skills of expression just as demanding as those needed for forging an unforgettable sentence.

A Proposal Is Clearly Organized

Proposals are written in conventional prose style, and thoughts are expressed in simple paragraph form. In professional writing, headings and subheadings are the single most commonly used strategy to express the writer’s overall organizational scheme. Examine your textbooks—also examine current articles in popular magazines—and you will discover how frequently headings are used to indicate the organizational structure of what has been written. You should communicate the outline of your thoughts to your own readers in the same explicit fashion.

If you are currently working on a master’s thesis or doctoral dissertation, your faculty advisor and committee may have a particular organizational scheme they want you to follow, possibly including certain chapter titles and within-chapter headings. Alternatively, if you are writing a grant proposal for a public or private funding agency, it is likely that the agency mandates that a proposal be divided into specific, labeled sections (e.g., “Research Objective,” “Relevant Literature,” “Proposed Method,” “Implications for Professional Practice”).

Organizing and Writing a Research Proposal

Proposals follow a simple, logical train of thought. Although there are conceivably many ways to arrange the various topics within a proposal, most proposals use similar formats, especially in quantitative studies. The following is an example of a format you might use in a proposal for a quantitative research study:

- I. The problem and its setting
 - A. The statement of the problem and subproblems
 - B. The hypotheses

- C. The delimitations
- D. The definitions of terms
- E. The assumptions
- F. The importance of the study
- II. The review of the related literature
- III. The data and the treatment of the data
 - A. The data needed and the means for obtaining the data
 - B. The research methodology
 - C. The specific treatment of the data for each subproblem
 - 1. Subproblem 1 (*The subproblem presented in Part I above is restated here.*)
 - a. The data needed to address the subproblem
 - b. The treatment of the data
 - 2. Subproblem 2 (*The same format for Subproblem 1 is followed here.*)
 - 3. *Additional subproblems are discussed in the same manner.*
- IV. The qualifications of the researcher and any assistants
- V. An outline of the proposed study (steps to be taken, timeline, etc.)
- VI. References
- VII. Appendixes

Proposals for qualitative studies sometimes use a slightly different format. The following format is an example of an outline you might use for a qualitative proposal:

- I. Introduction
 - A. General background for the study
 - B. Purpose of the study
 - C. Guiding questions
 - D. Delimitations and limitations
 - E. Significance of the study
- II. Methodology
 - A. Theoretical framework
 - B. Type of design and the assumptions that underlie it
 - C. Role of the researcher (including qualifications and assumptions)
 - D. Selection and description of the site and participants
 - E. Data collection strategies
 - F. Data analysis strategies
 - G. Methods of achieving validity
- III. Findings
 - A. Relationship to literature
 - B. Relationship to theory
 - C. Relationship to practice
- IV. Management plan, timeline, feasibility
- V. References
- VI. Appendixes

One rule governs the writing of proposals and final documents: *The arrangement of the material should be presented in such a manner that it forms for the reader a clear, progressive presentation.* It keeps items together that belong together—for example, the problem and its resultant subproblems, the subproblems and their corresponding hypotheses.

Formatting Headings and Subheadings

You must use different formats to indicate the different levels of headings you use. For example, if you have five different levels of headings, the American Psychological Association's *Publication Manual* (2010) specifies the following formats:

- **Level 1 headings**—the most important ones—are in **Boldface Uppercase and Lowercase Letters** and are **centered** on the page. These are headings of the largest units of writing;

for instance, they may be the titles of the various chapters in a proposal or research report. They correspond with Roman numerals I, II, III, and so on, in an outline.

- *Level 2 headings* are in **Boldface Uppercase and Lowercase Letters** that begin at the **left side** of the page—in other words, they are *flush left*. They correspond with the capital letters A, B, C, and so on, in an outline.
- *Level 3 headings* are in **Boldface first-letter-only-uppercase, ending with a period**. They are **indented** to the same degree that a paragraph is indented, and the first paragraph in the section follows on the same line. (Such headings are sometimes known as *run-in headings*.) They correspond with the numbers 1, 2, 3, and so on, in an outline.
- *Level 4 headings* are in ***Italicized boldface first-letter-only-uppercase, ending with a period***. They are placed, *indented*, at the beginning of the first paragraph in the section. They correspond with the lowercase letters a, b, c, and so on, in an outline.
- *Level 5 headings* are in ***Italicized non-boldface first-letter-only-uppercase, ending with a period***. They are placed, *indented*, at the beginning of the first paragraph in the section. They correspond with the numbers (1), (2), (3), and so on, that you sometimes see in an outline.

If you used this format, your various headings would look like this on the page:

First Level Heading

The first paragraph of this section begins here....

Second Level Heading

The first paragraph of this section begins here....

Third level heading. The first paragraph of this section begins here....

Fourth level heading. The first paragraph of this section begins here....

Fifth level heading. The first paragraph of this section begins here....

To help the headings stand out on the page, you may want to have an extra space (that is, an empty line) immediately preceding each one.

The format we suggest here is not the only one you might use. When choosing appropriate formats for your headings, you should check with any style manuals in your discipline and, if you are a student, with any graduate school requirements.

Above all, you should be *consistent* in how you format your headings. We have seen too many proposals, theses, and dissertations in which headings of equal importance appear in ALL CAPITALS and in Capitals and Lowercase, or perhaps they appear both

Centered

and

Flush Left.

Such inconsistency points to a careless, sloppy writer and, a proposal reviewer might think, perhaps an equally careless and sloppy researcher.

PRACTICAL APPLICATION Writing Your Proposal

Challenging as writing a proposal can be, especially for the beginning researcher, it isn't rocket science. Here we offer two sets of guidelines, one each for writing a first draft and for revising your proposal.

GUIDELINES Writing the First Draft

The following suggestions are based both on our own experiences as proposal writers and as faculty members who have advised numerous master's and doctoral students.

1. *Use a computer or other word processor.* Whether you begin writing your proposal on a word processor or on paper should depend on which medium allows you to think and write most easily. For example, if you have had considerable experience using a keyboard and can type as fast as you write, you will undoubtedly want to use a word processor from the get-go. In contrast, if you use the more tedious hunt-and-peck approach and have considerable trouble finding such letters as Q, X, and Z, you might want to start off with paper and pencil. At some point, however, you should put your first draft on a word processor to facilitate those inevitable revisions (there will be many!).

Early in the game, take the time to learn any special features of your word processing software that you will need for typing your proposal or your final research report. For example, learn how to insert tables, graphs, footnotes, and other specially formatted features. If you will be including words with accent marks or using punctuation marks different from those in English (e.g., *déjà*, *señor*, *Günter*, ζ), find out how to type them. If you will need to use certain symbols (e.g., α , Σ , π) or mathematical formulas, learn how to include them in your document.

2. *Adhere to any guidelines required by the institution, organization, or funding agency to which you are submitting the proposal.* If the group to which you are submitting the proposal requires that you (a) use certain headings, (b) follow a particular style manual, or (c) include certain information, do it! Blatant disregard for such guidelines is, for many proposal reviewers, a red flag that the researcher may not have his or her act together sufficiently to conduct the proposed research.

As their names imply, most style manuals also prescribe a certain writing style—for instance, whether to describe the researcher in first person (“I will conduct interviews . . .”) or third person (“The researcher will conduct interviews . . .”) and whether to use active voice (“The researcher will instruct participants to . . .”) or passive voice (“Participants will be instructed to . . .”) in describing procedures. Various academic disciplines have different style preferences, and you should not stray too far from the style typically used in your own field.

3. *When writing the first draft, focus more on organization and logical thought sequences than on exact wording, grammatical correctness, spelling, and nitty-gritty formatting details.* In Chapter 1 we mentioned that human beings can think about only a limited number of things at one time. All of the processes that skillful writing involves—organizing thoughts, following a logical sequence, expressing ideas clearly and succinctly, using acceptable grammar and punctuation, spelling words correctly, and so forth—may far exceed that capacity. In other words, you may very well *not* be able to do it all at once!

In the first draft, then, you should focus your attention on the big picture—that is, on presenting your ideas in a logical, organized, and coherent fashion. At this point, don't worry about picky details. If you can't think of the right word, leave a blank where it should go and move on. If you can't remember how a word is spelled, spell it in whatever way you can for the time being and then perhaps indicate your uncertainty by adding a question mark in parentheses. If you're not sure about where commas should go, either insert them or don't, and then check a style manual later on.

As you write, you may even discover that you are missing an important piece of information, perhaps something that you need to track down online or at the library. Never mind; let it go for now. Leave a blank and make a note of what you need. Chances are that you will need several such bits of information. You can track them all down later, *after* you have finished your first draft.

4. *Present the research problem at the beginning of the first chapter.* As we stated earlier in the chapter, *always* lead off with your research problem. The problem is at the very center of—and so drives—the entire project.

5. *Provide a context for your research problem.* A good proposal places the research problem within a specific context that helps the reader understand why the problem to be investigated *is*

a problem in need of solution. For example, perhaps the problem reflects an alarming state of affairs in our physical or social environment—say, an increase in sightings of frogs with birth defects or a high incidence of eating disorders in adolescent girls. Perhaps the problem involves inadequacies in an existing theory or conflicting findings in previous research. Perhaps the problem is a need to evaluate the effectiveness (or lack thereof) of a particular intervention—say, a new medical procedure or instructional method. Whatever form the context for the problem takes, it should be documented with appropriate citations of relevant research, theory, and other literature.

If you are writing a three-chapter proposal for a master's thesis or doctoral dissertation, you should include literature and citations that are key and central to your research problem near the beginning of the very first chapter. A more in-depth review of related literature should be presented a bit later, perhaps in the second chapter.

6. *Convince the reader of the importance of your project.* You must convince your readers that your planned research is not a trivial, meaningless undertaking—that, on the contrary, it can potentially make a substantial contribution to the body of human knowledge and may even, in some small way, help make the world a better place. Although you won't want to get emotional in your presentation, you nevertheless want to generate interest in what you are doing: You want your readers to *want to know* what your project's outcome will be.

As shown in the proposal outlines presented earlier in the chapter, researchers sometimes include a section specifically called "Importance of the Study," "Significance of the Study," or something of that ilk. In other cases researchers simply make a study's importance crystal clear within the introductory discussion of the overall context for the problem.

7. *Assume that your readers know nothing whatsoever about your proposed project.* Novice researchers often leave out critical pieces of information, assuming, somehow, that their readers are already aware of these things. (We have found this to be especially true for students who are writing a proposal for a faculty committee that already has some knowledge about the planned research.) Such omissions can lead to many misunderstandings along the way, and these misunderstandings may get you in trouble later on.

Your proposal is the mechanism through which you describe, in a permanent written form, what you intend to do from beginning to end. In this respect, it is very much like a contract to which you and your reviewers will ultimately agree. Accordingly, leave nothing unsaid, no question unanswered.

8. *Communicate that you have an open mind about what you will find.* Researchers often embark on research studies with the hope that they will uncover evidence in support of their hypotheses. But some novice researchers go too far and assert that they *will* find such evidence. Such statements as "I will show that" or "I will prove that" imply that the results of the study are already known. If the results can actually be predicted with 100% accuracy ahead of time, then what is the point of conducting the research? Truly objective, open-minded researchers place no bets in advance; they keep all options on the table. For instance, they might say "The purpose of this study is to determine *whether* . . ." or "The proposed research project is designed to investigate the *possible* effects of . . ."

9. *Describe your proposed methodology with as much detail and precision as possible.* The extent to which you can describe your methodology will depend to some degree on whether you are using a quantitative or qualitative approach. If you are using a quantitative approach, you will need to specify your sample, measurement instruments, and procedures in the utmost detail. If you are using a qualitative approach, you will probably be making some sampling and procedural decisions as the study proceeds. Nevertheless, at the proposal stage you should outline your sources of data and procedures as specifically as possible. Remember, the more information your reviewers have about your proposed project, the better position they will be in to determine its worth and potential contributions.

10. *If you intend to use data that already exist, describe where the data are located and how you plan to obtain them.* In some studies, and especially in historical research, a researcher may need

certain kinds of records. In such a situation, the researcher should know their exact location. Many novice researchers begin research projects by assuming that records are available but learn too late that either no records exist or the needed records are in an inaccessible location or under such heavy restriction that they are not available. Answer the question, Where are the data located? in no uncertain terms, and determine that you have access to them.

Suppose that the necessary data are letters written by an important historical figure and that they are in the possession of the person's family. You may know where the letters are located, but do you know how you will get them for your research purposes? Perhaps, in a case like this—or in any situation in which records are under other people's control—you might provide the names and addresses of the individuals who possess the data. You might also state that these custodians of the data have consented to your using the data for research purposes. Such details should be clearly stated in the proposal so that your sponsor, your academic committee, individuals at a funding agency, or whoever else is reading your proposal can feel confident that you will have ready access to the data you need.

11. *Describe how you will use the data to address your research problem.* Even though you have not yet collected your data, you will nevertheless need to describe how you intend to organize, analyze, and interpret them so that you can solve your research problem. Do not assume that others will know what you intend to do. Spelling out the treatment and interpretation of the data is a tedious, time-consuming process. But the alternative—attempting to relegate it to the broad sweep, the quick and easy statement, the careless approach—almost invariably courts disaster. Interpretation of the data is the step that gives meaning to the entire enterprise and makes it a genuine research endeavor, and it must therefore be planned and specified well in advance.

To see how some novice researchers fail to answer this most important question—How will the data be interpreted?—let's examine an excerpt from an economics doctoral student's proposal for a dissertation about labor relations. The excerpt appears in Figure 5.1. The student's main research problem is to "analyze the attitudes of professional employees toward certain aspects of management policy and to evaluate the relationship between these attitudes and the responsibility of management to articulate such policy for its employees." The student has organized his discussion of the data in terms of specific subproblems, describing both data collection and data interpretation with respect to each subproblem. In the excerpt, we see how the student says he will resolve the following subproblem:

What does an analysis of the attitudes of employees toward management policy for salary increases and merit pay reveal?

First read the student's restatement of the subproblem, his description of the data needed to resolve the problem, and his discussion of how he intends to secure those data; information about all of these issues appears under appropriate headings in Figure 5.1. Now, with such information in mind, read the section "How the Data Will Be Interpreted." What does the researcher really intend to do? Is he really going to *interpret* the data, to derive meaning from them? Is he going to "determine" anything *through an analysis of employee responses*?

Unfortunately, the student is not talking about interpreting the data. He is merely planning to *tabulate* and *graph* the data. He will rearrange them and present them in another form. The data remain almost as raw as they were originally in employees' responses to the questionnaire. The researcher also tells us that he will find two points of central tendency ("averages") for the data: the median and the mean. The median and mean of *what*? The frequencies? The percentages? Both? And *why* will he calculate the median and mean? What do these statistics tell us about "attitudes of employees toward certain aspects of management policy"? These are critical questions that should be answered *in the proposal*. In the student's proposal as it presently exists, there is no discussion of how the data relate to *attitudes of employees*, even though an understanding of these attitudes is central to resolving the subproblem.

What might the student do to interpret his data? After tabulating the data in the manner he describes, he might collapse the responses into two categories—or perhaps into a continuum of categories—that reflect either support of or opposition to management policies. Then he could carefully review each category to identify the characteristics of each. Were people who supported

FIGURE 5.1

Where is the interpretation of the data? An excerpt from a student's proposal

Restatement of Subproblem 1. The first subproblem is to determine through an analysis of employee responses the attitudes of employees toward certain aspects of management policy for salary increases and merit pay.

The Data Needed

The data needed to resolve this subproblem are those employee responses to questions concerning salary increases and merit pay.

Where the Data Are Located

The data are located in the employee responses to questions 3, 7, and 13 of the questionnaire, "Survey of Employee Attitudes Toward Management."

How the Data Will Be Secured

The data will be secured by accurately tabulating all of the responses of employees to the above questions on the questionnaire.

How the Data Will Be Interpreted

From the responses of the questions, a table will be constructed similar to the following structural model. It will indicate the employee attitudes, their frequency, and the percentages of these attitudes of the total attitude response to each question.

Attitude	Frequency	Percentage
Totals		

A graph will then be constructed to show which attitudes received the greatest number of reactions and which had the least number of reactions. The median and the mean will also be found for the total group as a basis for comparison.

management lukewarm in their support? What keywords did they use in their responses? What did the overall category responses indicate about the employees' attitudes?

Despite its obvious weakness, the excerpt in Figure 5.1 does illustrate one effective strategy for discussing the analysis and interpretation of the data. In particular, it can be quite helpful to *address each subproblem separately*. For each subproblem, you might:

- Restate the subproblem.
- Clearly identify the data that relate to the subproblem.
- Explain fully and unequivocally how you intend to analyze and interpret the data to resolve the subproblem.

More generally, *the plan for the treatment of the data should be so unequivocal and so specific that any other qualified person could carry out your research project solely by following your proposal*. Every contingency should be anticipated; every methodological problem should be resolved. The degree to which you delineate how the data will be interpreted will play a significant role in the success or failure of your research endeavor. The method of data interpretation is the key to research success, and it should be described with utmost care and precision.

12. *Use appendices to present informed consent letters, specific measurement instruments, and other detailed materials.* Although you need to describe your procedures precisely and completely, too much detail all in one place can interfere with the overall flow of your writing. Appendices provide an easy way to present any necessary details that are not central to the points you are trying

to make. Simply refer to each appendix as it is relevant to your discussion, perhaps like this: "To recruit participants, the nature of the study will be described, and volunteers will be asked to read and sign an informed consent letter (see Appendix A)." If you have more than one appendix, assign them letters that reflect the order in which you refer to them in the text: The first appendix you mention should be labeled "Appendix A," the second should be labeled "Appendix B," and so on.

GUIDELINES Revising Your Proposal

You must remember that your first draft will almost certainly *not* be your last one. This is true not only for research proposals and reports but for other major writing projects as well. Here we offer some suggestions for polishing your proposal into its final form.

1. *Set the proposal aside for a few days.* After writing your first draft, put it aside for a while so that, later, you can approach it with a fresh eye. If you reread it too soon, you will read it with what you *thought you had said* still fresh in your mind and so won't necessarily read what you *actually wrote*.

2. *Read a paper (rather than electronic) copy of your first draft.* As we mentioned in Chapter 1, paper copies often reveal problems with a text that somehow escape our attention on the computer screen. We aren't sure why this is, but we have repeatedly found it to be so.

Your proposal should, at this point at least, be double-spaced rather than single-spaced and have wide margins, leaving you lots of room for writing corrections and comments. You should expect that *you will write all over your first draft*. Figure 5.2 presents many commonly used editing marks for small-scale changes. For more significant changes (e.g., adding and moving text), you may want to use arrows, indicate pages where sentences or paragraphs should be moved to or from, and have blank sheets of paper or your computer nearby for major rewrites.

3. *Carefully scrutinize what you have written, looking for disorganized thoughts, illogical thinking, and inconsistencies in terminology.* Look for places where you move unpredictably from one topic to another, go off on unnecessary tangents, or draw unwarranted conclusions. Also, look at each paragraph under each one of your headings: All paragraphs under a particular heading should deal specifically with the topic that the heading identifies.

In addition, strive for consistency in your terminology. Novice researchers sometimes bounce from one label to another when talking about a single characteristic or variable, and the inconsistency can be confusing for readers. For instance, imagine that an educational researcher is investigating an aspect of human motivation—in particular the kinds of goals that students set for themselves as they study academic material. One such goal is the desire to truly learn and understand the topic of study—a goal that some motivation theorists call a *learning goal* and others call a *mastery goal*. In writing a research proposal, then, the researcher decides to appease both groups of theorists, using one term in some places and the other term in other places. Readers of the proposal are apt to be perplexed, perhaps thinking, "What's the difference between these two?" when there *isn't* a difference between them. Better to choose one term or the other and stick to it.

Consistency is important, too, when referring to the different groups of participants you might have in a study. For instance, imagine a medical researcher who wants to study the effects of a new pain reliever for alleviating the symptoms of chronic arthritis. The researcher plans for some arthritis sufferers to take the pain reliever and for others to take a sugar pill that, on the surface, looks identical to the pain reliever.¹ The researcher might call the first group *Group 1*, the

¹In such a study, the researcher must not deceive the sugar-pill recipients that they are getting a pain reliever. Such a deception would be a violation of basic ethical standards for research (see Chapter 4, especially the section "Voluntary and Informed Participation"). Instead, all participants in the study should be informed of the nature of the study: a comparison of the effects of a new pain reliever with those of a placebo. They do not need to be informed about which kind of pill they are taking, as such information might affect their subjective *perceptions* of pain and thereby distort the results of the study.

FIGURE 5.2

Commonly used editing marks

⧻ or ⧷ or ⧶	delete; take it out
⌵	close up; print as <u>one</u> word
^ or > or <	caret; insert <u>here something</u>
#	insert <u>aspace</u>
stet	let marked <u>text</u> stand as set
tr	transposes, change <u>order the</u>
/	used to separate two or more marks and often as a concluding stroke at the end of an insertion
¶	begin a new paragraph
Ⓟ	spell out (set <u>5 lbs.</u> as five pounds)
cap	set in <u>capitals</u> (CAPITALS)
lc	set in <u>Lowercase</u> (lowercase)
bf	set in <u>boldface</u> (boldface)
= or -/ or ⧄ or ⁄	hyphen
√	superscript or superior ($\sqrt{2}$ as in πr^2)
^	subscript or inferior (H_2O)
^	comma
√	apostrophe
o	period
;	semicolon
:	colon
↯ or ↮	quotation marks
⌈ ⌋	parentheses

Treatment Group, or the *Experimental Group*, and might call the second group *Group 2*, the *Placebo Group*, or the *Control Group*. The researcher should decide which terms to use and be consistent in using them. Moving back and forth willy-nilly among the different labels might lead a befuddled reader to conclude that the proposed research project will have six groups instead of two!

4. *Look for places where you are not as clear as you might be.* Ambiguous phrases and sentences—those with several possible meanings and those with no obvious meaning at all—significantly weaken the power of a research proposal. As an example, consider this excerpt from a literature review written by a former master's student:

It appears to be the case that not only is note taking superior for recall immediately after lecture, but that the entire memory storage and recall process is further strengthened as time goes on when one takes notes on a lecture. And, of course, this is generally how American college students are tested.

What does the student mean by the phrase “the entire memory storage and recall process is further strengthened”? And to what does the word *this* refer in the second sentence? Even though one of us authors is an educational psychologist who knows a great deal about both human memory processes and American testing procedures, neither of us has any idea what this student was trying to communicate.

5. *Keep your sentences simple and straightforward.* As a general rule, try to keep your sentences short. Vary the length, of course, but break up those long, contorted sentences into shorter, more

succinct ones. Be alert to how and where you use adjectives, adverbs, and other modifiers. Misplaced phrases and clauses can wreak havoc with the thoughts you want to communicate. As a simple, nonacademic example, consider this example of misplaced modification in a classified ad: "Piano for sale by a woman with beautifully carved mahogany legs that has arthritis and cannot play anymore." Move the prepositional phrase and add a comma, and the ad makes more sense: "FOR SALE: A piano with beautifully carved mahogany legs, by a woman who has arthritis and cannot play anymore."

6. *Choose your words carefully.* A thesaurus—perhaps a book, the "thesaurus" feature in your word processing software, or an online thesaurus (e.g., www.merriam-webster.com)—can help you find the exact word you need. Never use a long word where a short one will do. In a straightforward discussion, use one- or two-syllable words rather than longer ones. Use professional jargon only when you need it to relate your ideas to existing theories and literature in the discipline.

7. *Check carefully for errors in grammar, punctuation, spelling, and formatting.* Now is the time to attend to grammar, punctuation, spelling, correct heading formats, and other minor details. Ultimately you want your proposal to be, if not perfect, then as close to perfect as any human being can reasonably make it. Careless errors and other signs of sloppiness may suggest to your reviewers that the way you conduct your research project may be equally careless and sloppy.

Your word processing software will, of course, be helpful in this respect. For instance, *use the grammar checker*. Grammar checkers can search for word expressions, clichés, multiple negation, too many prepositional phrases, and other common problems. Some word processors even have a built-in function to measure the reading level of your writing; such information might be helpful in ensuring that you are writing at the appropriate level for your audience.

In addition, *use the spell checker, but don't rely on it exclusively*. As we pointed out in Chapter 1, a spell checker typically does nothing more than check each word to see if it is a "match" with a word in the English language or in some other list you specify. It will not tell you whether you have used the *right* words in every case. So even if you take advantage of the spell checker, always follow up by reading your document, word for word, to be sure that every word is spelled correctly. If you are a poor speller, then ask someone else—a good speller, obviously—to proof-read the entire document for errors.

You or your proofreader should be alert not only for spelling errors but also for the use of *homonyms*—sound-alike words that are spelled differently and have different meanings—in place of words you actually intended to use. Following are several commonly misused homonyms that we authors have often seen in research proposals and research reports:

it's versus its

it's is a contraction for "it is"

its is the possessive form of the pronoun *it*

there versus their versus they're

there is typically used either as (a) an adjective or adverb referring to a location or (b) an indefinite pronoun that begins a sentence (e.g., "There is no reason to . . .")

their is the possessive form of the pronoun *they*

they're is a contraction for "they are"

affect versus effect

affect as a *verb* means to have an influence on (e.g., "motivation affects learning")

affect as a *noun* is a synonym for emotion (e.g., "sadness and guilt are both forms of unpleasant affect")

effect as a *verb* means to bring something about (e.g., "to effect change")

effect as a *noun* means the causal result of an event (e.g., "rainfall has an effect on crop production")

The difference between *affect* and *effect* can be especially troublesome, in large part because *affect* as a verb and *effect* as a noun both involve an influence of some sort. But using the incorrect

word instead of its more appropriate homonym, especially when done frequently throughout the proposal, communicates to the reader that you have not taken the proposal-writing task seriously.

Speaking of *it's* versus *its*, we urge you to watch your use of apostrophes to indicate possessive nouns. In general, an apostrophe comes before the *s* in a singular noun (e.g., "a person's income level") but after the *s* in a plural noun (e.g., "companies' marketing strategies"). However, when a plural noun has no *s* unless it is possessive, the apostrophe comes before the *s* (e.g., "children's behaviors," "people's attitudes"). And when a singular noun ends in *s* even when it is *not* a possessive, you should add an apostrophe and an *s* (in that order) to indicate possession (e.g., "in Dr. Strauss's research").

8. *Make sure that items in bulleted lists are parallel in structure.* In writing a research proposal, a sequence of bullets or numbered items is often an effective way to present such things as the definitions of key terms and the major assumptions underlying a research project. All of the items in the sequence should have the same general grammatical structure—at a minimum by all being short phrases *or* by all being complete sentences. Mixing complete and incomplete sentences within a single list is frowned on. For example, in his dissertation proposal regarding cartographers and the Strong Vocational Interest Blank, which we previously mentioned in Chapter 3, Arthur Benton defined two of his terms as follows:

Cartographer. A cartographer is a professional employee who engages in the production of maps, including construction of projections, design, drafting (or scribing), and preparation through the negative stage for the reproduction of maps, charts, and related graphic materials.

Discrete interests. Discrete interests are those empirically derived qualities or traits common to an occupational population that serve to make them distinct from the general population or universe.

Notice how both definitions use complete sentences to describe the meanings of terms. Alternatively, the author might have used *incomplete* sentences, perhaps like the following:

Cartographer. A professional employee who engages in the production of maps, including construction of projections, design, drafting (or scribing), and preparation through the negative stage for the reproduction of maps, charts, and related graphic materials.

Discrete interests. Those empirically derived qualities or traits common to an occupational population that serve to make them distinct from the general population or universe.

Careful attention to such parallelism in form is yet another sign of a careful, meticulous researcher.

9. *Make sure there is a one-to-one correspondence between the citations in the text and the references in the reference list.* Every source you cite in the text or in footnotes should be included in more complete form in the proposal's reference list—no matter whether the source is a book, journal article, conference presentation, Internet website, or some other entity to which you are giving credit. Furthermore, every item in the reference list should be cited at least once in the text. The formats for citations and reference lists should be consistent with the style manual typically used in your particular academic discipline. The four most commonly used styles are listed in Table 12.1 in Chapter 12.

10. *Consider the feasibility of your project once again.* Now that you have laid everything out in the proposal, check one more time to be sure you have the time, resources, and energy to do everything you say you are going to do.

11. *Print out your second draft, and read your proposal carefully once again.* Look critically at each thought as it stands on paper. Do the words say exactly what you want them to say? Read carefully phrase by phrase. See whether one word will carry the burden of two or more. Throw out superfluous words.

12. *Seek the feedback of others.* We cannot stress this point enough. No matter how hard you try, you cannot be as objective as you would like to be when you read your own writing. Ask people to read and critique what you have written. Don't ask friends or relatives who are likely to give you a rubber stamp of approval. Instead, ask people who will read your work thoroughly, give you critical feedback, and make constructive suggestions.

One final comment: *Get used to writing.* Researchers write continuously—sometimes to communicate with others, at other times to facilitate their own thinking. Paper or a word processor can be effective for personal brainstorming sessions. Take time to sit back and use pencil or keyboard to help you clarify your thoughts and ideas.

PRACTICAL APPLICATION Strengthening Your Proposal

Not all research proposals get approved, of course. This is certainly the case for many proposals requesting funding from a private or government agency. But some proposals to conduct non-funded research get turned down as well, usually for one or more very good reasons. In Figure 5.3 we list shortcomings that experienced proposal reviewers have often observed. Proposals submitted by students for academic research projects (e.g., for theses and dissertations) tend to have a number of these shortcomings.

Once you have written what you believe to be your final proposal, you should scrutinize it one more time, preferably after you have set it aside for a few more days. *Take a critical approach, looking for what's wrong rather than what's right.* The following checklist can provide guidance about what to look for.



CHECKLIST

Features Detracting from Proposal Effectiveness

Check each item to be sure that your proposal exhibits *none* of the following characteristics:

FOR ANY RESEARCH PROPOSAL

- _____ 1. The statement of the problem is vague, or it is so obscured by discussions of other topics that it is impossible to find.
- _____ 2. The methodology is incompletely described; an explanation of exactly how the research will be conducted is not specifically delineated.
- _____ 3. The proposed treatment of each subproblem is general and cursory; it does not convey clearly how the data will be used and interpreted to resolve the subproblem or the overall research problem.
- _____ 4. Criteria for the admissibility of the data are weak or nonexistent.
- _____ 5. The proposal lacks sharpness. It is not logically organized. Without clear divisions that set forth the areas of the research project, it rambles. The reader has difficulty isolating the discussion of the problem, the subproblems, the related studies, the methodology, the interpretation of the data, and other important parts of the proposal.
- _____ 6. The proposal is phrased in terms that are too general, ambiguous, or imprecise to be useful for evaluation. Such phrases as "tests will be given" and "measurements will be taken" are largely meaningless.
- _____ 7. The format of the proposal deviates from the guidelines set forth by the approval group or funding agency.
- _____ 8. Some cited sources do not appear in the reference list; alternatively, the sources are incompletely or incorrectly cited.

FOR A PROPOSAL TO A FUNDING AGENCY

- _____ 9. The problem does not address the research area outlined by the funding agency.
- _____ 10. The proposal is too ambitious for the grant money available.
- _____ 11. Items included in the budget are disallowed by the terms of the grant.
- _____ 12. A clear and explicit budget statement outlining program expenditures is lacking, or the summary of estimated costs is ambiguous and indefinite.
- _____ 13. The section of the proposal explaining the study's importance is not set forth clearly enough for the funding agency to see a relationship of the study to the purpose for which the grant is awarded.

FIGURE 5.3

Common weaknesses in research proposals

Weaknesses Related to the Research Problem:

- The description of the project is so nebulous and unfocused that the purpose of the research is unclear.
- The problem as presented is not empirically testable.
- The problem is not framed within an appropriate context.
- The problem is unimportant or unlikely to yield new information.
- The hypothesis is ill-defined, doubtful, or unsound; it is based on insufficient evidence or illogical reasoning.
- The problem is more complex than the investigator realizes.
- The problem is of interest only to a particular, localized group, or in some other way has limited relevance to the field as a whole.

Weaknesses Related to the Research Design and Methodology:

- The description of the design and/or method is so vague and unfocused as to prevent adequate evaluation of its worth.
- The proposed methodology violates basic ethical standards.
- The data the investigator wishes to use are either difficult to obtain or inappropriate for the research problem.
- The proposed methods, measurement instruments, or procedures are inappropriate for the research problem (e.g., proposed measurement instruments may have poor reliability and validity).
- Appropriate controls are either lacking or inadequate.
- The equipment to be used is outdated or inappropriate.
- The statistical analysis has not received adequate consideration, is too simplistic, or is unlikely to yield accurate and clear-cut results.

Weaknesses Related to the Investigator:

- The investigator does not have sufficient training or experience for the proposed research.
- The investigator appears to be unfamiliar with important literature relevant to the research problem.
- The investigator has insufficient time to devote to the project.

Weaknesses Related to Resources:

- The institutional setting is inadequate or unsuitable for the proposed research.
- Proposed use of equipment, support staff, or other resources is unrealistic.

Weaknesses Related to the Quality of Writing:

- The proposal does not stay focused on the research problem; it rambles unpredictably.
- The proposal inadequately or incorrectly cites related literature.
- The proposal does not adhere to the appropriate style manual.
- The proposal has grammatical and/or spelling errors.

Sources: Based on Allen, 1960; Cuca & McLoughlin, 1987; Davitz & Davitz, 1996; Wong, n.d.

Final Thoughts about Proposal Writing

Possibly the most challenging part of a proposal, at least for novice researchers, is the discussion of how the data will be handled. Spelling out every minute detail of planned data analyses and supporting one's choices with solid rationales can be tedious and time-consuming. All too often, novice proposal writers leave out critical details about data analysis procedures, assuming or hoping that readers will mentally fill in reasonable analysis procedures. But in a good proposal, you cannot ask your reader to make such a journey of faith.

When drawing up a contract, an attorney meticulously includes all of the rights and obligations of the parties included in the contract. The proposal writer should prepare a proposal with the same precision. In a sense, a proposal is, under certain circumstances, a form of contract, or what we might call a *quasi* contract.

Are you submitting a proposal for a grant to underwrite a research project? If so, you (as the party of the first part) are proposing to undertake a research project in exchange for a monetary consideration from the agency providing the grant (the party of the second part). Regarded from a legal standpoint, your proposal, on acceptance by the granting agency, is a formal contractual relationship.

Now let's look at the situation from an academic standpoint. Certainly there are differences between a proposal presented to a funding agency and a proposal presented by a student to an academic advisor. Yet in another way the two kinds of proposals are very similar: In both cases, the basic elements of the research problem, the methodology, the data, and any other factors critical to conducting the inquiry must be clearly set forth and mutually agreed on before the actual research activity can begin.

Most faculty advisors will want to review the proposal periodically as it is being developed; they will also want to monitor your progress as you proceed with your study. Such a process of ongoing guidance from an experienced professional and researcher is to be welcomed, not avoided. It is perhaps the single best way you can learn the tricks of the research trade.

Proposals are mandatory for many academic research projects. For instance, any thesis or dissertation must begin with a proposal, and any project involving human subjects must get IRB approval before it ever gets off the ground. But even when a proposal is not mandatory, it is always *advisable*, regardless of the magnitude of the project or its academic sophistication. From a student's perspective, a proposal has two distinct advantages:

1. It helps the student organize the research activity.
2. It communicates to the student's advisor what the student intends to do, thereby enabling the advisor to provide counsel and guidance in areas that may pose exceptional difficulty.

A proposal for any research endeavor merits words that are carefully chosen, a style that is clear and concise, attention to the most minute procedural detail, and for each procedure, a rationale that is logically and clearly stated. All of this is a tall order, but the result reveals the scholarship of the proposal author as perhaps no other academic assignment can ever do.

It is an awesome fact to contemplate, but to no small degree *your proposal is you!* It defines your ability to think critically and to express your thoughts clearly. *It is the practical application of your research competence laid bare on a sheet of paper.*

A Sample Research Proposal

We conclude this chapter by presenting an example of what a research proposal—in this case, a proposal for a doctoral dissertation at the University of Northern Colorado—might look like. The author, Rosenna Bakari, uses the very first paragraph of the proposal to present the research problem clearly and concisely:

Attitudes that teachers bring into the classroom are a critical factor in the academic failure of African American students (Irvine, 1990). Preliminary research suggests that many in-service and prospective teachers do not hold positive attitudes toward teaching African American students (Irvine, 1990). As a result, many researchers see attitudes and values clarification of preservice teachers concerning race as a critical aspect of multicultural teacher education (Gay, 1994; Wiggins & Follo, 1999; Zeichner, 1996). However, there are no adequate instruments available to measure preservice teachers' attitudes about teaching African American students. Hence, the intent of this research is to develop and validate an instrument to measure preservice teachers' attitudes toward teaching African American students. (p. 1)

We now fast-forward to Bakari's methodology section. As we have in previous chapters, we present the proposal itself on the left and add our observations on the right.

Dissertation ANALYSIS 2

METHODOLOGY

This study is intended to develop and validate a survey instrument that assesses preservice teachers' attitudes toward teaching African American students. The survey instrument was developed based on educational recommendations and research literature indicating that culture is an important consideration in educating African American students effectively. Two pilot studies were conducted as preliminary investigations. This chapter will summarize the pilot studies and discuss the methodology of the current study.

[The student describes the two pilot studies she conducted previously relative to her present study. We pick the proposal up later, as she describes her proposed sample, measurement instruments, data collection, and data analysis.]

Sample

Three sub-groups will be solicited for participation. The first group will represent institutions where preservice teachers have little exposure to African American issues in education. The majority of participants are expected to be White and have little exposure to African American populations.

In the second group, preservice teachers will be solicited from teacher education programs that have more program goals or objectives related to teaching African American students. For example, diversity courses may be a requirement for graduation. In addition, preservice teachers are likely to have greater exposure to African American student populations during student teaching, in their university courses, or in their living communities than group one. However, the majority of participants are still expected to be White.

The third group of preservice teachers will be solicited from historically Black colleges or universities (HBCUs). Although HBCUs may differ in many respects, their focus is a "commitment, dedication, and determination to enhance the quality of life for

Comments

The author begins by reminding the reader of the purpose of the proposed research. The repetition of the research problem at this point, though not essential, is helpful to the reader, who can then put the procedures that follow into proper perspective.

The first paragraph is an advance organizer for the reader, who then can follow the author's subsequent train of thought more easily.

Earlier in the proposal the author presented her rationale for giving the instrument to three different groups. She predicted that the three groups would, on average, respond differently to the instrument, thereby providing evidence for the validity of the instrument.

Although the author is expecting the three groups to have different proportions of students from different racial groups, she will nevertheless seek information in support of her prediction through a demographic information sheet that she describes later in her proposal.

Always spell out what an abbreviation stands for before using it. For instance, here

African Americans" (Duhon-Sells, Peoples, Moore, & Page, 1996, p. 795). The majority of participants from this group are expected to be African American.

A minimum of 100 students will be solicited from each group. Sample size is critical because it provides a basis for the estimation of sampling error (Hair, Anderson, Tatham & Black, 1995). A sample size of at least 100 is recommended to conduct a confirmatory factor analysis because a sample less than 100 may not provide enough statistical power to reject the null hypothesis. A small sample could lead to acceptance of a model which is not necessarily a good fit, simply because there was not enough statistical power to reject the model. On the other hand, if the sample is too large, the model may be rejected due to sensitivity in detecting small differences, because the larger the sample, the more sensitive the test is to detecting differences (Hair, Anderson, Tatham & Black, 1995). Hair, Anderson, Tatham, and Black (1995) recommend a sample size between 100 and 200.

In order to achieve the minimum participant requirement for each group, involvement from more than one university may be necessary. For instance, four universities may represent group one while two universities may represent group two. This flexibility is important due to the variability in size of teacher education programs. Moreover, the reliance on instructors' willingness to contribute class time to this research may minimize the number of participants. All participants will be undergraduate or graduate preservice teachers and enrolled in a required course for a teacher preparation program. Graduate students must also be in pursuit of initial teacher certification. Students may be in any phase of their teacher preparation program. Preservice teachers who are not currently enrolled in any of the classes where the instrument will be distributed will not be selected for participation. Further, only students who are in attendance on the day the instrument is distributed will be selected for participation. For those students solicited to participate, participation will be voluntary, and anonymous.

Instrumentation

Four instruments will be employed for data collection in this research. They include the demographic data sheet, Teaching African American Students Survey, Responding Desirably on Attitudes and Opinions measurement (RD-16), and General Attitudes toward Teaching Survey. The demographic data sheet and the Teaching African American Students Survey (TAASS) were both designed by the researcher for this particular study. The General Attitudes toward Teaching Survey is an adaptation from a published Teacher Efficacy Scale and the TAASS. The RD-16 is a published instrument designed to measure social desirability. A description of the four instruments follows.

[Under four separate subheadings, the author then describes each instrument in detail, including the specific items that each one includes and any known information about validity and reliability.]

Data Collection

Participants will be contacted in their classes, where the instructor has agreed to allow class time to conduct this research. Participants will be told that the objective of the research is to gather information about future teachers, particularly, who they are

the author refers to "historically Black colleges or universities" and identifies the abbreviation HBCU in parentheses. She can then use "HBCU" in future discussions and her readers will know to what she is referring.

Here the author provides a justification for her sample size. We discuss the issue of statistical power in Chapter 11; at that point, we also revisit the concept of a null hypothesis.

The author explains why she is drawing her sample from several universities. It appears that she is predicting, and then answering, the kinds of questions the reader might have about her method.

The author gives enough information about her sample to enable any qualified reader to conduct the study she proposes. In addition, by describing the nature of her sample, she provides information about the population to which her study's results could reasonably be generalized.

Once again we see an advance organizer for the discussion that follows.

The author uses abbreviations (TAASS and RD-16) for two of her instruments. To be consistent, she should probably introduce them both in the second sentence of the paragraph, rather than leave TAASS for the third sentence as she does here.

The heading "Procedure" is more commonly used than "Data Collection" in human subjects research, but the latter is acceptable as well.

(demographics) and what they believe about teaching. To avoid a social desirability response set, participants will not be informed about the specific focus of the study (developing and validating an instrument to measure preservice teachers' attitudes toward teaching African American students). A statement will be read aloud to the class that informs students of their right to refuse to participate without any negative consequences, as well as the possibility of being requested to participate in follow-up research (test-retest reliability for the TAASS).

Requests for names and student identification numbers will be prohibited as any part of the data collection. However, participants will be asked to create identifications for themselves that cannot be traced to them by others. Pseudo-identification is necessary for students to remain anonymous, yet allows the researcher to conduct a retest for reliability measures. Examples of anonymous identifications will be given, such as a favorite noun, verb, or adjective (chair, jump, lazy). Students will be duly cautioned about selecting identifications that can be traced to them, such as mothers' maiden names, any part of their social security numbers, or nicknames.

Individual surveys will not be seen by anyone other than the participant once they are completed. Students will be requested to place their completed surveys in a designated "return" envelope. The last student to return the surveys will be requested to seal the return envelope. Only when the last survey is placed in the return envelope, and the envelope is sealed, will the administrator be permitted to handle the materials.

In classes where the researcher does not administer the instruments, research packets will be prepared for the person who does. Research packets will contain a disclosure statement to be read aloud to the participants. In addition to the disclosure sheet, the packets will include a demographic information sheet, the TAASS, and only one of the validity criteria instruments. Half of the participants will receive the RD-16 in their packet, and the other half will receive the General Attitudes Scale. The order of the instruments will also vary in the packets, with the exception of the demographic data sheet. The demographic data sheet will always appear last. Administrators will be instructed to avoid interpreting items on any of the three survey instruments. If students ask for interpretation of any items on the surveys, administrators will be instructed to respond, "Please use your best interpretation to answer all the items." However, clarifications may be made about the demographic information, if requested.

Three weeks after the initial research data have been collected, classes will be selected (based on availability) for retest of the TAASS. Participants will be solicited in a minimum of three classes that participated in the initial research. Only the TAASS will be administered for the retest. Students will be required to use the pseudo-identification selected in the initial research.

Data Analysis

LISREL and SPSS statistical software will be used for all analyses. As Hair, Anderson, Tatham, and Black, (1995) point out, there is no method of dealing with missing data that is free of disadvantages. Anytime missing data are imputed there is a risk of biasing the results (e.g., distributions, or correlation). Even the option of using the complete case approach has disadvantages. When only completed data are used, there is a

Here the author describes her procedures regarding informed consent.

Here she describes the steps she will take to ensure participants' right to privacy.

Notice how the author uses future tense to describe her proposed methodology. Later, when she rewrites the methodology section for her final research report, she will, of course, change her description of procedures to past tense.

The author will vary the order in which participants respond to the instrument, presumably as a way of determining whether taking one instrument affects how a participant responds to the instruments that follow.

The author is taking steps to increase the reliability of the instrument by standardizing its administration.

The author will administer the TAASS to some participants twice so that she can determine its test-retest reliability.

Notice that the author describes her proposed methods of data analysis as well as her methods of data collection. By doing so, she helps the reader determine whether her analyses will be appropriate for her research questions.

risk of reducing the sample size to an inappropriate number. Moreover, the results may no longer be generalizable to the intended population if the missing data are systematized rather than randomized (Hair, Anderson, Tatham & Black, 1995). Before any approach will be decided as to how to handle missing data, the missing data will be reviewed for systematic avoidance of response.

[The author then describes the specific analyses she plans to conduct and how they relate to her research problem.]

NOTE: Excerpt is from a research proposal submitted by Rosenna Bakari to the University of Northern Colorado, Greeley, in partial fulfillment of the requirement for the degree of Doctor of Philosophy. Reprinted with permission.

Notice, too, that the author will consider the nature of the data before and during her data analyses.

A book by Hair, Anderson, Tatham, and Black (1995) is cited several times in the methodology section. To be consistent with APA style (which she adheres to in her proposal), the author should list all four authors only for the first citation; after that, she can shorten the citation to "Hair et al. (1995)."

For Further Reading

- Bernstein, T. M. (1977). *The careful writer: A modern guide to English usage*. New York: Atheneum.
- Davitz, J. R., & Davitz, L. L. (1996). *Evaluating research proposals: A guide for the behavioral sciences*. Upper Saddle River, NJ: Prentice Hall.
- DeBakey, L. (1976). The persuasive proposal. *Journal of Technical Writing and Communication*, 6(1), 5–25.
- Johnson-Sheehan, R. (2008). *Writing proposals* (2nd ed.). New York: Longman.
- Krathwohl, D. R., & Smith, N. L. (2005). *How to prepare a dissertation proposal: Suggestions for students in education and the social and behavioral sciences*. Syracuse, NY: Syracuse University Press.
- Miner, L. E., & Miner, J. T. (2008). *Proposal planning and writing* (4th ed.). Westport, CT: Greenwood.
- Ogden, T. E., & Goldberg, I. A. (Eds.). (2002). *Research proposals: A guide to success* (3rd ed.). San Diego, CA: Academic Press.
- Rudestam, K. E., & Newton, R. R. (2007). *Surviving your dissertation: A comprehensive guide to content and process* (3rd ed.). Thousand Oaks, CA: Sage.
- Smith, M. C., & Carney, R. N. (1999). Strategies for writing successful AERA proposals. *Educational Researcher*, 28(1), 42–45, 58.

Now go to MyEducationLab at www.myeducationlab.com to take assessments to evaluate your mastery of chapter concepts. Assignments and Activities exercises and Building Skills in Reading Research exercises are also available to help you master concepts and skills. Feedback for many of these exercises is provided so that you can see why your answers are correct or incorrect.

6

Qualitative Research

To answer some research questions, we cannot skim across the surface. We must dig deep to get a complete understanding of the phenomenon we are studying. In qualitative research, we do indeed dig deep: We collect numerous forms of data and examine them from various angles to construct a rich and meaningful picture of a complex, multifaceted situation.

The term **qualitative research** encompasses several approaches to research that are in some respects quite different from one another. Yet all qualitative approaches have two things in common. First, they focus on phenomena that occur in natural settings—that is, in the “real world.” And second, they involve capturing and studying the complexity of those phenomena. Qualitative researchers rarely try to simplify what they observe. Instead, they recognize that the issue they are studying has many dimensions and layers, and they try to portray it in its multifaceted form.

As noted in Chapter 1, most researchers strive for objectivity in their research. They believe their observations should be influenced as little as possible by any perceptions, impressions, and biases they may have. By maintaining objectivity, they hope to maximize their chances of determining the ultimate Truth of which we spoke in Chapter 4. In contrast, some qualitative researchers—certainly not all of them—argue against this objectivity principle. Although objective methods may be appropriate for studying events in the physical world, these researchers say, an objective approach to studying human events—interpersonal relationships, social structures, creative products, and so on—is neither desirable nor, perhaps, even possible (e.g., Creswell, 2009; Eisner, 1998; Wolcott, 1994). From this perspective, the researcher’s ability to interpret and make sense of what he or she observes is critical for understanding any social phenomenon. In a sense, *the researcher is an instrument* in much the same way that an oscilloscope, sociogram, or rating scale is an instrument.

Furthermore, some qualitative researchers believe that there isn’t necessarily a single, ultimate Truth to be discovered. Instead, there may be multiple perspectives held by different individuals, with each of these perspectives potentially having equal validity, or truth (Creswell, 2007; Guba & Lincoln, 1988). One goal of a qualitative study, then, might be to reveal the nature of these multiple perspectives.

Qualitative research can be found in many academic disciplines, including anthropology, sociology, psychology, biology, history, political science, education, and medicine. In some disciplines, such as psychology and education, qualitative approaches were once frowned on (largely because of their subjective nature) and have only recently gained wide acceptance as legitimate research. Yet we should hardly think of qualitative research as being “new” or “modern.” In fact, many researchers believe that all inquiry starts out in a qualitative form (e.g., Lauer & Asher, 1988). When little information exists on a topic, when variables are unknown, when a relevant theory base is inadequate or missing, a qualitative study can help define what is important—that is, *what needs to be studied*. The field of medicine, as one example, makes extensive use of qualitative methods when unique or puzzling cases are first observed. Biologists’ efforts to classify newly observed species, create taxonomies, and describe the social behaviors of primates and certain other animal species are largely qualitative efforts. Many analyses of historical data are almost entirely qualitative. And social scientists often look subjectively for patterns in the complex phenomena they observe, sometimes using qualitative methods exclusively and sometimes combining qualitative and quantitative methods into a *mixed-methods design* (more on mixed-methods research in Chapter 10).

In this chapter we give you a general idea of what qualitative research is and what it strives to accomplish, with a particular focus on studies of human beings and their creations. Included in the chapter are descriptions of five kinds of qualitative studies: case studies, ethnographies, phenomenological studies, grounded theory studies, and content analyses. We describe a sixth kind, historical research, in Chapter 7.

The Nature of the Research Problem and Planning in Qualitative Research

In Chapter 2 we emphasized the importance of pinning down the problem with utmost precision. But here, too, we sometimes find an exception in qualitative research. Qualitative researchers often formulate only general research problems and ask only general questions about the phenomenon they are studying. For example, they might ask, What is the nature of the culture of people living in Samoa? or What is it like to live with someone who has Alzheimer's disease? Such research problems and questions do not remain so loosely defined, however. As a study proceeds, the qualitative researcher gains increasing understanding of the phenomenon under investigation and so becomes increasingly able to ask more specific questions—and occasionally can formulate and test specific hypotheses as well.

Because qualitative researchers tend to ask open-ended questions at the beginning of an investigation, they sometimes have difficulty identifying—at the outset—the exact methods they will use. Initially, they may select only a general approach suitable for their purpose, perhaps choosing a case study, ethnography, or content analysis. As they learn more about what they are studying and can thereby ask more specific questions, so, too, can they better specify what methods they should use to answer those questions.

The methodology in a qualitative study, then, may continue to evolve over the course of the investigation. Despite this fact, we must emphasize that *qualitative research requires considerable preparation and planning*. The researcher must be well trained in observation techniques, interview strategies, and whatever other data collection methods are likely to be necessary to address the research problem. The researcher must have a firm grasp of previous research related to the problem so that he or she knows what to look for and can separate important information from unimportant details in what he or she observes (some grounded theory studies are exceptions, for reasons you will discover shortly). And the researcher must be adept at wading through huge amounts of data and finding a meaningful order in what, to someone else, might appear to be chaos. For these reasons, a qualitative study can be a challenging task indeed. It is definitely *not* the approach to take if you're looking for quick results and easy answers.

When to Choose a Qualitative Approach

Qualitative research studies typically serve one or more of the following purposes (Peshkin, 1993):

- *Description.* They can reveal the multifaceted nature of certain situations, settings, processes, relationships, systems, or people.
- *Interpretation.* They enable a researcher to (a) gain new insights about a particular phenomenon, (b) develop new concepts or theoretical perspectives about the phenomenon, and/or (c) discover problems that exist within the phenomenon.
- *Verification.* They allow a researcher to test the validity of certain assumptions, claims, theories, or generalizations within real-world contexts.
- *Evaluation.* They provide a means through which a researcher can judge the effectiveness of particular policies, practices, or innovations.

As a general rule, qualitative studies do *not* allow the researcher to identify cause-and-effect relationships—to answer questions such as What caused what? or Why did such-and-such happen? You will need quantitative research, especially experimental studies, to answer questions of this kind.

Qualitative Research Designs

In this section, we describe five common qualitative research designs. We give you enough information to help you determine whether one of these approaches might be suitable for your research question, and we briefly describe the specific nature of the method, data analysis, and research report for each design. Later in the chapter, we discuss data collection strategies and data analysis strategies that are more broadly applicable to qualitative designs.

A single chapter cannot cover everything you would need to know to carry out a solid qualitative research project. Should you choose to conduct a qualitative study, we urge you to take advantage of the resources listed in the “For Further Reading” section at the end of the chapter.

Remember, too, that of all the designs we describe in this book, qualitative research methods are the least prescriptive (Eisner, 1998). There are no magic formulas, no cookbook recipes for conducting a qualitative study. This book, as well as any others you may read, can give you only general guidelines based on the experiences of those qualitative researchers who have gone before you. In a qualitative study, the specific methods you use will ultimately be constrained only by the limits of your imagination.

Case Study

In a **case study**—sometimes called *idiographic research*—a particular individual, program, or event is studied in depth for a defined period of time. For example, a medical researcher might study the nature, course, and treatment of a rare illness for a particular patient. An educator might study and analyze the instructional strategies that a master teacher uses to teach high school history. A political scientist might study the origins and development of a politician’s campaign as he or she runs for public office. Case studies are common not only in medicine, education, and political science, but also in law, psychology, sociology, and anthropology.

Sometimes researchers focus on a single case, perhaps because its unique or exceptional qualities can promote understanding or inform practice for similar situations. At other times researchers study two or more cases—often cases that are either similar or different in certain key ways—to make comparisons, build theory, or propose generalizations; such an approach is called a *multiple* or *collective* case study.

A case study may be especially suitable for learning more about a little known or poorly understood situation. It can also be appropriate for investigating how an individual or program changes over time, perhaps as the result of certain conditions or interventions. In either circumstance, it is useful for generating or providing preliminary support for hypotheses. Its major weakness is that, especially when only a single case is involved, we cannot be sure that the findings are generalizable to other situations.

Method. In a case study, the researcher collects extensive data on the individual(s), program(s), or event(s) on which the investigation is focused. These data often include observations, interviews, documents (e.g., newspaper articles), past records (e.g., previous test scores), and audiovisual materials (e.g., photographs, videotapes, audiotapes). In many case studies, the researcher may spend an extended period of time on site and interact regularly with the person or people being studied.

The researcher also records details about the context surrounding the case, including information about the physical environment and any historical, economic, and social factors that have bearing on the situation. By identifying the context of the case, the researcher helps others who later read the research report to draw conclusions about the extent to which the study’s findings might be generalizable to other situations.

Data analysis. Data analysis in a case study typically involves the following steps (Creswell, 2007; Stake, 1995):

1. *Organization of details about the case.* The specific “facts” about the case are arranged in a logical (e.g., chronological) order.

2. *Categorization of data.* Categories are identified to help cluster the data into meaningful groups. (For instance, a researcher studying the course of a political campaign might think in terms of “campaign strategies,” “fund-raising activities,” “news media accounts,” “setbacks,” etc.)
3. *Interpretation of single instances.* Specific documents, occurrences, and other bits of data are examined for the specific meanings that they might have in relation to the case.
4. *Identification of patterns.* The data and their interpretations are scrutinized for underlying themes and other patterns that characterize the case more broadly than a single piece of information can reveal.
5. *Synthesis and generalizations.* An overall portrait of the case is constructed. Conclusions are drawn that may have implications beyond the specific case that has been studied.

Especially when only a single case is studied, any generalizations made are, of course, tentative and must await further support from other studies—perhaps from additional case studies, other kinds of qualitative studies, or experimental research.

A case study researcher often begins the data analysis process during data collection; preliminary conclusions are likely to influence the kinds of data he or she seeks out and collects in later parts of the study. Ultimately the researcher must look for convergence (*triangulation*) of the data: Many separate pieces of information must all point to the same conclusion.

The research report. If you conduct a case study, you will probably want to include the following in your report:

1. *A rationale for studying the case.* Explain why the case was worthy of in-depth study—in other words, how it will significantly advance human beings’ knowledge about some aspect of the world.
2. *A detailed description of the facts related to the case.* Describe the specific individual(s), program(s), or event(s) you studied, as well as the setting and any other uncontested facts about the case. Your description should be as thorough and objective as possible.
3. *A description of the data you collected.* Tell your readers what observations you made, whom you interviewed, what documents you examined, and so on.
4. *A discussion of the patterns you found.* Describe any trends, themes, personality characteristics, and so on, that the data suggest. At this point, you are going beyond the facts themselves to your *interpretation* of the facts. Support each pattern you identify with sufficient evidence to convince the reader that the pattern does, in fact, accurately portray the data. If some data contradict the patterns you propose, however, you should describe those as well. *Even though you are interpreting as well as reporting data, you want to present as complete and unbiased an account of the case as you possibly can.*
5. *A connection to the larger scheme of things.* In some way, you need to answer the question, So what? In what way does the case study contribute to our collective knowledge about some aspect of the world or human experience? The connection(s) you make here might take one or more of several forms. You might compare the case with other, previously reported cases and note similarities and dissimilarities. You might argue that the case either supports or disconfirms an existing hypothesis or theory. Or you might use the case to support your contention that a particular intervention—perhaps a medical treatment, teaching method, or campaign strategy—can be a highly effective one.

Ethnography

In a case study, the researcher looks in considerable depth at a particular person, program, or event. In contrast, in an **ethnography**, the researcher looks in depth at an *entire group*—more specifically, a group that shares a common culture. (The word *ethnography* comes from *ethnos*, Greek for “a nation or other close-knit group of people,” and *graph*, “something written or recorded.”) The ethnographic researcher studies a group in its natural setting for a lengthy time period, often several months or several years. The focus of investigation is on the everyday behaviors of the people in the group (e.g., interactions, language, rituals), with an intent to identify

cultural norms, beliefs, social structures, and other patterns. Ideally the ethnographic researcher identifies not only explicit cultural patterns—those readily acknowledged by group members or easily observable in objects or behaviors—but also *implicit* patterns—those beliefs and assumptions that have such a below-the-surface, taken-for-granted quality that even group members aren't always consciously aware of them.

Ethnographies were first used in cultural anthropology, but they are now seen in sociology, psychology, education, and marketing research as well. The conception of the type of “culture” that can be studied has also changed over time: Whereas ethnographies once focused on long-standing cultural groups (e.g., people living on the island of Samoa), more recently they have been used to study such “cultures” as those of adult work environments, elementary school classrooms, violent adolescent groups, and Internet-based communities (e.g., Bender, 2001; Kozinets, 2010; McGibbon, Peter, & Gallop, 2010; Mehan, 1979).¹

An ethnography is especially useful for gaining an understanding of the complexities of a particular sociocultural group. It allows considerable flexibility in the methods used to obtain information, which can be either an advantage (to the astute researcher who knows what to look for) or a disadvantage (to the novice who may be overwhelmed and distracted by unimportant details). Hence, if you decide that an ethnography is the approach most suitable for your research problem, we urge you to get a solid grounding in cultural anthropology before you venture into the field (Creswell, 2007).

Method. Site-based fieldwork is the *sine qua non*—the essence—of any ethnography. Prolonged engagement in the group's natural setting gives the researcher time to observe and record processes that would be almost impossible to learn about by using any other approach.

The first step in an ethnographic study is to gain access to a site appropriate for answering the researcher's general research problem or question. Ideally, the site should be one in which the researcher is a “stranger” and has no vested interest in the outcome of the study. A site that the researcher knows well (perhaps one that involves close acquaintances) may be more accessible and convenient, but by being so close to the situation, the researcher may have trouble looking at it with sufficient detachment to gain a balanced perspective and portray an accurate picture of the processes observed (Creswell, 2007).

To gain access to a site, the researcher must often go through a **gatekeeper**, a person who can provide a smooth entrance into the site. This individual might be a tribal chief in a community in a developing country, a principal or teacher in a school or classroom, or a program director at a homeless shelter. Then, after gaining entry into the site, the researcher must establish rapport with and gain the trust of the people being studied. At the same time, the researcher must be open about why he or she is there. The principle of *informed consent* described in Chapter 4 is just as essential in an ethnography as it is in any other type of research.

Initially, the researcher casts a broad net, intermingling with everyone and getting an overall sense of the social and cultural context. Gradually, the researcher identifies **key informants** who can provide information and insights relevant to the research question and can facilitate contacts with other helpful individuals.

In some ethnographic studies, the researcher engages in **participant observation**, becoming immersed in the daily life of the people. In fact, over the course of the study, the researcher's role may gradually change from “outsider” to “insider.” The advantage here is that the researcher might gain insights about the group and its behaviors that could not be obtained in any other way. The disadvantage is that he or she may become so emotionally involved as to lose the ability to assess the situation accurately. In some situations, the researcher may even “go native,” joining the group and therefore becoming unable to complete the study (Creswell, 2007).

Throughout the fieldwork, the researcher is a careful observer, interviewer, and listener. Furthermore, he or she takes extensive field notes (written either on site at the time or in private later in the day) in the forms of dialogues, diagrams, maps, and so forth. Lengthy conversations and significant events can be recorded using audiotapes and videotapes. The researcher may also

¹ See Kraut and colleagues (2004) for a good discussion of the research possibilities, potential pitfalls, and ethical issues related to studying people's postings on the Internet.

collect artifacts (e.g., tools, ritualistic implements, artistic creations) and records (e.g., accounting ledgers, personal journals, lesson plans) from the group. In order to test hypotheses about a group's unconsciously shared beliefs or assumptions, some ethnographic researchers occasionally conduct *breaching experiments*—that is, they intentionally behave in ways they suspect might violate an unspoken social rule—and observe people's reactions (Mehan & Wood, 1975).

We must caution you that conducting a good ethnography requires both considerable patience and considerable tolerance. One experienced ethnographer has described the process this way:

It requires a great patience under any circumstances for me to "sit and visit." A rather inevitable consequence of being inquisitive without being a talker is that my conversational queries usually prompt others to do the talking. During fieldwork, I make a conscious effort to be sociable, thus providing opportunities for people to talk to me. My work ethic takes over to help me become not only more social but more attentive and responsive, and out pour the informants' stories and explanations so essential to good fieldwork.

(Parenthetically, I note my suspicion that many fieldworkers talk too much and hear too little. They become their own worst enemy by becoming their own best informant. . . .)

. . . I never confront informants with contradictions, blatant disbelief, or shock, but I do not mind presenting myself as a bit dense, someone who does not catch on too quickly and has to have things explained. . . . (Wolcott, 1994, p. 348)

Data analysis. As is true in case study research, data collection and data analysis in an ethnographic study often occur somewhat simultaneously. The analysis typically proceeds in the following sequence (Wolcott, 1994):

1. *Description.* The information obtained is organized into a logical structure. Ethnographers have used a variety of strategies to organize and describe the groups they have observed, including the following:
 - Describing events in chronological order
 - Describing a typical day in the life of the group or of an individual within the group
 - Focusing on a critical event for the group
 - Developing a story, complete with plot and characters
2. *Analysis.* The data are categorized according to their meanings. Patterns, regularities, and critical events are identified.
3. *Interpretation.* The general nature of the group and its practices is inferred from the categories, meanings, and patterns identified in Step 2. Existing theoretical frameworks in one's discipline may lend structure and support during the interpretation process.

Experienced ethnographers readily admit that it is virtually impossible—and perhaps not even desirable—to analyze their data with total objectivity. Wolcott (1994) has proposed that the researcher should instead strive for *rigorous subjectivity*; to achieve this end, he has suggested, the researcher should aim for balance, fairness, completeness, and sensitivity in the final analysis and interpretation of the data. Even so, any researcher conducting an ethnographic study must continually acknowledge—both to self and to others—that personal attitudes and opinions are inevitably creeping into and biasing observations and interpretations.

The research report. The final report for an ethnography may or may not be written in the impersonal style that is typical for many other forms of research. Sometimes it is, instead, a personal, literary narrative designed to engage the reader's attention and interest. Some experienced ethnographers are storytellers as much as they are scholars and researchers.

When you write about an ethnographic study, you will want to include the following information:

1. *An introduction that provides a rationale and context for the study.* Present your research question at the beginning of the report (at this point in the book, such a suggestion should hardly be a surprise!), and describe the nature of your study as it relates both to your question and to one or more theoretical perspectives. More generally, explain why the study was an important one for you to conduct and for others to read about.

2. *A description of the setting and methods.* Describe the group you studied and the methods you used to study it. Go into considerable detail about what people do and say, how they interact with one another, what systems and rituals they have in place, and so on. In other words, engage in the *thick description* described in Chapter 4. Ideally, an ethnographic report should “place the reader figuratively in the setting . . . transport the reader to the actual scene . . . make it real” (Creswell, 2012, p. 472).
3. *An analysis of the group studied.* Describe the patterns and themes you observed (e.g., the stated or unstated norms and conventions for behavior, the social hierarchy, the belief system). Present evidence (e.g., descriptions of artifacts, conversations with group members) to support your claims. To the extent that different group members have different perspectives, you should present those perspectives. Use the participants’ actual words—perhaps including their language or dialect as well—to give your account realism and “life.” Ethnographers give their research participants *voice*: They often use participants’ own words to convey a sense of what it is like to live and work within the group.
4. *A conclusion.* Relate your findings to your research question and to concepts, theories, and previous research in your discipline. In the process, also be explicit about how your background and perspectives may have colored your analysis and conclusions.

Although it may be impossible for you to be completely objective when you describe the group you have studied, you should nevertheless conscientiously try to avoid making judgments. Even small changes in wording can make a significant difference in this regard. For instance, rather than saying, “Only one villager had ever graduated from high school,” you might say, “One villager had graduated from high school” (Wolcott, 1994, pp. 352–353). And rather than saying, “Few pupils were at task,” you might instead say, “Five pupils appeared to be engaged in the assignment” (Wolcott, 1994, p. 353). Such words as *only* and *few* can imply such meanings as “insufficient” or “disappointing”—value judgments that an impartial researcher tries to avoid.

Ultimately you want to construct an in-depth portrait of a sociocultural group in all its complexities. Your portrait should give your readers a better understanding of the culture and help them look at it from the perspectives of the group’s members.

Phenomenological Study

In its broadest sense, the term *phenomenology* refers to a person’s perception of the meaning of an event, as opposed to the event as it exists external to the person. A **phenomenological study** is a study that attempts to understand people’s perceptions, perspectives, and understandings of a particular situation. In other words, a phenomenological study tries to answer the question, What is it like to experience such-and-such? For instance, a researcher might study the experiences of people caring for a chronically or terminally ill relative, living in an abusive relationship, or home-schooling a child.

In some cases, the researcher has had personal experience related to the phenomenon in question and wants to gain a better understanding of the experiences of others. By looking at multiple perspectives on the same situation, the researcher can then make some generalizations of *what something is like* from an insider’s perspective.

Method. Phenomenological researchers depend almost exclusively on lengthy interviews (perhaps 1 to 2 hours in length) with a carefully selected sample of participants. A typical sample size is from 5 to 25 individuals, all of whom have had direct experience with the phenomenon being studied (Creswell, 2007; Polkinghorne, 1989).

The actual implementation of a phenomenological study is as much in the hands of the participants as in the hands of the researcher. The phenomenological interview is often a very unstructured one in which the researcher and participants work together to “arrive at the heart of the matter” (Tesch, 1994, p. 147). The researcher listens closely as participants describe their everyday experiences related to the phenomenon; the researcher must also be alert for subtle yet meaningful cues in participants’ expressions, pauses, questions, and occasional sidetracks.

A typical interview looks more like an informal conversation, with the participant doing most of the talking and the researcher doing most of the listening.

Throughout the data collection process, the researcher suspends any preconceived notions or personal experiences that may unduly influence what the researcher “hears” the participants saying. Such suspension—sometimes called *bracketing* or *epoché*—can be extremely difficult for a researcher who has personally experienced the phenomenon under investigation. Yet it is essential if the researcher is to gain an understanding of the typical experiences that people have had.

Data analysis. The central task during data analysis is to identify common themes in people’s descriptions of their experiences. After transcribing the interviews, the researcher typically takes steps such as the following (e.g., Creswell, 2007):

1. *Identify statements that relate to the topic.* The researcher separates relevant from irrelevant information in the interview and then breaks the relevant information into small segments (e.g., phrases or sentences) that each reflect a single, specific thought.
2. *Group statements into “meaning units.”* The researcher groups the segments into categories that reflect the various aspects (“meanings”) of the phenomenon as it is experienced.
3. *Seek divergent perspectives.* The researcher looks for and considers the various ways in which different people experience the phenomenon.
4. *Construct a composite.* The researcher uses the various meanings identified to develop an overall description of the phenomenon as people typically experience it.

The final result is a general description of the phenomenon as seen through the eyes of people who have experienced it firsthand. The focus is on common themes in the experience but also with consideration of diversity in the individuals and settings studied.

The research report. There is no specific structure for reporting a phenomenological study. As is true for virtually any form of research, you will want to present the research problem or question, describe your methods of data collection and analysis, draw a conclusion about the phenomenon you have studied (in the form of a composite of your participants’ experiences), relate your findings to an existing body of theory and research, and discuss any practical implications of your findings. Your report should be sufficiently vivid that your readers come away feeling that “I understand better what it is like for someone to experience that” (Polkinghorne, 1989, p. 46).

Grounded Theory Study

Of all the research designs we describe in this book, a **grounded theory study** is the one *least* likely to begin from a particular theoretical framework. On the contrary, the major purpose of a grounded theory approach is to *begin with the data and use them to develop a theory*. The term *grounded* refers to the idea that the theory that emerges from the study is derived from and rooted in data that have been collected in the field rather than taken from the research literature. Grounded theory studies are especially helpful when current theories about a phenomenon are either inadequate or nonexistent.²

Typically, a grounded theory study focuses on a *process* related to a particular topic—including people’s actions and interactions—with the ultimate goal of developing a theory about that process. The approach has its roots in sociology (Glaser & Strauss, 1967) but is now also used in such fields as anthropology, geography, education, nursing, psychology, and social work. It has been used effectively for a wide range of topics—for instance, to study children’s eating habits, college students’ thoughts and feelings during classroom discussions, and workers’ stress levels in public service agencies (Do & Schallert, 2004; Kime, 2008; Skagert, Dellve, Eklöf, Pousette, & Ahlborg, 2008).

²Some researchers associate grounded theory studies with a particular method of data analysis—in particular, that of Corbin and Strauss (2008; Strauss and Corbin, 1990)—and suggest the term *emergent theory* as a broader, less prescriptive label for this approach (e.g., Jaccard & Jacoby, 2010).

Virtually all experts agree that a grounded theory researcher should have a firm grasp of general concepts and theoretical orientations in his or her discipline as a whole; hence, an in-depth literature review early in the process is essential. However, experts disagree about whether the researcher should look closely at previous findings *directly related to the present research problem* before collecting and analyzing data. For example, Glaser (1978) has argued that too much advance knowledge of earlier research regarding a topic may limit a researcher's ability to be open-minded about how to analyze and interpret his or her own data. In contrast, many others suggest that the advantages of conducting a relatively thorough literature review outweigh the disadvantages; in particular, previous works and writings about a topic can often help a researcher think more clearly and insightfully about the data collected (e.g., Hesse-Biber, 2010; Jaccard & Jacoby, 2010). Our own advice is to learn as much as you can about your research topic through a thorough review of the related literature but *to refrain from forming specific hypotheses about what you yourself might find*.

Method. As is true for the qualitative designs previously described, data collection is field-based, flexible, and likely to change over the course of the study. Interviews typically play a major role in data collection, but observations, documents, historical records, videotapes, and anything else of potential relevance to the research question might also be used. The only restriction is that the data collected *must* include the perspectives and voices of the people being studied (Charmaz, 2002, 2006; Corbin & Strauss, 2008).

Data analysis begins almost immediately, at which point the researcher develops *categories* to classify the data. Subsequent data collection is aimed at *saturating* the categories—in essence, learning as much about them as possible—and at finding any disconfirming evidence that may suggest revisions in the categories identified or in interrelationships among them. This process of moving back and forth between data collection and data analysis, with data analysis driving later data collection, is sometimes called the **constant comparative method**. The theory that ultimately evolves is one that includes numerous concepts and interrelationships among those concepts; in other words, it has *conceptual density* (Schram, 2006).

Data analysis. Experts disagree about the best approach to analyzing data in a grounded theory study (e.g., see Charmaz, 2006; Corbin & Strauss, 2008; Glaser, 1992). One widely used approach is that proposed by Corbin and Strauss (2008; Strauss and Corbin, 1990; also see Neuman, 2011), who suggest the following steps:

1. *Open coding.* The data are divided into segments and then scrutinized for commonalities that reflect categories or themes. After the data are categorized, they are further examined for *properties*—specific attributes or subcategories—that characterize each category. In general, open coding is a process of reducing the data to a small set of themes that appear to describe the phenomenon under investigation.
2. *Axial coding.* Interconnections are made among categories and subcategories. Here the focus is on determining more about each category in terms of
 - The conditions that give rise to it
 - The context in which it is embedded
 - The strategies people use to manage it or carry it out
 - The consequences of those strategies

The researcher moves back and forth among data collection, open coding, and axial coding, continually refining the categories and their interconnections—and perhaps combining or subdividing some of the categories—as additional data are collected.

3. *Selective coding.* The categories and their interrelationships are combined to create a *story line* that describes “what happens” in the phenomenon being studied.
4. *Development of a theory.* A theory, in the form of a verbal statement, visual model, or series of hypotheses, is offered to explain the phenomenon in question. The theory depicts the evolving nature of the phenomenon and describes how certain conditions lead to certain actions or interactions, how those actions or interactions lead to *other* actions, and so on, with the typical sequence of events being laid out. No matter what form the theory takes, *it is based entirely on the data collected*.

We have described these steps only in the most general terms. Corbin and Strauss's *Basics of Qualitative Research* (2008) offers more specific guidance and some helpful examples.

The steps just listed provide a structured and relatively systematic way of boiling down a huge body of data into a concise conceptual framework that describes and explains a particular phenomenon; as such, it has a semblance of rigor and objectivity that many researchers find appealing. Yet in some experts' eyes, these steps are *too* structured, to the point that they limit a researcher's flexibility and may predispose the researcher to identify categories prematurely (Charmaz, 2000; Glaser, 1992). Should you decide that a grounded theory study is the best way to tackle your research problem, we urge you to read diverse descriptions of the form that such a study might take (e.g., see Charmaz, 2000, 2002, 2006; Corbin & Strauss, 2008; Glaser, 1992).

The research report. The style of writing used to describe a grounded theory study is typically objective and impersonal. Building on Creswell's (2007) recommendations, we suggest that you include the following in your report:

1. *A description of the research question.* Describe your general research problem and explain how you delineated it more precisely over the course of the study.
2. *A review of the related literature.* Do *not* use the literature to provide the specific categories or themes for coding your data, but *do* use it to provide a rationale and context for your study.
3. *A description of your methodology and data analysis.* Describe the approach you took at the beginning of the study and how your approach evolved over time. Outline the nature of the sample and setting, as well as the specific methods (interviews, observations, etc.) you used. Explain the categories and properties you identified. Describe how your data collection was driven by your data analysis.
4. *A presentation of your theory.* Present the theory you have developed in a verbal or visual form, or, even better, both verbally and visually. Use some of your actual data (e.g., excerpts from interviews) to illustrate and support the theory.
5. *A discussion of implications.* Show how your theory is similar to or dissimilar from other theoretical perspectives. Explain how it relates to existing knowledge about the topic. Discuss potential implications of the theory for practice or future research.

Content Analysis

A **content analysis** is a detailed and systematic examination of the contents of a particular body of material for the purpose of identifying patterns, themes, or biases. Content analyses are typically performed on *forms of human communication*, including books, newspapers, personal journals, legal documents, films, television, art, music, videotapes of human interactions, transcripts of conversations, and Internet blog and bulletin board entries.³ For instance, a researcher might use a content analysis to determine what religious symbols appear in works of art, in what various activities teachers spend their time in the classroom, or what attitudes are reflected in the speeches or newspaper articles of a particular era in history. As you might infer from these examples, content analyses are found in a wide variety of disciplines, including psychology, history, art, education, journalism, and political science.

Of the five designs described in this chapter, a content analysis involves the greatest amount of planning at the front end of the project. The researcher typically defines a specific research problem or question at the very beginning (e.g., "Do contemporary children's books reflect traditional gender stereotypes?", "What religious symbols appeared in early Byzantine architecture, and with what frequency, during the years 527–867?"). The researcher also identifies the sample to be studied and the method of analysis early in the process.

Content analyses are not necessarily stand-alone designs. For example, a systematic content analysis might be an integral part of the data analysis in a phenomenological study (e.g., see Wennick, Lundqvist, & Hallström, 2009). A content analysis might also be used to flesh out the

³ Again we refer you to Kraut and colleagues (2004) regarding ethical issues related to studying people's postings on the Internet.

complex, multidimensional aspects of a descriptive or experimental study, resulting in a *mixed-methods design* with both qualitative and quantitative elements (see Chapter 10).

Method. As a general rule, a content analysis is quite systematic, and measures are taken to make the process as objective as possible. The following steps are typical:

1. The researcher identifies the specific body of material to be studied. If this body is relatively small, it is studied in its entirety. If it is quite large (e.g., if it consists of all newspaper articles written during a particular time period), a sample (perhaps a random sample) is selected.
2. The researcher defines the characteristics or qualities to be examined in precise, concrete terms. The researcher may identify specific examples of each characteristic as a way of defining it more clearly.
3. If the material to be analyzed involves complex or lengthy items (e.g., works of literature, transcriptions of conversations), the researcher breaks down each item into small, manageable segments that are analyzed separately.
4. The researcher scrutinizes the material for instances of each characteristic or quality defined in Step 2. When judgments are entirely objective (e.g., when the study involves looking for the appearance of certain words in a text), only one judge, or *rater*, is necessary. When judgments are more subjective (e.g., when the study involves categorizing discrete teacher behaviors as reflecting various teaching activities), two or three raters are typically involved, and a composite of their judgments is used.

Data analysis. Almost invariably, one crucial step in a content analysis is to tabulate the frequency of each characteristic found in the material being studied. Thus, virtually any content analysis is quantitative as well as qualitative. In some situations, appropriate statistical analyses are performed on the frequencies or percentages obtained to determine whether significant differences exist relevant to the research question. The researcher then uses such tabulations and statistical analyses to interpret the data as they reflect on the problem under investigation.

The research report. If you conduct a content analysis, either as your sole methodology or in combination with other designs, you should include the following in your research report:

1. *A description of the body of material you studied.* Describe the overall body of material you wanted to investigate and any sampling procedures that you used to select specific items or artifacts from it.
2. *Precise definitions and descriptions of the characteristics you looked for.* Define each characteristic precisely enough that another researcher could replicate your study. Consider using specific examples from your data to illustrate each characteristic.
3. *The coding or rating procedure.* Describe the procedure that the rater(s) used to evaluate the material and, if applicable, how multiple ratings were combined. If two or more raters were involved, also report the consistency with which they coded the data (see the discussions of *interrater reliability* and *correlation coefficient* in Chapters 4 and 11, respectively).
4. *Tabulations for each characteristic.* Report frequencies or percentages (or both) for each characteristic. Consider using tables or graphs as a way of reporting this information in a concise, organized fashion.
5. *A description of patterns that the data reflect.* Identify themes or trends in the material (e.g., as reflected in your tabulations).

Table 6.1 summarizes the nature of the five designs described in the preceding sections. Keep in mind, however, that these designs are not necessarily as distinctly different as the table might indicate. Any particular study may include elements of two or more qualitative designs. Remember, much qualitative research is, by its very nature, flexible and open-ended, and so it may continue to evolve over the course of a project. To the extent that your research question leads you to believe that two or more designs are equally relevant to your purpose, think creatively about how you might combine them into a single study.

TABLE 6.1 Distinguishing characteristics of different qualitative designs

Design	Purpose	Focus	Methods of Data Collection	Methods of Data Analysis
Case study	To understand one person or situation (or perhaps a very small number) in great depth	One case or a few cases within its/their natural setting	• Observations • Interviews • Appropriate written documents and/or audiovisual material	• Categorization and interpretation of data in terms of common themes • Synthesis into an overall portrait of the case(s)
Ethnography	To understand how behaviors reflect the culture of a group	A specific field site in which a group of people share a common culture	• Participant observation • Structured or unstructured interviews with "informants" • Artifact/document collection	• Identification of significant phenomena and underlying structures and beliefs • Organization of data into a logical whole (e.g., chronology, typical day)
Phenomenological study	To understand an experience from the participants' points of view	A particular phenomenon as it is typically lived and perceived by human beings	• In-depth, unstructured interviews • Purposeful sampling of 5–25 individuals	• Search for <i>meaning units</i> that reflect various aspects of the experience • Integration of the meaning units into a seemingly typical experience
Grounded theory study	To derive a theory from data collected in a natural setting	A process, including human actions and interactions and how they result from and influence one another	• Interviews • Any other relevant data sources	• Prescribed and systematic method of coding the data into categories and identifying interrelationships • Continual interweaving of data collection and data analysis • Construction of a theory from the categories and interrelationships
Content analysis	To identify the specific characteristics of a body of material	Any verbal, visual, or behavioral form of communication	• Identification and possible sampling of the specific material to be analyzed • Coding of the material in terms of predetermined and precisely defined characteristics	• Tabulation of the frequency of each characteristic • Descriptive or inferential statistical analyses as needed to answer the research question

Such flexibility should *not*, however, lead you to believe that you can conduct a qualitative research project in a sloppy, poorly-thought-through manner. On the contrary, the flexible nature of a qualitative study makes it just that much more challenging, especially for the novice researcher. *For anything you do in a qualitative study, you must have a definite rationale and a distinct purpose, and you must keep your overall goal—to answer your research question—clearly in sight at all times.*

CONCEPTUAL ANALYSIS EXERCISE Choosing a Qualitative Research Design

Following are brief summaries of five potential research projects. Identify the qualitative methodology that is probably most appropriate for each project. The answers appear after the "For Further Reading" list at the end of the chapter.

1. In an effort to learn the nature and appeal of long-standing social groups among American men, a researcher plans to spend a nine-month period with a local chapter of the Benevolent and Protective Order of Elks. By observing and interacting with the

- Elks, he hopes to observe the chapter's meetings, rituals, and charitable activities and to discover the chapter's beliefs, values, goals, and interpersonal dynamics.
2. A researcher wants to determine to what degree and in what ways television commercials might portray men and women in traditionally gender-stereotypical ways (e.g., how often men versus women are shown cleaning house, how often men versus women are making important business decisions).
 3. In order to learn how grassroots political parties emerge and develop over time, a researcher wants to study the origins and evolution of three recently established "Tea Party" groups, one in her own state and two in neighboring states.
 4. A researcher is intrigued by Asperger syndrome, a cognitive disability in which people have average or above-average intelligence and language skills but poor social skills and little or no ability to interpret other people's nonverbal social cues (body language, etc.). The researcher wonders what it must be like to be an adolescent with this syndrome—how a teenager is apt to feel about having few or no friends, being regularly excluded from classmates' social activities, and so on.
 5. A researcher wants to determine how doctors, nurses, and other hospital staff members coordinate their actions when people with life-threatening traumatic injuries arrive at the emergency room. The researcher can find very little useful research on this topic in professional journals.
-

Collecting Data in Qualitative Research

As you have seen, qualitative researchers often use multiple forms of data in any single study. They might use observations, interviews, objects, written documents, audiovisual materials, electronic entities (e.g., e-mail messages, Internet websites), and anything else that can help them answer their research question. Furthermore, many qualitative studies are characterized by an **emerging design**: Data collected early in the investigation often influence the kinds of data the researcher subsequently gathers.

In qualitative research, the potential sources of data are limited only by the researcher's open-mindedness and creativity. For example, in a school setting, a researcher might consider where various students are seated in the lunch room, what announcements are posted on the walls, or what messages are communicated in graffiti (Eisner, 1998). In an ethnographic study of a cultural group, a researcher might ask one or more participants to keep a daily journal or to discuss the content and meaning of photographs and art objects (Creswell, 2007).

Regardless of the kinds of data involved, data collection in a qualitative study takes a great deal of time. The researcher should record any potentially useful data thoroughly, accurately, and systematically, using field notes, audiotapes, sketches, photographs, or some combination of these. As they collect data, many qualitative researchers also begin jotting notes—sometimes called **memos**—about their initial interpretations of what they are seeing and hearing.

It is essential that data collection methods be consistent with the ethical principles presented in Chapter 4. The people being studied must know the nature of the study and be willing participants in it (this is *informed consent*), and any data collected should not be traceable back to particular individuals (thus maintaining participants' *right to privacy*). One common way of keeping personal data confidential is to assign various pseudonyms to different participants and to use those pseudonyms both during data collection and in the final research report.

Common to all qualitative studies is a need to identify an appropriate *sample* from which to acquire data. Another feature that most qualitative studies share (content analyses excepted) is heavy reliance on *observations*, *interviews*, or both, as a source of data. We now look at each of these three topics more closely.

Sampling

Qualitative researchers typically draw their data from many sources—not only from a variety of people but perhaps also from objects, text materials, and audiovisual and electronic records. The particular entities they select for analysis comprise their **sample**, and the process of selecting them is called **sampling**.

Only rarely—for instance, when a researcher conducts a content analysis of a small group of items—can a researcher look at *everything* that has potential relevance to the research problem. More typically, the researcher must be choosy about the data to gather and analyze and, as a result, will get an incomplete picture of the phenomenon in question. One experienced qualitative researcher has described the situation this way:

Whether observing, interviewing, experiencing, or pursuing some combination of strategies, you cannot be everywhere at once or take in every possible viewpoint at the same time. Instead . . . you develop certain perspectives by engaging in some activities or talking to certain people rather than others. . . . You build assertions toward the never-quite-attainable goal of “getting it right,” approximating realities but not establishing absolutes.

Your task, both derived from and constrained by your presence, is thus inherently interpretive and incomplete. The bottom line is that there is no bottom line: It is not necessary (or feasible) to reach some ultimate truth for your study to be credible and useful. (Schram, 2006, p. 134)

How you identify your sample must depend on the research question(s) you want to answer. If you want to draw inferences about an entire population or body of objects, you must choose a sample that can be presumed to *represent* that population or body. Ideally, this sample is chosen through a completely random selection process or through a process that incorporates appropriate proportions of each subgroup within the overall group of people or objects. We consider several such sampling strategies in our discussion of descriptive quantitative research in Chapter 8. (Truly effective researchers often draw on methodologies from diverse research traditions.)

Often, however, qualitative researchers are intentionally *nonrandom* in their selection of data sources. Instead, their sampling is *purposeful*: They select those individuals or objects that will yield the most information about the topic under investigation. For example, grounded theory researchers tend to engage in **theoretical sampling**, choosing data sources that are most likely to help them develop a theory of the process in question. Later, they may employ **discriminant sampling**, returning to particular data sources that can help them substantiate the theory.

A novice qualitative researcher might ask: How large should my sample be? How much is enough? There are no easy, cut-and-dried answers to these questions, but we offer three suggestions that might be helpful in decision making:

1. Be sure that the sample includes not only seemingly “typical” but also seemingly “non-typical” examples.
2. When a power hierarchy exists—as it does in the workplace and in many clubs and communities—sample from various levels in the hierarchy. For example, in the workplace, interview both bosses and employees; in a club or community, interview not only highly active, influential members but also less involved individuals (e.g., see Becker, 1970).
3. Actively look for cases that can potentially discredit emerging hypotheses and theories (see the description of *negative case analysis* in Chapter 4).

Ideally the sample should provide information not only about how things are *on average* but also about how much *variability* exists in the phenomenon under investigation.

Observations

The qualitative researcher may make observations either as a relative outsider or, especially in the case of an ethnography, as a participant observer. Unlike observations conducted in quantitative studies (see Chapter 8), observations in a qualitative study are intentionally unstructured and free-flowing: The researcher shifts focus from one thing to another as new and potentially

significant objects and events present themselves. The primary advantage of conducting observations in this manner is flexibility: The researcher can take advantage of unforeseen data sources as they surface.

Such an approach has its drawbacks, of course. The researcher (especially a novice researcher) won't always know what things are most important to look for, especially at the beginning, and so may waste considerable time observing and recording trivialities while overlooking entities that are more central to the research question. A second disadvantage is that *by his or her very presence*, the researcher may influence what people say and do or may change how significant events unfold (recall the discussion of *reactivity* in Chapter 4).

Recording events can be problematic as well. Written notes are often insufficient to capture the richness of what one is observing. Yet audiotapes and videotapes aren't always completely dependable either. Background noises may make tape-recorded conversations only partially audible. A video camera can capture only the events happening in a small, focused area. And the very presence of tape recorders and video cameras may make some participants uncomfortable.

If you decide to conduct observations as part of a qualitative study, we offer these suggestions:

1. Before you begin your study, experiment with various data recording strategies (field notes, audiotapes, videotapes), identify the particular methods that work best for you, and practice using them in other contexts.
2. When you begin your observations, have someone introduce you to the people you are watching. At this point, you should briefly describe your study and get participants' informed consent.
3. As you observe, remain relatively quiet and inconspicuous, yet be friendly to anyone who approaches you. You certainly don't want to discourage people from developing relationships with you and—perhaps later—taking you into their confidence.
4. If you take field notes, consider dividing each page of your notebook or word processing document into two columns. Use the left column to record your observations (making notes, drawing maps, etc.), and use the right column to write your preliminary interpretations—the *memos* we spoke of earlier.

The last suggestion is a particularly important one. *It is essential that you not confuse your actual observations with your interpretations of them*, for two reasons. First, you need to be as objective as you can in the records you keep. And second, your interpretations of what you have seen and heard are apt to change over the course of the study.

Interviews

Interviews can yield a great deal of useful information. The researcher can ask questions related to any of the following (Silverman, 1993):

- Facts (e.g., biographical information)
- People's beliefs and perspectives about the facts
- Feelings
- Motives
- Present and past behaviors
- Standards for behavior (i.e., what people think *should* be done in certain situations)
- Conscious reasons for actions or feelings (e.g., why people think that engaging in a particular behavior is desirable or undesirable)

However, keep in mind that, especially when a researcher asks about *past* events, behaviors, and perspectives, interviewees must rely on their memories, and human memory is rarely as accurate as a tape recorder or video recorder might be. In fact, people's memories are subject to considerable distortion: People are apt to recall what *might* or *should* have happened (based on their attitudes or beliefs) rather than what actually *did* happen (e.g., see Brainerd & Reyna, 2005; Schwarz, 1999). And even when people are talking about present circumstances, they aren't always terribly insightful—and sometimes they're intentionally dishonest—about their attitudes, feelings, and motives (Corallo, Sackur, Dehaene, & Sigman, 2008; Uziel, 2010). Thus, the shrewd

researcher will seek to substantiate participants' introspective reports with observations of their behavior and other relevant data (Locke, 2009).

Interviews in a qualitative study are rarely as structured as the interviews conducted in a quantitative study (more about quantitative study interviews in Chapter 8). Instead, they are either open-ended or semistructured, in the latter case revolving around a few central questions. Unstructured interviews are, of course, more flexible and more likely to yield information that the researcher hadn't planned to ask for. Their primary disadvantage is that the researcher gets different information from different people and thus may not be able to compare the responses of various interviewees.

In some cases, a researcher may want to interview several participants simultaneously in a **focus group**. To conduct a focus group, the researcher gathers several people (usually no more than 10 or 12) to discuss a particular issue for 1 to 2 hours. A moderator—someone who may or may not be the researcher—introduces the issues to be discussed, makes sure that no one dominates the discussion, and keeps people focused on the topic. Focus groups are especially useful when time is limited, group members feel comfortable sharing their thoughts and feelings with one another, and the group interaction might be more informative than individually conducted interviews (Creswell, 2007; Neuman, 2011).

PRACTICAL APPLICATION Conducting Interviews in a Qualitative Study

Conducting an informative interview is not as easy as it might seem. In this section we offer guidelines for the novice researcher and then suggest computer technology that can make the task easier.

GUIDELINES Conducting a Productive Interview

The following suggestions are based partly on our own experiences and partly on guidance offered by experts in qualitative research (Creswell, 2007, 2009; Eisner, 1998; Shank, 2006; Silverman, 1993):

1. *Identify some questions in advance.* Conducting an unstructured interview effectively requires considerable experience and skill: The researcher must sense when the conversation is drifting in an unproductive direction and gently guide it back on course. Novice researchers often have greater success when they prepare a few questions in advance and make sure that all of the questions are addressed at some point during the interview. These questions should, of course, be related to the research questions and overall research problem. As an example, in a qualitative study she conducted for her doctoral dissertation, Debby Zambo examined how children with reading disabilities believe their minds work when they read. She worked with and extensively studied 11 children in grades 5 through 9, interviewing them 10 to 15 times over the course of her investigation. Figure 6.1 presents an excerpt from her dissertation, in which she shows how her interview questions aligned with her research questions.

For any single interview, limit your list of questions to a small number, perhaps five to seven of them. (Although Debby Zambo had many more questions than this, she spread them throughout a dozen or so interviews with each child.) You will find that you won't necessarily need to ask every question explicitly, as the answers to some may emerge while a participant is responding to others.

Ideally, interview questions should encourage people to talk about a topic *without* hinting that they give a *particular* answer. In other words, avoid leading questions. Questions such as "What is going on now?" "What is it like to work here?" and "What's a typical day like?" can stimulate an informative conversation without suggesting that one kind of response is somehow more desirable than another (Shank, 2006).

2. *Consider how participants' cultural backgrounds might influence their responses.* In an effort to ascertain men's beliefs about ideal family size for a research project in what is now Bangladesh, Howard Schuman (1967) asked a seemingly simple question: "Suppose you had no children. How many would you like to have?" Most men responded, "As many as God wills." This

FIGURE 6.1

Example of how a researcher might align interview questions with research questions

From *Uncovering the Conceptual Representations of Students with Reading Disabilities* (pp. 140–142) by D. Zambo, 2003, unpublished doctoral dissertation, Arizona State University, Tempe. Reprinted with permission.

Research Question	Interview Question
1. What do students with reading disabilities think about reading and themselves? a. What are their thoughts about reading? b. What are their ideas about themselves and reading?	What do they think reading is all about? What do they find easy/difficult to read? Who do they think good/poor readers are and what do good/poor readers do? How do you become good/poor at reading? What are they reading? What do they think is easy/difficult to read? What goes on in their head when they read easy/difficult things? What is their activity level (calm/fidgety) when they read? What body parts do they use when they read? How do they think reading has/will impact their lives in the past, present, and future?
2. What emotions are evoked when they read?	Do they get frustrated when they read? What other emotions may be involved when they read? Does believing they can get better at reading help them be a better reader? Does hoping they can get better at reading help them be a better reader? Does wishing they can get better at reading help them be a better reader?
3. What do children with reading difficulties know about the cognitive processes of reading? a. What do they know about attention? b. What do they know about their memory and reading?	What is attention? Do they recognize that they must focus their attention when they read? What do they focus on? Why do they focus on that? Do they have difficulty with attention? If so, what do they do? Is their attention easy or difficult to capture when they read? Can they sustain their attention enough when they read? What do they do to sustain their attention? How consistent is their attention? What do they do to make their attention consistent? Is their attention better on some days and when is it better? What do they do if their attention is better on some days? What distracts them when they read? Do ideas and memories pop into their heads and distract them when they read? What do they know about memory in general? What do they do to put things into their memory? What do they do to keep things in their memory? How do they remember what they read? How do they remember/understand what they have read?

(continued)

FIGURE 6.1

Continued

Research Question	Interview Question
4. What do students with dyslexia know about the brain and reading?	Do they understand the brain is interconnected with external body parts? Analogy—Can they create an analogy for the brain? Metacognition—Thinking about thinking—What do they wonder about their mind/brain? [What do they] think about their thinking? Can they differentiate mental entities (thoughts, dreams, and memories) from close imposters?
5. What do children with dyslexia know about their dyslexic mind?	How do their brains work when they read? Are their brains like or different from other's brains when they read? Do they listen/see/feel things in their brains when they read? How do they do this? Do they think their minds are active when they read? What happens in their minds when they read? What do they do to make this happen? Are they aware of what is in their minds as they read? Are their minds excited when they read? How do things get from a book to their brains?

response reflected a widespread cultural tradition at the time: to leave one's fate in the hands of God, or at least to *say* that one's fate is in God's hands. Wisely, Schuman revised the question: "Suppose you had no children. If God wished to give you as many children as you wished, how many would you wish for?" (p. 22). This revision yielded responses that were far more useful in addressing Schuman's research question.

As Schuman discovered, participants' cultural backgrounds can influence their interview responses in ways you haven't necessarily anticipated. For instance, if you are interviewing people from Asian cultures, you should be aware that they are less likely to brag about their individual accomplishments than Westerners are (Heine, 2007). A naive researcher might erroneously conclude that Asian individuals are less productive than Western individuals, when in reality Asian individuals are merely less *boastful* than their Western counterparts.

Social scientists have only begun to identify the many ways in which people from various cultural groups may respond differently in interview situations. Thus we can give you only general, nonprescriptive advice here: Be sensitive to the fact that culture may play a significant role in how your participants interpret and respond to your questions, and experiment with multiple ways of asking for the kinds of information you ultimately want to obtain.

3. *Make sure your interviewees are representative of the group.* You should choose people whom you expect to give you typical perceptions and perspectives. But as noted in the earlier discussion of sampling, you may also intentionally pick a few "extremists" or other exceptional individuals; however, when you do so, you should identify them as such in your notes.

4. *Find a suitable location.* In theory, you can conduct an interview anywhere that people are willing to talk to you. But you will probably have a more successful interview if you find a quiet place where you and your interviewee are unlikely to be distracted or interrupted.

5. *Get written permission.* Explain the nature of the study and your plans for using the results. Ask the participant (or, in the case of a child, the participant's parent or legal guardian) to sign an informed consent form. Offer to provide an abstract or copy of the research report once you have completed the study.

6. *Establish and maintain rapport.* Begin the conversation with small talk that can break the ice. Be courteous and respectful at all times. Show genuine interest in what the person has to say.

Interviews in qualitative studies are typically quite informal, to the point where they may appear similar to casual conversation. There is one critical difference between a qualitative

interview and normal dialogue, however: The researcher wants to gain information from the interviewee without also revealing his or her own perspectives. In other words, a critical element of most intimate conversations—disclosure of one's thoughts, beliefs, and feelings—is lopsided, with only one member of the pair doing the disclosing. To maintain rapport and general feelings of closeness and trust, therefore, you must show compassion and interest in other ways, perhaps through body language (smiling, maintaining eye contact, leaning forward) and such neutral encouragements as "Go on" and "What do you mean?" (Shank, 2006).

7. *Focus on the actual rather than on the abstract or hypothetical.* You are more likely to get revealing information if you ask what a person *does or would do in a specific situation*. For example, if you are interviewing a teacher, ask questions about specific teaching strategies rather than about educational philosophy. Otherwise, you might get nothing more than what Eisner (1998) has described as "pious, canned proclamations that seem as though they had been snatched from a third-rate philosophy of education text" (p. 183).

8. *Don't put words in people's mouths.* Let people choose their own way of expressing their thoughts. A good interviewer is, above all, a good listener who lets people say what they want to say in the way they want to say it. Furthermore, a good interviewer recognizes that people may reveal inconsistencies in their recollections, attitudes, and logic: Their perceptions won't necessarily all fit together in a neat little package (Kvale, 1996).

9. *Record responses verbatim.* Whether you use handwritten notes, shorthand, a tape recorder, or laptop computer, capture everything the person says, especially if the interview is an unstructured one. If you suspect that an interviewee may have said something other than what he or she intended to communicate, read or play back the response and ask if it accurately reflects his or her thoughts.

10. *Keep your reactions to yourself.* Although you won't necessarily want to maintain a continual "poker face," you're more likely to get accurate information if you don't show surprise, agreement, or disapproval of what someone tells you.

11. *Remember that you are not necessarily getting the facts.* As confident and convincing as some of your participants may be, you should always treat their responses as *perceptions* rather than as facts.

12. *When conducting a focus group, take group dynamics into account.* Whenever you gather two or more individuals into a single interview, these individuals will rarely act as true equals. Some participants are likely to dominate the conversation. Others may be reluctant to express their views, perhaps because they are shy or feel uncertain about the validity of their perspectives. In most cases, you will get more representative data—and hence more *useful* data—if you make sure that everyone in the group has a chance to answer each question. Accordingly, you should keep your list of questions for a focus group quite short. And if you are recording the focus group session, ask participants to identify themselves by name at the beginning of the session; having them do so will help you identify different speakers when you transcribe the session later on.



Using Technology to Facilitate Data Collection and Transcription

With appropriate hardware and software, most laptops can be configured to serve as audio recorders. And, of course, videos recorded on a camcorder can be easily downloaded to a personal computer. Meanwhile, transcription software, such as HyperTRANSCRIBE, lets you mark key points in a videotaped or audiotaped interview, retrieve desired pieces of information quickly, and slow down what you have recorded so that you can transcribe it more easily. Other software programs will even do your transcribing for you. The capabilities of this software are expanding all the time. We urge you to look at such software as Dragon Dictate, IBM VoiceType Simply Speaking, Naturally Speaking, and ViaVoice to see how their latest versions might make your recording and interviewing tasks easier and less time-consuming.

In some cases, you can conduct qualitative interviews over long distances through various Internet mechanisms, including e-mail, instant messaging, or video conferencing. Keep in

mind, however, that ethical standards don't fly out the window simply because you are conversing with people in cyberspace rather than in the same room. You must still seek participants' (or parents') informed consent, and you must protect participants' privacy. Furthermore, you must ensure that participants have appropriate characteristics and qualifications for your investigation—something that may be difficult to determine if you never see these individuals in the flesh.

Organizing and Analyzing Qualitative Data

As you have undoubtedly realized by this time, there is usually no single “right” way to analyze the data in a qualitative study. The researcher begins with a large body of information and must, through inductive reasoning, sort and categorize it and gradually boil it down to a small set of abstract, underlying themes. Even in content analysis—an approach that, on the surface, may seem quite straightforward and matter-of-fact—the researcher often determines the specific characteristics to be studied only after carefully scrutinizing the body of material in search of potentially meaningful characteristics to identify and count.

In the quantitative designs we examine in Chapters 8 and 9, data analysis and data interpretation are, in large part, two separate steps, with numerical data being mathematically manipulated and statistically analyzed, followed by interpretations of those manipulations and analyses in order to address the original research questions and hypotheses. In most qualitative research, however, data analysis and interpretation are closely interwoven, and both are often enmeshed with data collection as well. Schram (2006) has expressed this idea quite eloquently:

Experiences do not speak for themselves; nor do features within a research setting directly or spontaneously announce themselves as worthy of your attention. As a qualitative fieldworker, you cannot view your task simply as a matter of gathering or generating “facts” about “what happened.” Rather, you engage in an active process of *interpretation*: noting some things as significant, noting but ignoring others as not significant, and missing other potentially significant things altogether. (p. 11)

Creswell (2007) has described a **data analysis spiral** that is, in our view, equally applicable to a wide variety of qualitative studies. Using this approach, you go through the data several times, taking the following steps:

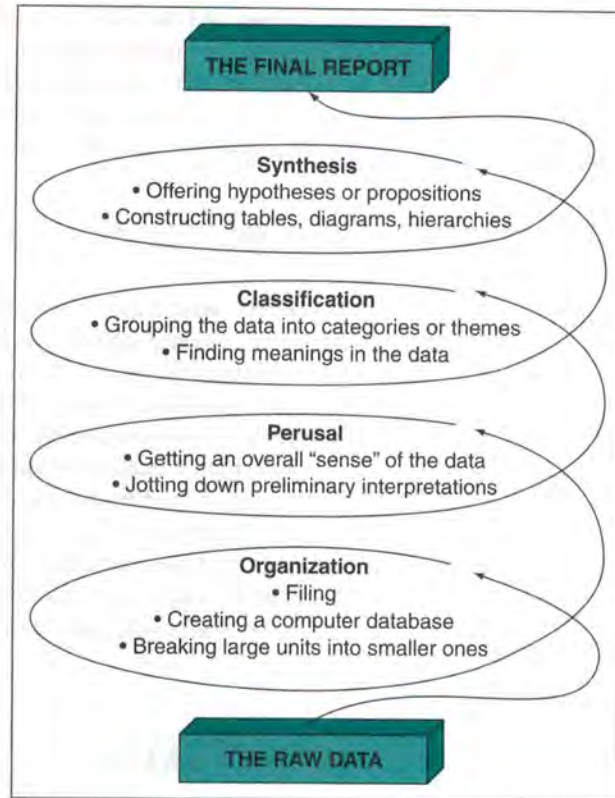
1. Organize the data, perhaps using index cards, manila folders, or a computer database. You may also break down large bodies of text into smaller units, perhaps in the form of stories, sentences, or individual words.
2. Peruse the entire data set several times to get a sense of what it contains as a whole. In the process, you should jot down a few memos that suggest possible categories or interpretations. If your data are in paper form, you might write comments in the margins or use Post-it notes to capture your preliminary thoughts. If your data are in electronic form, you might use the *insert comment* feature available in many software programs, or you might add your initial impressions in a different font or color or, for a spreadsheet or database, in a separate column or field.
3. Identify general categories or themes, and perhaps subcategories or subthemes as well, and then classify each piece of data accordingly. At this point, you should be getting a general sense of patterns—a sense of *what the data mean*.
4. Integrate and summarize the data for your readers. This step might include offering propositions or hypotheses that describe relationships among the categories. It might also involve packaging the data into an organizational scheme such as a table, figure, matrix, or hierarchical diagram.

We depict this spiral graphically in Figure 6.2.

No matter how you proceed, the data analysis for a qualitative study is a complex and time-consuming process. You must wade through a great deal of information, some of which will be

FIGURE 6.2

The data analysis spiral
(based on Creswell, 2007)



useful and some of which will not. Furthermore, the data you obtain are apt to be multifaceted and may simultaneously reflect several distinct meanings.

In a qualitative study, the interpretation of the data will inevitably be influenced by the researcher's biases and values to some extent, reflecting the notion of *researcher as instrument* mentioned earlier. Nevertheless, we urge you to do as much as you can to minimize the extent to which your prior expectations and opinions enter into your final analysis, perhaps by using some or all of the following strategies:

- Collect two or more different kinds of data (e.g., observations, interviews) related to any particular phenomenon.
- Get multiple and varying perspectives on any single issue or event.
- Make a concerted effort to look for evidence that contradicts your hypotheses.
- In your final research report, acknowledge any biases you have, so that your readers can take them into account when reading the report.

Using Computer Databases to Facilitate Data Organization and Interpretation



We authors are guessing that you will use word processing or similar software to record interviews and perhaps some of your other data as well. By storing your data on a computer, you can easily retrieve any piece of information using a relevant keyword, and you can sort your data quickly and in multiple ways. As a precaution against some unforeseen catastrophe (e.g., a flood or fire), you should back up your files on a flash drive or other external storage device that you store in a safe location (e.g., in a safe deposit box or at the home of a trustworthy relative).

We suggest that you also consider using computer software to help you organize and interpret your data. For some studies, a simple spreadsheet program such as Excel may suffice (see Appendix A). Other software programs are especially suited for complex qualitative research studies (e.g., ATLAS.ti, Ethnograph, SuperHyperQual, HyperRESEARCH, Kwalitan, MAXQDA, NVivo). Such programs provide a ready means of storing, segmenting, and organizing lengthy

field notes, and they are designed to help you find patterns in your notes. Typically you can transfer—in computer language, you can *import*—word processing files into the programs; some programs let you include photographs, audiotapes, and videotapes as well. Any one of these programs may take some study and practice to master, but keep in mind that the time you spend learning how to use it effectively is likely to *save* you time in the long run.

PRACTICAL APPLICATION Planning a Qualitative Study

As you have seen, a qualitative research project is not something to be entered into casually. The flexibility of qualitative methodologies is an advantage for experienced researchers but often a disadvantage for novices, who may not have sufficient background or training to make wise decisions about how to proceed. Furthermore, data collection and data analysis may be far more complex and time-consuming than a researcher anticipated.

If you think a qualitative approach might be suitable for your purposes, you may want to do a pilot study first to find out whether you feel comfortable with the ambiguity and lack of structure in the process. We urge you, too, to learn as much as you can about qualitative research strategies, perhaps by reading some of the sources in the “For Further Reading” list at the end of this chapter. Once you have determined that you have both the time and skills to conduct a qualitative study, you may find the following checklist helpful in your planning.



CHECKLIST

Pinning Down the Details of a Qualitative Study

WHAT IS THE PURPOSE OF THE PROJECT?

- _____ 1. What is the current status of knowledge pertaining to the question?

- _____ 2. Why is the study important?

WHAT IS THE SPECIFIC FOCUS AND DESIGN OF THE PROJECT?

- _____ 3. Will the focus be on individuals, groups, cultures, experiences, processes, or content?

- _____ 4. Will the design be a case study, ethnography, phenomenological study, grounded theory study, content analysis, a combination of two or more of these, or none of these?

WHAT DATA ARE NEEDED?

- _____ 5. How will you gain access to the site?

_____ 6. How much time will you need?

_____ 7. What resources are needed and available?

_____ 8. Are there any existing constraints on data collection?

HOW WILL THE DATA BE COLLECTED?

_____ 9. How will the participants or materials be sampled?

_____ 10. What role will you, as the researcher, assume?

_____ 11. How will you ensure anonymity and confidentiality for the participants?

_____ 12. What procedures will you follow, and in what order?

HOW WILL THE ANALYSIS BE CONDUCTED?

_____ 13. What is the unit of analysis (person, event, story, artifact, etc.)?

_____ 14. What methods of analysis will you use?

_____ 15. How will you make sure that you and others can have confidence in your findings?

HOW WILL THE FINDINGS BE COMMUNICATED?

_____ 16. How will you describe the context?

_____ 17. How will you convey the participants' perspectives?

_____ 18. What format(s) will you use to synthesize the data?

Criteria for Evaluating Qualitative Research

How do readers, reviewers, and practitioners assess the worth of a qualitative proposal or research study? What characteristics are essential to a good study? What makes one study “excellent” and another study only “marginal”?

Experienced qualitative researchers have offered a variety of standards that might be used to evaluate a qualitative study (Altheide & Johnson, 1994; Creswell, 2007, 2009; Eisner, 1998; Gall, Gall, & Borg, 2007; Glaser, 1992; Howe & Eisenhardt, 1990). We have boiled down their suggestions to nine general criteria:

1. **Purposefulness.** The research question drives the methods used to collect and analyze data, rather than the other way around.
2. **Explicitness of assumptions and biases.** The researcher identifies and communicates any assumptions, beliefs, values, and biases that may influence data collection and interpretation.
3. **Rigor.** The researcher uses rigorous, precise, and thorough methods to collect, record, and analyze data. The researcher also takes steps to remain as objective as possible throughout the project.
4. **Open-mindedness.** The researcher shows willingness to modify hypotheses and interpretations when newly acquired data conflict with previously collected data.
5. **Completeness.** The researcher depicts the object of study in all of its complexity. The researcher spends sufficient time in the field to understand all nuances of a phenomenon; describes the physical setting, behaviors, and perceptions of participants; and gives readers an in-depth, multifaceted picture of the phenomenon (i.e., *thick description*).
6. **Coherence.** The data yield consistent findings, such that the researcher can present a portrait that “hangs together.” Multiple data sources converge onto consistent conclusions (*triangulation*), and any contradictions within the data are reconciled.
7. **Persuasiveness.** The researcher presents logical arguments, and the weight of the evidence suggests one interpretation to the exclusion of others.
8. **Consensus.** Other individuals, including the participants in the study and other scholars in the discipline, agree with the researcher’s interpretations and explanations.
9. **Usefulness.** The project yields conclusions that promote better understanding of the phenomenon, enable more accurate predictions about future events, or lead to interventions that enhance the quality of life.

PRACTICAL APPLICATION Evaluating Qualitative Research Studies

Drawing from the preceding criteria, as well as from guidelines presented by Good (1993), we offer the following checklist to help you evaluate a final report for a qualitative research project. Some of the items in the checklist are equally applicable to a proposal for a future qualitative project.



CHECKLIST

Evaluating a Qualitative Study

METHODS

		YES	NO
_____	1. Is the context/setting of the study adequately described?	_____	_____
_____	2. Are techniques for data collection appropriate for the research problem? Are they thoroughly and precisely described?	_____	_____
_____	3. Are multiple data sources used?	_____	_____
_____	4. Are sufficient data collected from a variety of participants over an appropriate length of time?	_____	_____
_____	5. Are criteria for the selection of participants, informants, or materials presented? Is the sample described in sufficient detail?	_____	_____
_____	6. Are the roles of the researcher and participants made clear?	_____	_____
_____	7. Does the researcher identify any assumptions, beliefs, values, or biases that might influence data collection or analysis?	_____	_____

FINDINGS AND INTERPRETATIONS

_____	8. Are the data analysis techniques appropriate for the research question, methodology, and theoretical framework?	_____	_____
_____	9. Are data analysis techniques explicitly described?	_____	_____
_____	10. Do data analysis techniques allow for revision and reinterpretation as new data come to light?	_____	_____
_____	11. Is triangulation of the various data sources addressed?	_____	_____
_____	12. If used, are tables, figures, and other graphics easy to read and interpret? Do they enhance the reader's ability to understand the study?	_____	_____
_____	13. Are sufficient data reported to support the conclusions drawn?	_____	_____
_____	14. Are any irrelevant and unnecessary data reported? If so, what should be deleted? _____	_____	_____
_____	15. Are discrepant data discussed and reconciled?	_____	_____
_____	16. Have the setting and observations been sufficiently described to present a convincing case?	_____	_____
_____	17. Are participant "voices" used to support the assertions and present multiple perspectives?	_____	_____
_____	18. Is the report detailed enough that the findings can be compared to other studies in other contexts?	_____	_____
_____	19. Is the discussion congruent with the research question and rationale for the study?	_____	_____
_____	20. Are implications for theory and/or practice discussed?	_____	_____

_____	21. Have other scholars in the field reviewed the proposal or report? If so, do they agree that the approach, methodology, and conclusions are appropriate?	_____	_____
_____	22. Have participants in the project read the report? Do they agree with its findings?	_____	_____
WRITING STYLE			
_____	23. Is the writing style (e.g., expository, narrative) appropriate for the study?	_____	_____
_____	24. If a narrative is used, are the stories understandable? Are they authentic?	_____	_____
_____	25. Is the writing style concise? Is the argument clear?	_____	_____
_____	26. Are the writer's arguments logical and persuasive?	_____	_____

A Sample Dissertation

As an illustration of how the methodology of a qualitative research study might look, we present excerpts from Robin Smith's doctoral dissertation conducted at Syracuse University (Smith, 1999). The study was a multiple case study that also incorporated elements of grounded theory research and content analysis.

The study focused on five high school students who had significant intellectual disabilities. In particular, it examined the nature of the students' involvement and participation in high school classrooms. It looked, too, at teachers' perceptions and interpretations of the students' disabilities and academic performance.

The dissertation's "Method" chapter begins with an overview of the research strategies used and a rationale for selecting the individuals to be studied. It then presents more specific information about each of the five students: Gerald, Trish, Nick, Tyrone, and Abe (all pseudonyms). We pick up the chapter at the point where it begins a discussion of data collection. As we have done in preceding chapters of this book, we present excerpts on the left and a running commentary on the right.

Dissertation ANALYSIS 3	
DATA COLLECTION PROCESSES	Comments
Data gained in the varied academic settings of the five students assisted in understanding the patterns of academic participation and the meanings and relationships of the five students regarding their academic participation in high school. I gathered data from the following sources: Observations Over three school semesters, I conducted observations of five high school students who were attending high school and enrolled in at least one academic subject in the general high school curriculum. These observations totaled 52 visits ranging in length ...	<i>Here the author provides information about the amount of time she spent in the field. Her observations varied considerably in duration depending on the situation; we are more apt to see such flexibility in a qualitative study than in a quantitative one.</i>

the shortest was 15 minutes . . . the longest, 6 hours. . . . (The author continues with a detailed discussion of the kinds of observations made and the circumstances in which she made them.)

Conversations and Interviews

I had conversations and interviews with adults involved and concerned with the students, such as general and special education teachers, assistants, and parents. I recorded and described these conversations in field notes and transcriptions. . . .

The semi-structured interviews with the parents of each student included the following kinds of questions:

1. Tell me about the history of your child's schooling.
2. What are the child's strengths? That is, what is he or she good at?
3. Where does it get hard for the student?
4. How does he or she like high school? How can you tell?
5. What do you see your child learning?
6. What are your goals and dreams for your child?
7. What else should I know about your child to better understand what is happening for him or her in school?

I conducted similar interviews with the special education teachers, which included discussion of their educational goals for the student. I conducted one formal interview with each special education teacher, with further interviews as necessary to enhance my understanding of my data. These other interviews were often in the form of brief conversations during or in between class, interviews by appointment, and phone conversations.

I also conducted interviews with the general education teachers in the form of formal, informal, or brief conversations that fit into the teachers' schedules. . . . I also had some conversations with the general education teachers by staying a few minutes after class and asking them questions about what I had observed that day or how they thought the student was doing. . . .

I taped and had transcribed in-depth interviews, and I embedded observer comments in the transcribed text as I reviewed it. I wrote down informal conversations as soon as possible, and when possible, wrote during the conversations according to the comfort level of the participants with note taking. I used a Hewlett Packard 200LX-palmtop computer, which enabled me to take legible and detailed notes and add more detail soon after an observation.

Official Records and Documents

Official records and documents were another source of information. At the very end of my study I went to the district office of special education, which kept the official records of all five of the students. I looked in each file to learn what I could about the students' grades and progress reports, along with the professional assessments and recommendations regarding the students' schooling. I took notes on my hand-held computer and read long quotes into my tape recorder for

The interviews were presumably structured in this manner so that similar kinds of information would be obtained about each child.

The author used follow-up unstructured interviews to gather additional information as needed. This strategy is consistent with a grounded theory approach, in which the researcher moves back and forth between data collection and data analysis.

The author used audiotapes and transcriptions to capture the details of in-depth interviews; she also wrote notes about shorter, more informal conversations as soon after they took place as possible. The phrase "according to the comfort level of the participants with note taking" might have been better worded as "to the extent that participants felt comfortable with my note taking." However, the phrase reflects an appropriate sensitivity about taking notes only when it did not make a participant feel uneasy. And notice the author's use of a small computer to facilitate data collection!

Why did the author wait until the end to look at school records and documents? Later in this excerpt you will see her reason: She was worried that early knowledge of these records would bias her interpretations of what she observed in the classroom.

later reference and transcription. I took notes on students' work in class and from some student work I found in the files, and collected samples of their work where possible.

Finally, I relied heavily on very detailed field notes. At first I wrote everything I saw. As I narrowed my focus I consistently included the students' interactions with adults and peers, their reactions to what was going on, and what other students were doing at the same time. Describing interactions of the nonspeaking students was challenging; due to the crowded conditions of several of the general education classes and my being in a wheelchair, I was not always able to be close enough to the student to observe facial expressions. Fortunately, each student was accessible to me most of the time, especially when I was well into the study and a couple of sympathetic teachers invited the student to sit where I could be close by. Thus many of my observations were able to include whispered dialogue between the student and support person helping with an assignment.

CODING AND ANALYSIS

...As I collected and analyzed data from preliminary observations, I found issues to explore ... questions arose that created a need for further observing or interviewing. Using the constant comparative method of analysis (Glaser & Strauss, 1967), I collected data, looked for emerging themes and recurrent events, categorized them, and reevaluated my themes and categories. As I collected more data, I wrote analytic memos about my data, and reevaluated my previous theories as I compared old data with new (Bogdan & Biklen, 1992, pp. 72-75). The themes of academic engagement, generated by my pilot study, continued to expand in depth and breadth, and they generated more themes that guided the development of my study.

For example, Nick, one of the students I observed, sat with his assistant in the last row by the door, separated by another row of desks from the class; he seemed an observer in class lectures and discussions. When his assistant supported him to participate in hands-on activities, the assistant did the task for the student. The educators in the room said to me, "He doesn't understand much of what's going on," and they did not expect him to benefit from the actual curriculum content ("He's not getting much out of it."). In contrast, Trish, a student with even less physical coordination and verbal expression, followed a full academic schedule, and many of her teachers considered her to be involved, interested, and learning. This led me to look for signs of expectations of the student and how people evaluated the students. Thus, early data codes such as "expectation," "perception," and "assessment" led to a chapter regarding expectations and another regarding types of assessments.

I used Q.S.R. Nudist (QSR-NUD*IST, 1995) to code my data. This program enabled me to identify text segments in various ways, including participants' names and roles, as well as assigned categories such as "engaged," "disengaged," and "academics," that resulted in 98 data codes. A few of these original codes survived my ongoing revisions and collapsing of categories to my final analysis. I printed categories out in groups and coded them again by hand, testing new coding categories by merging several categories and reexamining the data. For example, many of the text segments that I had labeled "expectation" evolved into "assessment." Once I had determined

This narrowing of focus as the study proceeds is frequently found in qualitative research.

Notice how the author is looking for nonverbal as well as verbal information. Notice, too, how cooperative participants (in this case, some "sympathetic teachers") can facilitate data collection.

The constant comparative method is central to grounded theory research.

Here we see that the author conducted an earlier pilot study—something we urge any beginning researcher to do, particularly when planning a qualitative study.

Notice the author's attention to Nick's physical distance from other students—a clear, nonverbal indicator that Nick is essentially a nonparticipant, an outsider, in this classroom.

The ability to contrast one situation with another is a key advantage of a multiple case study.

Here we see open coding, the first step in data analysis in a grounded theory study.

*NUD*IST was an early, groundbreaking computer database program especially suited for data collection in qualitative research.*

that assessment was an important category, I subdivided it into "formal," "informal," and "professional," each with its own set of categories which are explained in my data chapter, "Patterns of Assessment." ... Further hand coding yielded the categories I finally used in the chapter on "participation." ... (The author continues the discussion of data coding and other issues and then turns to the subject of values.)

Exposing Researcher Values

During this research I have continuously inspected my expectations and values as a continuing reminder of the role that values have in inquiry. ... Ongoing self-reflection in memos and discussions with mentors throughout the course of the study helped me identify and account for the interference of my assumptions in my study. ... For example, sometimes I was tempted to express findings about expectations in cause and effect terms. ...

I expose my values in my narrative as playing a significant role in my inquiry. In sharing my values in the introduction, and further here, I have attempted to take them into account as I share my data and analysis. For example, as a disability rights advocate, I have hoped that my research regarding students with disabilities would be a contribution toward achieving equality and full integration of people with disabilities. I remained aware of my bias against the self-contained setting, where four of the students in the study were based, in order to see what might actually benefit the students in that setting. I am aware that my bias is related to my advocacy stance against segregation and to the negative accounts of friends who have experienced segregated special education. I also had a prejudice against professional assessments along with the likelihood ... that I might be influenced by the contents if I read them early in the study. To counter inappropriate influence of this prejudice, I read the assessments at the end of my study and took a class in how to administer psychoeducational assessments. ...

As I listened to my informants, I was aware of my own assumption that students benefit from academic inclusion and that all students have the right to attain knowledge ... for my observations and interviews, I kept an open mind to the notion that special education settings do not preclude learning, may even enhance it, and that observing the special education academic experiences could also inform me about student engagement and how they [students] participated in the academic activities.

LEAVING THE FIELD

The process of leaving the field was gradual. I was learning less and less from observations by the end of spring. Completing ceasing the first school year observation was precipitated by the beginning of the university summer session and my assignment to spend all day in a suburban school as a student teacher. I was assigned to Trish's summer school class the second summer session and took notes on that experience. I visited her twice in the fall but was excluded from her general education classes due to overcrowding. Also in the fall, I spent two days with Tyrone. ... By then I had been analyzing data and felt the main thing lacking was the assessment of material from official records. Waiting until the following summer to look into

Here data analysis has moved on to axial coding, where the author is refining her categories and their interconnections.

In this section, the author reveals her biases and the strategies she used to counteract those biases. Regular conversations with her university advisor and others helped her identify assumptions she didn't initially realize she was making.

Here the author describes her bias in favor of inclusion, where students with disabilities learn in general education classrooms alongside their nondisabled peers, rather than in self-contained classrooms, in which students with disabilities are segregated from nondisabled students.

Here we discover why the author waited until the end of her study to look at school records.

Here the author is looking for disconfirming evidence, one effective strategy for minimizing the influence of a researcher's biases on data interpretation.

In grounded theory terminology, the author has probably saturated her categories at this point: Any additional information is shedding little or no new light on the subject matter.

the records proved wise, as I was able to find them a rich source of data. I actually eased my way out of the field (Bogdan & Biklen, 1992, pp. 104-105) rather than leaving, keeping contacts with many of my informants and calling to find out what is going on with a student or to clarify a question.

NOTE: Excerpt is from *Academic Engagement of High School Students with Significant Disabilities: A Competence-Oriented Interpretation* (pp. 18–30) by R. M. Smith, 1999, unpublished doctoral dissertation, Syracuse University, Syracuse, New York. Reprinted with permission.

Notice that the author didn't just disappear from the scene. Instead, she continued to maintain contact with her participants after her research was completed.

For Further Reading

- Agar, M. H. (1996). *The professional stranger: An informal introduction to ethnography* (2nd ed.). San Diego, CA: Academic Press.
- Barlow, D. H., & Nock, M. K. (2009). Why can't we be more idiographic in our research? *Perspectives on Psychological Science*, 4, 19–21.
- Bernard, H. R. (2000). *Social research methods: Qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.
- Birks, M., & Mills, J. (2011). *Grounded theory: A practical guide*. Thousand Oaks, CA: Sage.
- Bloomberg, L. D., & Volpe, M. F. (Eds.). (2008). *Completing your qualitative dissertation: A roadmap from beginning to end*. Thousand Oaks, CA: Sage.
- Butler-Kisber, L. (2010). *Qualitative inquiry: Thematic, narrative and arts-informed perspectives*. London: Sage.
- Chandler, S. (Ed.). (1992). Qualitative issues in educational research. *Theory into Practice*, 31, 87–186.
- Corbin, J., & Strauss, A. C. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Thousand Oaks, CA: Sage.
- Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage.
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Upper Saddle River, NJ: Pearson/Allyn and Bacon.
- Denzin, N. K., & Lincoln, Y. S. (2008). *Strategies of qualitative inquiry* (3rd ed.). Thousand Oaks, CA: Sage.
- Eisner, E. W. (1998). *The enlightened eye: Qualitative inquiry and the enhancement of educational practice*. Upper Saddle River, NJ: Prentice Hall.
- Fetterman, D. M. (2010). *Ethnography: Step by step* (3rd ed.). Thousand Oaks, CA: Sage.
- Glaser, B. (1992). *Basics of grounded theory analysis*. Mill Valley, CA: Sociology Press.
- Graneheim, U. H., & Lundman, B. (2003). Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today*, 24(2), 105–112.
- Grbich, C. (2007). *Qualitative data analysis*. Thousand Oaks, CA: Sage.
- Hammersley, M. (2008). *Questioning qualitative inquiry: Critical essays*. Thousand Oaks, CA: Sage.
- Hammersley, M., & Atkinson, P. (2007). *Ethnography: Principles in practice* (3rd ed.). New York: Routledge.
- Harch, A. J. (2006). *Early childhood qualitative research*. New York: Routledge.
- Heidegger, M. (2005). *Introduction to phenomenological research*. Bloomington: Indiana University Press.
- Hutchinson, S. (1988). Education and grounded theory. In R. R. Sherman & R. B. Webb (Eds.), *Qualitative research in education: Focus and methods* (pp. 123–140). London: Falmer.
- Jaccard, J., & Jacoby, J. (2010). *Theory construction and model-building skills*. New York: Guilford. (See Chapter 10, "Grounded and Emergent Theory.")
- Jorgensen, D. L. (2007). *Participant observation: A methodology for human studies*. Thousand Oaks, CA: Sage.
- King, N. (2010). *Interviews in qualitative research*. Thousand Oaks, CA: Sage.
- Latimer, J. (Ed.). (2003). *Advanced qualitative research for nursing*. New York: Wiley.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Thousand Oaks, CA: Sage.
- Locke, L. F. (1989). Qualitative research as a form of scientific inquiry in sport and physical education. *Research Quarterly for Exercise and Sport*, 60, 1–20.
- Merriam, S. B. (1995). What can you tell from an N of 1? Issues of validity and reliability in qualitative research. *PAACE Journal of Lifelong Learning*, 4, 51–60.
- Merriam, S. B. (1997). *Qualitative research and case study applications in education* (rev. ed.). New York: Wiley.
- Morgan, D. L. (1998). *The focus group guidebook*. Thousand Oaks, CA: Sage.
- Moustakas, C. (1994). *Phenomenological research methods*. Thousand Oaks, CA: Sage.
- Neuman, W. L. (2011). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Boston: Allyn & Bacon. (See Chapter 13, "Field Research and Focus Group Research," and Chapter 15, "Analysis of Qualitative Data.")
- Peshkin, A. (1988). Understanding complexity: A gift of qualitative research. *Anthropology and Education Quarterly*, 19, 416–424.
- Richardson, J. T. E. (1999). The concepts and methods of phenomenological research. *Review of Educational Research*, 69, 53–82.
- Sandelowski, M. (2000). Whatever happened to qualitative description? *Research in Nursing and Health*, 23, 334–340.
- Schram, T. H. (2006). *Conceptualizing and proposing qualitative research* (2nd ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.
- Shank, G. D. (2006). *Qualitative research: A personal skills approach* (2nd ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

- Smith, J. A., Flowers, P., & Larkin, M. (2009). *Interpretative phenomenological analysis: Theory, method and research*. London: Sage.
- Smith, M. L. (1987). Publishing qualitative research. *American Educational Research Journal*, 24, 173–183.
- Stake, R. E. (2006). *Multiple case study analysis*. New York: Guilford.
- Stake, R. E. (2010). *Qualitative research: Studying how things work*. New York: Guilford.
- Strauss, A., & Corbin, J. (1994). Grounded theory methodology: An overview. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 273–285). Thousand Oaks, CA: Sage.
- Wertz, F. J., Charmaz, K., McMullen, L. M., Josselson, R., Anderson, R., & McSpadden, E. (2011). *Five ways of doing qualitative analysis: Phenomenological psychology, grounded theory, discourse analysis, narrative research, and intuitive inquiry*. New York: Guilford.
- Wolcott, H. F. (1994). *Transforming qualitative data: Description, analysis, and interpretation*. Thousand Oaks, CA: Sage.
- Yin, R. K. (2008). *Case study research: Design and methods* (4th ed.). Thousand Oaks, CA: Sage.
- Yin, R. K. (2011). *Qualitative research from start to finish*. New York: Guilford.

ANSWERS TO THE CONCEPTUAL ANALYSIS EXERCISE “Choosing a Qualitative Research Design” on page 150

1. The researcher wants to learn about the general *culture* of an Elks chapter; hence an *ethnography* is most appropriate.
2. A *content analysis* is called for here—in particular, a systematic sampling and analysis of television commercials that are broadcast within a specified time period.
3. By focusing on three specific examples of a grassroots political party, the researcher is presumably intending to conduct a *multiple case study*.
4. The focus here is on how adolescents *perceive* their situation, making a *phenomenological study* especially relevant to the research problem.
5. Because the research question involves a process—human interaction—and very little literature exists to shed light on the question, a *grounded theory* study is probably in order here.

Now go to MyEducationLab at www.myeducationlab.com to take assessments to evaluate your mastery of chapter concepts. Assignments and Activities exercises and Building Skills in Reading Research exercises are also available to help you master concepts and skills. Feedback for many of these exercises is provided so that you can see why your answers are correct or incorrect.

7

Historical Research

Looking at a string of seemingly random events, the historical researcher develops a rational explanation for their sequence, speculates about possible cause-and-effect relationships among them, and draws inferences about the effects of events on individuals and the society in which they lived.

In and of itself, history consists of nothing more than an ever-flowing stream of events and continuing changes in human life and its institutions—its languages, customs, philosophies, religions, art, architecture, and so on. Historical research tries to make sense of this maelstrom. It considers the currents and countercurrents of present and past events, with the hope of discerning patterns that tie them all together. At its core, **historical research** deals with the *meaning of events*.

Many people have the impression that historical research involves gathering significant facts about a major event—perhaps a war, an economic depression, or the emergence of a new nation—and organizing these facts into a sequence, usually chronological. Such an enterprise may yield a historical narrative. It is not, however, true historical research. The heart of the historical method is, as with any other type of research, not the accumulation of the facts, but rather the *interpretation* of the facts. Interpretation of the data is central in all research. The task of the historical researcher is not merely to describe *what* events happened but to present a *factually supported rationale* to suggest *how* and *why* they may have happened.

Historical research is certainly not the domain of historians alone. On the contrary, it can be found in such disciplines as geography, anthropology, political science, economics, psychology, literature, and linguistics. For example, some social scientists engage in **comparative-historical research**, comparing historical events and processes across two or more societies or cultures, with the goal of identifying similarities, differences, and patterns that might possibly reflect cause-and-effect relationships.

As you will see, historical research is largely a qualitative endeavor, although historical researchers often make use of quantitative data as well.

Data Sources in Historical Research

In Chapter 4 we distinguished between *primary data* and *secondary data*, with the former being closer to the reality, or Truth, that the researcher ultimately wants to uncover. In historical research, this distinction is often referred to in terms of *primary sources* versus *secondary sources*. **Primary sources** are those that appeared first in time—in particular, when or soon after the events in question occurred. These sources take such diverse forms as letters, diaries, newspaper articles, sermons, laws, census reports, immigration records, probate documents, deeds, photographs, paintings, films, buildings, and tools. As an example, Matthew

McKenzie, a doctoral student in history at the University of New Hampshire, studied the impact of the Boston Marine Society on political decision making and scientific advancements in colonial America and in the early decades of the United States. McKenzie made extensive use of the society's minutes and other documents, as he explained in the following excerpt from his dissertation:

[A]s an organization of [sea] captains predicated upon fellowship and mutual aid and with a distinct role within the port [Boston Harbor], the Society went to great lengths to follow proper parliamentary procedures and to act only on decisions taken unanimously. As part of this process, the Society maintained meeting minutes recording the Society's (though not individuals') opinions, resolutions, and approved actions. Consequently, throughout its 250-year history, the society left committee reports, resolutions, and clear statements that reveal its collective will and motivations. These records allow historians to uncover not only what the organization did, but why. (McKenzie, 2003, pp. 2–3)

Another source of primary data, at least for events within the past few decades, can be found in interviews of people who participated in them. This approach is sometimes known as **narrative research** or **oral history**. A good example is Kevin Kearns's *Dublin Tenement Life* (1994), which pulls data from interviews with many residents of inner-city tenement buildings into a vivid description of inner-city Irish life in the early 1900s. We present a brief excerpt from Kearns's description of teenage courting practices to give you a taste of this approach:

[B]y age eighteen or thereabouts "marriage was their highest ambition," claims Peggy Pigott [a former tenement resident whom Kearns interviewed]. It was around this time that young women liked to go "clicking" in pairs. Clicking was an acceptable practice whereby respectable young women would stroll together along fashionable Dublin streets ostensibly window-viewing but in reality hoping to meet decent lads. When May Hanaphy [another interviewee] and one of her pals went clicking back in the 1920s, it was a perfectly proper way to meet a prospective husband:

Oh, clicking then was very popular. See, that's how flirting went on. That's how many a girl got her husband, going out at night time. Oh, you'd go out for that purpose at that time. We'd go clicking along mostly O'Connell Street or maybe down Henry Street, you know, slow walking . . . strolling, and two fellas come along and say 'there's two mots.' (Kearns, 1994, p. 46)

Interview data often give life to historical events. But just as is true in conducting any interview, the researcher must remember that participants' recollections are not always accurate. Only when several people recall events similarly can a researcher have reasonable confidence in what the interviews reveal. More generally, the guidelines for "Conducting a Productive Interview" presented in Chapter 6 are as applicable to historical researchers as to other qualitative researchers.

Historical researchers don't necessarily limit themselves to words, images, and objects; they often use *numbers* as well. For instance, they might draw inferences about people's interests during a particular time period by looking at the numbers of books on various topics that were sold during that period (Marius, 1989). Or they might examine the frequencies with which the Puritans of colonial America named their children after figures in the Bible, chart trends in these frequencies over time, and then speculate about what the trends might mean for religious practices and beliefs (Marius, 1989). In his study of the Boston Marine Society, McKenzie (2003) used early tax rolls to determine the wealth of society members, and he used society records of new members to show the decline in the society's popularity and influence in the early 1800s (see Figures 7.1 and 7.2). In research for his master's thesis in history, Peter Leavenworth (1998) used early real estate deeds to find patterns in land sales in New England during the 1600s, with a focus on land sold by Native Americans to British colonists. At one point in his analysis, Leavenworth plotted the frequency of land sales for each month of the year

FIGURE 7.1

Using tax rolls to determine the wealth of Boston Marine Society members

From *Vocational Science and the Politics of Independence: The Boston Marine Society, 1754–1812* (p. 25) by M. G. McKenzie, 2003, unpublished doctoral dissertation, University of New Hampshire, Durham. Reprinted with permission.

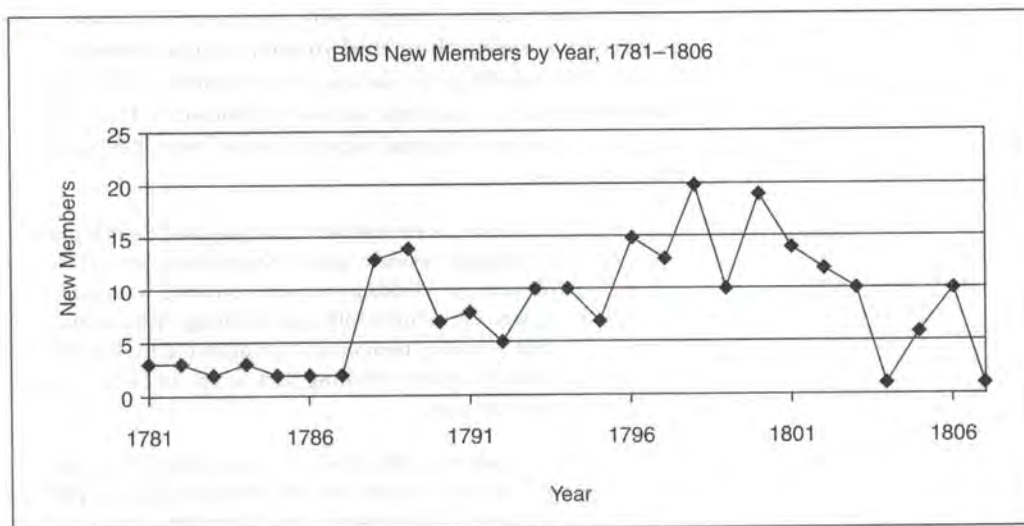
Wealth Bracket	No. in bracket	Assessed Tax in 1771 (£ s)	Proportion of Population
Top 20%	8	£37 12s–£46 13s	11%
2nd	6	£28 6s–£37 12s	8%
3rd	19	£18 18s–£28 4s	26%
4th	19	£9 10s–£18 16s	26%
5th	20	0–£9 8s	28%
Total Population	72		100%

Source: Pruitt, *Massachusetts Tax Evaluation*; and Baker, *Boston Marine Society*, 318–361

FIGURE 7.2

Using new membership data to reveal the decline in the Boston Marine Society's popularity and influence during the early 1800s

From *Vocational Science and the Politics of Independence: The Boston Marine Society, 1754–1812* (p. 198) by M. G. McKenzie, 2003, unpublished doctoral dissertation, University of New Hampshire, Durham. Reprinted with permission.



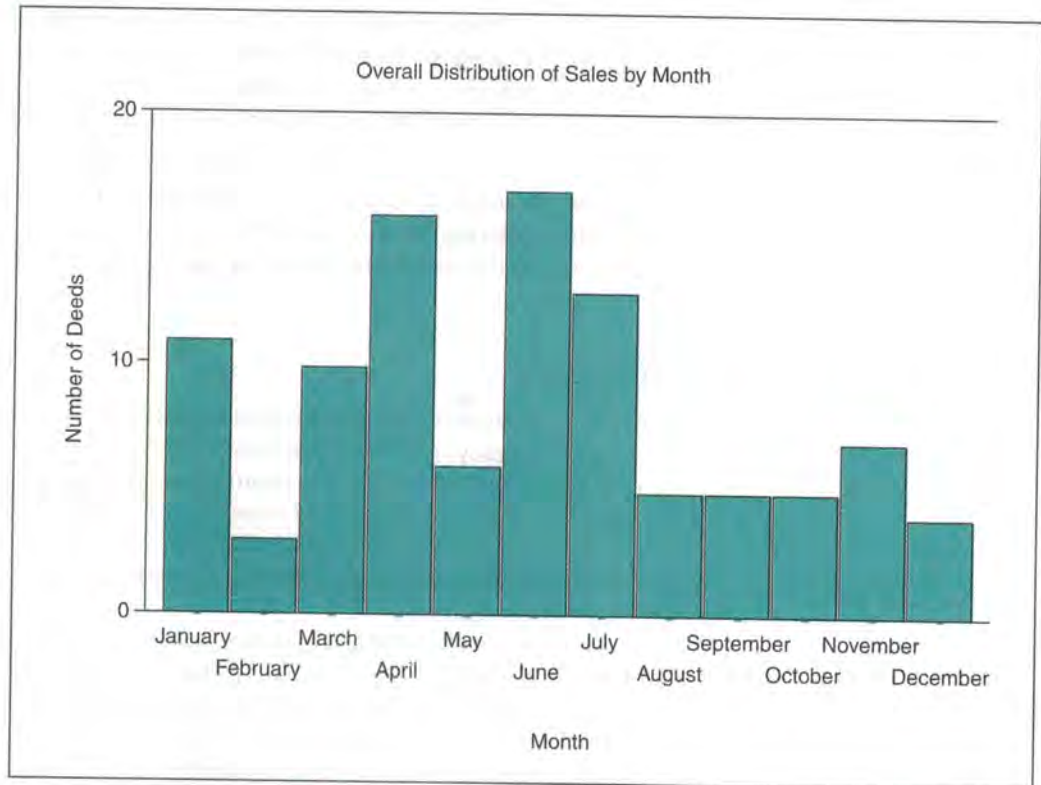
(see Figure 7.3). In the following excerpt from his thesis, he found considerable meaning in the month-to-month frequency data:

One last use of the deed index is to chart times of the year when transactions were more prevalent. . . . [T]he dips in overall Indian deed activity shown in February and May are consistent throughout the century. The February dip was either hunting- or weather-related, while the May dearth may have been either a time of Indian removal to summer habitations, or the period of spring fishing runs, or both. Many early accounts, including missionary John Eliot's, mention large annual spring gatherings of many bands at traditional fishing locations on the Merrimack and Piscataqua Rivers. When distribution is separated by decade, large spikes in land sales increasingly cluster in the spring later in the century. This may point to a growing native need for sustenance from the English marketplace after the hardships of the winter months, especially as their increasing proximity to white society did not raise their standard of living. Both Indian and English land sale patterns, not surprisingly, display a marked decrease at harvest time. (Leavenworth, 1998, pp. 88–89)

FIGURE 7.3

Using early deeds to track patterns in land sales, especially from Native Americans to British colonists, in New England during the 1600s

From "The Best Title That Indians Can Claim...": *Native Agency and Consent in the Transfer of Penacook-Pawtucket Land in the Seventeenth Century* (p. 88) by P. S. Leavenworth, 1998, unpublished master's thesis, University of New Hampshire, Durham. Reprinted with permission.



In the examples just presented, the raw numerical values sufficed for the researchers' purposes. In other instances, however, historical researchers may perform statistical analyses on the numerical data they collect; in such instances, historical research is truly a blend of qualitative and quantitative methodologies. An excellent resource for beginning historical researchers is Haskins and Jeffrey's *Understanding Quantitative History* (1990), which suggests many potentially useful sources of quantitative information and describes a variety of ways to analyze quantitative historical data.

In contrast with primary sources, **secondary sources** are the works of historians who have interpreted and written about primary sources. These include history textbooks, as well as more scholarly books and articles written about a particular event or time period. Secondary sources inevitably reflect the assumptions and biases of the people who wrote them. For this reason, good historical researchers rely on primary sources whenever possible.

Yet virtually *any* verbal report of an event—even a report that historians would consider to be a primary source—is apt to be somewhat distorted in line with a person's assumptions and biases. For example, a newspaper article describing a highly controversial event—perhaps an unprovoked attack that resulted in many human casualties—might very well be slanted in one direction or another, depending on the journalist's political leanings.

As noted in Chapter 4, researchers can probably never determine the ultimate, objective Truth that lies beneath a body of data. The historical researcher keeps this fact in mind, as illustrated in the following excerpt from a geography dissertation that analyzed changes in Jamaican agricultural practices during the early 1800s:

The sources of data are primarily printed documents published during the period of crisis, plus some surviving correspondence. Most important are island newspapers, the annual almanac, books, pamphlets, and government records. As with any historical data, [they] are incomplete in their coverage of contemporary events and may present a biased picture of the times. Some are also suspect as personal interpretations of interested parties participating in the events

described. These problems are not insurmountable so long as we are aware of them and interpret source information in terms of its contemporary context. In fact, the opinions and biases reflected in much of the data [are themselves] an important component in our portrait of adaptive change. (R. K. Ormrod, 1974, pp. 217–218)

Notice how the author acknowledges the probable presence of bias in some of his sources, even those published during the time period he was studying. In fact, he uses that bias to his advantage: as a way of getting a better handle on the prevailing perspectives of the time.

Collecting Historical Records



Historical researchers often find it helpful to make copies of the primary sources (letters, minutes, deeds, etc.) they find in archival collections at libraries, museums, historical societies, and the like. Many institutions provide photocopies for a small fee. An alternative approach is to photocopy the sources using a *digital camera*, which requires less light than photocopying and so is less likely to inflict environmental wear and tear on fragile documents. Photographs from digital cameras can be loaded directly onto a personal computer, allowing for easy cutting-and-pasting into a research report.

As researchers collect their documents, however, they should keep track of where each one came from, perhaps in a form similar to the footnotes they might eventually use to describe the documents in a final research report. For paper photocopies, small Post-it notes attached at a corner and written comments on the flip side are obvious strategies. When a digital camera is used, an ongoing log of what each shot is and where and when it has been taken is essential.

Online Databases for Historical Events



Some historical documents and records can be found only in the archives of various libraries, museums, and historical societies around the world, and you may have to travel a great distance to see them. But many others have been painstakingly captured in electronic form and made available online. For example, many university libraries subscribe to such online databases as the following:

- *Making of America*: Copies of thousands of books and tens of thousands of journal articles published in the 1800s.
- *Accessible Archives*: Copies of several American newspapers and magazines from the 1700s and 1800s.
- *New York Times Article Archive*: Every page of every issue of the *New York Times* since its first issue in 1851.
- *Ancestry Library*: A genealogical database that includes such documents as census records; birth, marriage, and death certificates; and immigration and naturalization records.

In addition, government websites offer many historical documents online. One especially helpful site is that of the U.S. Library of Congress (www.loc.gov), which provides documents related to both American and world history and also gives links to other helpful websites around the world.

Such sources open up many possibilities for researchers who might otherwise be restricted by location, disabilities, or other challenges. In addition, simple word processing functions (e.g., *copy* and *paste*) are invaluable time savers. As you locate relevant quotations, references, and so forth that you want to keep for future reference, you can highlight and copy desired sections and then paste them into your own database. If you do such copying and pasting, however, you must take appropriate precautions to avoid copyright infringement. Furthermore, remember that you must *not* use the words and ideas of others without proper citation.

PRACTICAL APPLICATION Handling Historical Data Systematically

In historical research, most of the data collected will be recorded in terms of hundreds, perhaps thousands, of notes. It is imperative that the researcher have some means of gathering and controlling the data in order to reap the greatest return from the innumerable hours spent in archives, document rooms, and libraries.

It is easy to read and take notes, but for many novice researchers, it is difficult to organize those notes into a format that facilitates interpretation. Historical data collection demands a systematic plan, not only for collecting data but also for retrieving and analyzing them. Before beginning historical research, therefore, you should have a specific plan for the acquisition, organization, storage, and retrieval of your data. Here we suggest two approaches you might take: a paper-and-pencil approach and a computerized approach.

A paper-and-pencil approach. If you have only limited computer skills, you might initially be tempted to make paper-and-pencil notes about the data collected. Such an approach is workable if your project is small and involves only a small amount of data. If you take a paper-and-pencil approach, you may want to modify the format presented in Figure 3.2 of Chapter 3 to fit the nature of historical data. For instance, your form might include such categories as "type of document," "date," "author(s)," "location," and so on.

Yet keep in mind the multidimensional nature of historical data. Any single historical artifact may lie simultaneously in the provinces of time, space, personality, and specific subject matter. Take, for example, a note you might make when studying Edgar Allan Poe's poem "Annabel Lee." Poe wrote the poem in 1849 while living in a small cottage in the Fordham section of the Bronx in what is now New York City. As a researcher, you might wish to look at the note within the contexts of (a) Poe's life in Fordham; (b) Poe's poetry written in 1849; (c) poems that Poe wrote about his wife, Virginia Clemm; and (d) all information collected about the poem "Annabel Lee."

If your research project involved a historical analysis of Poe's poetry, you would no doubt also want to look at each of his other poems from multiple angles. How, then, could you organize all of your paper-and-pencil notes? One strategy would be to make multiple copies, perhaps three or four, depending on your ultimate analysis of the data. For instance, you might want to organize the poems by (a) chronological order, (b) geographical location, (c) historical figures, and (d) specific subtopics within your overall research topic. Such an approach would obviously be both time- and space-consuming, and your photocopying expenses would quickly mount up. Furthermore, you would want to have backup-up copies of all the data you have collected. A more efficient alternative, then, is to use a computer to record and organize your data.



A computerized approach to collecting and organizing your data. Technology facilitates data collection in several ways. We begin with the note-taking process. Instead of using note-taking forms, take a cassette recorder to the library or archives or wherever else the data may be available. Dictate the data (notes) instead of transcribing them. Better still, take a laptop computer and either type or dictate your notes directly into it. (As noted in Chapter 6, some hardware and software lets you turn a laptop into a tape recorder.) When you return home, enter your notes into a database or word processing file.

From here on, the management of your data is limited only by the capabilities of your software. Word processing programs let you search through your documents for recurring words or phrases; you can then copy and paste these into another document. Many database software programs sort, alphabetize, create tables and summaries, and arrange data in an ascending or descending order by date and event. To get a sense of the various data management features your software package offers, browse through its paper or electronic manual. (You can typically find the electronic manual in a "Help" or similarly labeled pull-down menu at the top of your computer screen.) Alternatively, you might experiment with the many options in the software's pull-down menus and think creatively about how each feature might help you manage a large data set.

Evaluating and Interpreting Historical Data

After researchers have located historical data relevant to a research problem, they must decide what is fact and what is fiction. In other words, they must determine the *validity* of their data.

The data of historical research are subject to two types of evaluation. First, a researcher must judge whether a document or artifact is authentic. Second, if the item is indeed authentic, he or she must decide what it means. In these two situations, the researcher is reviewing the data to determine their **external evidence** and **internal evidence**, respectively. You may also see these concepts referred to as *external criticism* and *internal criticism*.

External Evidence

External evidence is primarily concerned with the question, Is the article genuine? Counterfeits and frauds are not uncommon, nor is their acceptance by the naive scholar and researcher unusual.

External evidence for the validity of a document is of paramount importance for the credibility of the research. Establishing authenticity of documents may in some cases involve carbon dating, handwriting analysis, identification of the type of ink and paper used, vocabulary usage and writing style, and other approaches. This aspect of historical methodology is a study in itself, and we cannot discuss it at length in a text as brief as this one.

Internal Evidence

Quite apart from the question of genuineness is the equally important question, What does it mean? When considering a manuscript or a statement, the researcher asks such questions as, What was the author trying to say? To what individuals or events do certain phrases refer? What interpretations can be extracted from the words?

Take a well-known utterance. The time is November 19, 1863. Abraham Lincoln is speaking at the dedication of a national cemetery in Gettysburg, Pennsylvania. In that brief but famous address, the president said, "But, in a larger sense, we cannot dedicate—we cannot consecrate—we cannot hallow this ground. The brave men, living and dead, who struggled here, have consecrated it, far above our poor power to add or detract."

What did Lincoln mean by "the brave men, living and dead"? Did he mean only the brave men of the Union forces? (We must remember that he was dedicating a Union cemetery.) Did he mean the brave Confederate men as well? Or did he mean brave men, indiscriminately, with no thought of North or South but merely of courage and valor? To a researcher studying the life of Abraham Lincoln, it is important to determine what Lincoln specifically meant by his words or, alternatively, whether he was intentionally being ambiguous.

The matter of internal evidence applies not only to articles from the distant past but to more contemporary documents as well. What does the decision of a court mean? What do the words of the decision convey about the intent and will of the court? The question comes up frequently in legal interpretation. In such instances, the primary question is, *What do the words mean?* This is the sole concern of internal evidence.

Considerable historical meaning can be found in graphic documents as well as verbal ones. For example, in his dissertation about the Boston Marine Society, McKenzie analyzed nautical charts of New England coastline created during the society's era. He noticed that one early map-maker, DesBarres—whose nautical charts were typically rendered in painstaking detail—had a glaring omission in a map of the Bar Harbor area of Maine. McKenzie interpreted the omission in light of political concerns of the time:

In at least one case, DesBarres consciously changed the shape of the coastline to suit imperial [British] needs, thus pitting local needs against imperial desires. In his chart of the coast of Maine from Frenchman Bay to Mosquito Harbor, DesBarres

failed to indicate Northeast Harbor, the best harbor in the region, or anything that might resemble a harbor along the southern coast of Mount Desert island. . . . DesBarres' omission was almost certainly intentional, as the rest of the island's features, including its topography, coves, and hazardous rocks, were laid out in DesBarres' characteristic detail, and in more detail than the rest of land areas on the chart. DesBarres most likely left this strategically important harbor out of consideration for military reasons. In this case, imperial concerns outweighed the need for accurate local charts for free commerce. (McKenzie, 2003, p. 79)

When interpreting historical data, the researcher must inevitably impose certain assumptions on them. For instance, when looking at the laws that a democratic government created during a particular era, the researcher may assume that the laws reflected the *needs and beliefs of the majority of voting citizens* (Marius, 1989). Or, more generally, when tracking the course of events in a particular society, a researcher may assume that the events reflected *economic or social progress* (Breisach, 1994). And many historical researchers from Western countries are apt to interpret events from a Western cultural perspective: Practices consistent with Western norms and values are seen as being in some way "better" than other practices (Myrdal, 1973).

A good historical researcher will identify, explicitly and concretely, the assumptions that guide his or her interpretation of historical data. As an example, we look once again at the previously excerpted geography dissertation, this time focusing on the researcher's interpretation of events during a crisis in Jamaican sugar planting practices during the early 19th century:

Our interpretation [of the data] will depend upon two primary assumptions: (1) an adaptation imperative existed which demanded that the island society respond to the events threatening its pattern of livelihood, and (2) most of the behaviors involved in the response, bounded by the constraints inherent in the functioning cultural ecosystem, were goal directed. These assumptions lead us to expect an orderly response to crisis rather than a random one and lead us to seek a behavior system which sought to relieve the stresses on the society. Although such a behavior system would function in a probabilistic manner rather than as a closely determined one, we should nevertheless be able to construct an orderly framework of interpretation around our data. (R. K. Ormrod, 1974, p. 227)

Psychological or Conceptual Historical Research

Thus far, we have discussed conventional historical research—the study of significant events and the individuals who played important roles in them. But this chapter would not be complete without a brief discussion of another type of historical research—namely, that concerned with the origin, development, and influence of ideas and concepts. Ideas and concepts can influence the course of history just as surely as events and people do.

As an example, the idea of *democracy* was born in Greece; its development has run parallel to the events of the Greco-Roman world, the Middle Ages, and modern times. Over the years the idea evolved into such concepts as *representative government* and *political campaign*. The initial concept of democracy that began in ancient Athens is, perhaps, found in its purest form today in the New England town meeting.

Consider other key ideas that have guided the course of civilization: capitalism, socialism, rationalism, individualism, communism, postmodernism. Each of them has its own developmental history, which is just as "real" as the history of Western civilization, the United States, or your hometown. Search any comprehensive paper or electronic encyclopedia for some of the principal ideas of civilization just listed, and you may well find discussions of how these ideas have evolved over time.

Look, too, at Arnold Toynbee's monumental *A Study of History* (1939–1961). Here, you will find not the traditional approach to history—the description and interpretation of

events—so much as the dynamic ideas that have powered the histories of nations and civilizations and that have been instrumental in bringing about cataclysmic changes in those histories.

Searching for Roots

In Chapter 4 we mentioned the research of John Livingston Lowes, which is presented in *The Road to Xanadu* (1927, 1955). Lowes's book is remarkable because it is, in a sense, research in reverse. In it, Lowes searches for "the genesis of two of the most remarkable poems in English, 'The Rime of the Ancient Mariner' and 'Kubla Khan'" (Lowes, 1927, p. 3).

This type of research, which is the counterpart of a genealogical search of one's family origins, begins with such questions as, Where did it come from? and How did it all begin? This is precisely the type of research that astronomers and astrophysicists conduct to try to account for the creation of the universe. One does not have to design a research project on such a cosmic scale to engage in the same kind of detective work that reels backward in search of answers instead of forward, as is usually done.

The process of beginning with a phenomenon and going backward in time to identify possible causal factors is sometimes called *ex post facto research*. We consider *ex post facto* within a different context—how people's backgrounds may potentially affect their current characteristics and behaviors—in Chapter 9.

PRACTICAL APPLICATION Historical Research Writing

Written accounts of historical research vary widely, depending on the researcher and his or her writing style. Research reports dealing with historical research need not be dull, and in fact those historians whose works often appear on best-seller lists infuse their descriptions of history with many colorful events and interesting personalities.

GUIDELINES Writing the Historical Research Report

Many of the suggestions we offer about writing in Chapters 1 and 5 apply to historical research as well as to any other type of research. In addition, Marius (1989) has offered several useful *rules for argument* that you should keep in mind when, in particular, you are writing about a historical research study.

1. *State your own argument early in the game.* Remember, you are not only presenting the data but you are also *interpreting* it. You should be up-front about your interpretation and not keep your readers guessing.
2. *Provide examples to support any assertion you make.* You make a more convincing case when you give examples of data that lend credence to your position.
3. *Give the fairest possible treatment of any perspectives different from your own.* You may very well be presenting an interpretation that differs from those of other scholars. Describe competing interpretations and provide evidence that supports them, as well as evidence that casts doubt on them.
4. *Point out the weaknesses of your own argument.* Better to shoot holes in your own case than to have others do it for you. You portray yourself as a credible researcher when you appear to be objective—rather than blindly one-sided—in your analysis and interpretation of your data.

With regard to this fourth guideline, we should remind you that your research project, while answering your initial research question, may also yield new, unanswered questions. You can turn any inconsistent findings that you uncover into "unresolved issues" or "suggestions for future research."

A Sample Dissertation

We now return to Matthew McKenzie's dissertation on the Boston Marine Society, which McKenzie completed for his doctoral degree in history at the University of New Hampshire (McKenzie, 2003). We have already shown brief snippets of the dissertation to illustrate certain aspects of historical research, but we now present a larger chunk with a running commentary. You will see that, overall, the dissertation has a different feel to it than the proposal and dissertation excerpts that appear in previous chapters. Unlike those earlier excerpts, this dissertation is written as a *narrative*, with historical events and interpretations seamlessly interwoven throughout the discussion. McKenzie spent an extra year overhauling major sections of his dissertation, and his efforts show clearly in the quality of his writing. Notice, too, how he used footnotes to identify his sources, reflecting the *Chicago* style that historians typically use in their research reports (more about various styles in Chapter 12).

Dissertation ANALYSIS 4

INTRODUCTION

In the spring of 1755, Captain Hector McNeill was in command of a merchant vessel in a small flotilla conveying an army up the Bay of Fundy. The fleet had left Boston a few days before with the task of safely delivering 2,000 New England soldiers to fight against their French imperial rivals at Fort Beaussejour. As the fleet sailed along the current-swept, rocky shores, [Colonel] Robert Monckton worried about the fate of his army. Back in Boston, there had been almost no charts for him to consult, and even fewer descriptions of the currents and tides that made this region so dangerous. Moreover, his and his army's fate rested in the hands of a few Boston merchant skippers, like Captain McNeill, none of whom likely knew the latest and best techniques in navigation.

Despite his fears, however, and the dangerous shoals and hazardous headlands, the fleet proceeded safely. When Monckton approached McNeill about their progress, curious as to how a colonial trading skipper could successfully undertake such a hazardous job, McNeill showed him information which no British commander in North America or London knew existed. Trading along the coast, McNeill had collected five years of nautical observations, including (presumably) tides, currents, coastal descriptions, and manuscript drawings. From these observations, McNeill had drawn a chart covering the coast from Cape Cod to Cape St. Mary's including the Bay of Fundy. McNeill's chart impressed the British commander. And shortly after the Boston skipper safely delivered his regiments, Monckton dislodged the French from Beaussejour.¹

McNeill was not alone in his interest in marine cartography in New England. In 1760, he joined a group of master mariners in Boston, called the Boston Marine Society (BMS), which had also been systematically collecting navigational observations since 1754. Both McNeill and the Marine Society understood that local navigational knowledge carried commercial, political, and imperial opportunities. Consequently, when the organization united senior captains for mutual aid, they also recognized that

Comments

In Chapter 6, we suggested that you always state your research problem at the very beginning of your research proposal. In a research report, however, researchers often begin with a few paragraphs of background information that provides a context for the research problem. A common strategy in historical writing is to begin with a story—a real-life drama of sorts—that draws readers in and motivates them to continue reading.

Here we get a glimpse of what will be one major thrust of the dissertation: describing and tracking the nature of early nautical charts, whose use and promotion were partly attributable to efforts of the Boston Marine Society.

Notice the use of footnotes to identify the sources for certain statements. The author is using the style required by the Journal of American History (available online on the journal's Website at www.journalofamericanhistory.org). Footnotes are also consistent with the Chicago Manual of Style (2010), which historians typically follow.

¹ Hector McNeill to Lord Colville, January 17, 1763, Boston Marine Society Papers (Massachusetts Historical Society, Boston, Mass.).

they stood in an important position between London imperial agents in North America and the coastline that interested them. Furthermore, they were actively collecting data as every member returned to Boston—a feature that they would try to barter for greater influence in Boston and within the [British] Empire.

Historians are fortunate in the Marine Society's meticulous record keeping and parliamentary procedure. Two key issues help modern researchers see the society's collective will and motivation. First, as membership was limited to captains alone, the Society was self-conscious that they spoke as an elite body in Boston's maritime community. Second, as an organization of captains predicated upon fellowship and mutual aid and with a distinct role within the port, the Society went to great lengths to follow proper parliamentary procedures and to act only on decisions taken unanimously. As part of this process, the Society maintained meeting minutes recording the Society's (though not individuals') opinions, resolutions, and approved actions. Consequently, throughout its 250-year history, the society left committee reports, resolutions, and clear statements that reveal its collective will and motivations. These records allow historians to uncover not only what the organization did, but why.

This is not the first study of the Boston Marine Society. Earlier studies of the Marine Society have cataloged in some detail the work the Marine Society undertook during its long history. Nathaniel Spooner stitched together a rough narrative in his 1879 *Gleanings of the Boston Marine Society* (Boston, 1879, 1999). In 1982, William A. Baker's *A History of the Boston Marine Society* (Boston, 1982) integrated the Marine Society's history more closely with changes in Boston politics and economics and assembled systematic information on the society's more than 3,000 members. Both of these works greatly aided the project that follows. Yet neither delved into the society's influence upon the history of American science, [and] with the exception of Baker's study of the Society during the American Revolution, neither Baker nor Spooner were interested in examining how the society operated as an active agent in Boston's historical development.

This study seeks to examine the society within the context of the history of American science. Academic centers and learned societies have been the focus for most considerations of American science because of their prominence in the nineteenth and twentieth centuries. The Marine Society's scientific interests indicate, however, that colonial groups could and did develop their own scientific agenda that they pursued through methods adapted from common vocational practices. In doing so, the Marine Society's navigational work draws important parallels to the history of colonial science in other areas during the late eighteenth century. In the simplest form, I argue that colonial Boston shipmasters were not dependent upon learned societies for their navigational research needs. Rather, they adapted their mutual aid society and developed methodologies to collect navigational observations, analyze them for reliability and accuracy, and in a few cases, publish their findings for the benefit of the community. Furthermore, given the close ties between seafaring, economic growth and political influence in a mercantile economy, the Marine Society's work in navigational research granted them social and political influence in Boston. With this added influence—power would be too strong a term for it—the Marine Society tried to stabilize post-Revolutionary Boston politics, and to legitimate their efforts to become one of the

Notice the smooth flow of the narrative from one event to another. In our experience, narrative writing is more challenging than traditional "scientific" report writing. However, when well executed (as is the case here), narrative reports are also more engaging than scientific reports tend to be.

Here we see the context in which the discussion of the society's minutes (excerpted earlier in the chapter) appeared.

Here the author explains how his own research extends the boundaries of what is known and believed about the Boston Marine Society's role in American history.

Here the author also explains how his research represents a divergence from the traditional approach to the history of science: Rather than studying the effects of traditional academic groups (universities and academic organizations), he is studying the impact of a less academic, yet definitely influential, group.

The author makes his central hypothesis clear at this point: He believes that early shipmasters relied on one another rather than on traditional scientific investigations to get the information they needed to travel safely along the northern Atlantic coastline.

town's new elites. Ultimately, the Marine Society lost its political influence as changes in navigational research, shifts in Boston and national politics, and new market centers for scientific information combined to weaken the society's position in both the political and navigational research world.

The Marine Society gives us a glimpse of the rise and fall of what I call "vocational science." In many previous studies discussed below, science and research were considered as a purely intellectual—"academic"—exercise, centered in learned academies, universities, and laboratories. I argue, to the contrary, that those who used navigation to carry their vessels safely into port, and expanded navigational knowledge, pursued science just as much as those who approached navigation from theoretical understandings of geodesy, mathematical astronomy, and spherical trigonometry. Whether using complex mathematical models to develop an absolute understanding of coastal features, or using piloting techniques, rule of thumb guidelines, simple instruments, and best-as-possible guess-work, both vocational and academic researchers formed part of a larger process by which the knowledge of New England's coast expanded.

The idea of vocational science also highlights an important mechanism by which specific groups used science to shore up their economic, social and political positions within their local area. While most prior work on American science has shown how the pursuit of scientific knowledge translated into improved cultural and social reputation, most have seen these efforts as a neutral desire to expand humanity's understanding of the world. Yet in this case, engagement in scientific research carried immediate economic, political and social benefits that were anything but neutral. As Joyce Chaplin has shown, colonial Carolina low-country planters sent botanical specimens to the Royal Society and the Royal Society of Arts in exchange for agricultural innovations. These innovations—seeds, water control mechanisms, and processing machinery—helped them secure political control over Carolina politics during the Early Republic and helped create the land-owning elite of the Ante-bellum south.² James McClellan argues that while French planters in Saint Domingue did not embrace science as openly as their Carolina counterparts, science did serve the mercantilist interests of the state, and helped perpetuate slavery in the French Caribbean.³ Finally, John Lauritz Larson has shown that experimental engineering designs for locks, dams, and internal waterways promised America's post-Revolutionary elite a means to promote private improvement schemes with public funds and in the face of public opposition.⁴ In all these situations, science—whether tied to European centers or not—worked to bolster a specific group's local political and economic positions. Not pursued solely for knowledge in its own right, science expanded knowledge of the natural world, yet at the same time advanced specific interests.

He posits a second hypothesis as well: The society's significant involvement in the local economy gave it considerable influence in early Boston.

The author introduces a new concept—vocational science—to describe the phenomenon he uncovers in his research.

The author contrasts his own viewpoint with more traditional views.

The author argues convincingly that, contrary to the popular perception of scientists as individuals who are more concerned about the general quest for knowledge than about their personal needs, these "vocational scientists" often had fairly self-serving motives at the root of their endeavors. He draws analogies to advancements in other locations and other times, where people may have been equally self-promoting. In doing so, he situates his research within a larger body of research literature that has preceded his own work.

²Joyce Chaplin, *An Anxious Pursuit: Agricultural Innovation and Modernity in the Lower South, 1730–1815* (Chapel Hill, 1993), 131–142.

³James E. McClellan III, *Colonialism and Science: Saint Domingue in the Old Regime* (Baltimore, 1993), 9, 289–292.

⁴John Lauritz Larson, *Internal Improvement: National Public Works and the Promise of Popular Government in the Early United States* (Chapel Hill, 2001), 1–37.

Readers will find the terms “science,” “navigational knowledge,” and “research” used quite liberally and perhaps over-interchangeably in the pages that follow. This is intentional. The structured and distinct practices that we associate with science today had yet to develop in the second half of the eighteenth century. The lines between “amateur,” “practitioner,” and “interested gentleman” were blurry to say the least. As others have shown, to impose such categories on inquiries into the natural world and the inquirers themselves clouds more than clarifies. Only after science underwent dramatic changes in the early nineteenth century would science have such clear structures.⁵

[The report continues with a discussion of earlier researchers’ explanations of the interplay among science, politics, and social dynamics in colonial American and the early decades following the American Revolution.]

⁵See McClellan, *Colonialism and Science*, 7; and Roy MacLeod, “On visiting the Moving Metropolis: Reflections on the Architecture of Imperial Science,” in *Scientific Aspects of European Expansion*, ed. William K. Storey (Hampshire, 1996), 24–27.

NOTE: Excerpt is from *Vocational Science and the Politics of Independence: The Boston Marine Society, 1754–1812* (pp. 1–6), by M. G. McKenzie, 2003, unpublished doctoral dissertation, University of New Hampshire, Durham. Reprinted with permission.

Here the author anticipates and addresses a potential source of confusion for his readers. In particular, he provides a reasonable rationale for why he will use several terms interchangeably.

We have said enough about the historical method for the purposes of the beginning researcher. Beginning with the next chapter, we turn to an entirely different approach to the discovery of knowledge. Whereas Chapters 6 and 7 have dealt with qualitative methodologies, Chapters 8 and 9 explore quantitative research, and Chapter 10 explains how qualitative and quantitative approaches might be combined in *mixed-methods* studies.

For Further Reading

- Barzun, J., & Graff, H. (2003). *The modern researcher* (6th ed.). Belmont, CA: Wadsworth. (This text provides a comprehensive survey of historians’ work.)
- Breisach, E. (2008). *Historiography: Ancient, medieval, and modern* (3rd ed.). Chicago: University of Chicago Press.
- Brundage, A. (2008). *Going to the sources: A guide to historical research and writing* (4th ed.). Wheeling, IL: Harlan Davidson.
- Clive, J. (1989). *Not by fact alone: Essays on the writing and reading of history*. New York: Knopf.
- Creswell, J. W. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (3rd ed.). Upper Saddle River, NJ: Pearson/Merrill/Prentice Hall. (See Chapter 16, “Narrative Research Designs.”)
- Edson, C. H. (1986). Our past and present: Historical inquiry in education. *Journal of Thought*, 21, 13–27.
- Errante, A. (2000). But sometimes you’re not part of the story: Oral histories and ways of remembering and telling. *Educational Researcher*, 29(2), 16–27.
- Floud, R. (2006). *An introduction to quantitative methods for historians*. New York: Routledge.
- Gilmore, M. B. (1992). Historians, books, computers, and the library. *Library Trends*, 40, 667–686.
- Gray, W. (1991). *Historian’s handbook: A key to the study and writing of history* (2nd ed.). Prospect Heights, IL: Waveland.
- Haskins, L., & Jeffrey, K. (1990). *Understanding quantitative history*. New York: McGraw-Hill.
- Hill, M. R. (1993). *Archival strategies and techniques*. Thousand Oaks, CA: Sage.
- Janesick, V. J. (2010). *Oral history for the qualitative researcher*. New York: Guilford.
- Knowles, A. K. (Ed.). (2008). *Placing history: How maps, spatial data, and GIS are changing historical scholarship*. Redlands, CA: ESRI Press.
- Lewis, M. J., & Lloyd-Jones, R. (2009). *Using computers in history: A practical guide to data presentation, analysis, and the Internet* (2nd ed.). New York: Routledge.
- Marius, R., & Page, M. E. (2006). *A short guide to writing about history* (6th ed.). New York: Longman.
- Neuman, W. L. (2011). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Boston: Allyn & Bacon. (See Chapter 14, “Historical-Comparative Research.”)
- Novick, P. (1988). *That noble dream: The “objectivity question” and the American historical profession*. Cambridge, England: Cambridge University Press.
- Slavens, T. P. (1994). *Sources of information for historical research*. New York: Neal-Schuman.

Now go to MyEducationLab at www.myeducationlab.com to take assessments to evaluate your mastery of chapter concepts. Assignments and Activities exercises and Building Skills in Reading Research exercises are also available to help you master concepts and skills. Feedback for many of these exercises is provided so that you can see why your answers are correct or incorrect.

8

Descriptive Research

To behold is to look beyond the fact; to observe, to go beyond the observation. Look at the world of people and objects around you, and you will be overwhelmed by what you see. But select from that world a well-chosen few, and observe them with insight, and they will tell you a great deal.

In this chapter we discuss types of quantitative study that fall under the broad heading *descriptive quantitative research*. This general category of research designs involves either identifying the characteristics of an observed phenomenon or exploring possible associations among two or more phenomena. In every case, descriptive research examines a situation *as it is*. It does not involve changing or modifying the situation under investigation, nor is it intended to determine cause-and-effect relationships.

As you proceed through the chapter, you will find several strategies—sampling, making observations, interviewing—that you encountered previously in the discussion of qualitative research in Chapter 6. This is old news, you might think. On the contrary, such strategies take on a very different form when we want them to yield quantitative data.

Descriptive Research Designs

In the next few pages, we describe observation studies, correlational research, developmental designs, and survey research, all of which yield quantitative information that can be summarized through statistical analyses. We devote a significant portion of the chapter to survey research, because this approach is used quite frequently in such diverse disciplines as business, government, public health, sociology, and education.

Observation Studies

In the qualitative studies described in Chapter 6, observations are usually recorded in great detail, perhaps with field notes or videotapes that capture the wide variety of ways in which people or other animal species act and interact. From these data, the researcher constructs a complex yet integrated picture of how certain humans or nonhumans spend their time.

In *quantitative* research, however, an **observation study** is quite different. Perhaps most importantly, an observation study does not necessarily involve members of the animal kingdom. Certainly it might involve humans or other animals. But it might, instead, be aimed at studying plant species, nonliving objects (e.g., rock formations, soil samples), or dynamic physical phenomena (e.g., weather patterns, black holes).

In addition, a quantitative observation study tends to have a particular, prespecified focus. When human beings are the topic of study, the focus is typically on a certain aspect of behavior. Furthermore, the behavior is quantified in some way. In some situations, each occurrence of the behavior is *counted* to determine its overall frequency. In other situations, the behavior is *rated* for accuracy, intensity, maturity, or some other dimension. But regardless of approach, the researcher strives to be *as objective as possible* in assessing the behavior.

being studied. To maintain such objectivity, he or she is likely to use strategies such as the following:

- Define the behavior being studied in such a precise, concrete manner that the behavior is easily recognized when it occurs.
- Divide the observation period into small segments and then record whether the behavior does or does not occur during each segment. (Each segment might be 30 seconds, 5 minutes, 15 minutes, or whatever other time span is suitable for the behavior being observed.)
- Use a rating scale to evaluate the behavior in terms of specific dimensions (more about rating scales a bit later in the chapter).
- Have two or three people rate the same behavior independently, without knowledge of one another's ratings.
- Train the rater(s) to use specific criteria when counting or evaluating the behavior, and continue training until consistent ratings are obtained for any single occurrence of the behavior.

A study by Kontos (1999) can give you a flavor for what a researcher might do in an observation study. Kontos's research question was this: What roles do preschool teachers adopt during children's free-play periods? (She asked the question within the context of theoretical issues that are irrelevant to our purposes here.) The study took place during free-play sessions in Head Start classrooms, where 40 preschool teachers wore cordless microphones that transmitted what they said (and what people near them said as well) to a remote audiotape recorder. Each teacher was audiotaped for 15 minutes on each of two different days. Following data collection, the tapes were transcribed and broken into 1-minute segments. Each segment was coded in terms of the primary role the teacher assumed during that time, with five possible roles being identified: *interviewer* (talking with children about issues unrelated to a play activity), *stage manager* (helping children get ready to engage in a play activity), *play enhancer/playmate* (joining a play activity in some way), *safety/behavior monitor* (managing children's behavior), or *uninvolved* (not attending to the children's activities in any manner). Two research assistants were trained in using this coding scheme until they were consistent in their judgments at least 90% of the time, indicating a reasonably high *interrater reliability*. They then independently coded each of the 1-minute segments and discussed any segments on which they disagreed, eventually reaching consensus on all segments. (The researcher found, among other things, that teachers' behaviors were to some degree a function of the activities in which the children were engaging. Her conclusions, like her consideration of theoretical issues, go beyond the scope of this book.)

As should be clear from the preceding example, an observation study involves considerable advance planning, meticulous attention to detail, a great deal of time, and, often, the help of one or more research assistants. Furthermore, a pilot study is essential for ironing out any wrinkles in identifying and classifying the behavior(s) or other characteristic(s) under investigation. Embarking on a full-fledged study without first pilot-testing the methodology can result in many hours of wasted time.

Ultimately, an observation study can yield data that portray much of the richness and complexity of human behavior. In some situations, then, it provides a quantitative alternative to such approaches as ethnographies and grounded theory studies.

Correlational Research

A **correlational study** examines the extent to which differences in one characteristic or variable are related to differences in one or more *other* characteristics or variables. A **correlation** exists if, when one variable increases, another variable either increases or decreases in a somewhat predictable fashion.

In correlational studies, researchers gather data about two or more characteristics for a particular group of people or other appropriate units of study. These data are numbers that reflect specific measurements of the characteristics in question. When human beings are the focus of investigation, the data might be test scores, ratings assigned by an expert observer, or frequencies of certain behaviors. Data in animal studies, too, might be frequencies of

particular behaviors, but alternatively they might be fertility rates, metabolic processes, or measures of health and longevity. Data in studies of plants, inanimate objects, or dynamic physical phenomena might be measures of growth, chemical reactions, density, temperature, or virtually any other characteristic that human measurement instruments can assess with some objectivity. Whatever the nature of the data, at least two different characteristics are measured in order to determine whether and in what way these characteristics are interrelated.

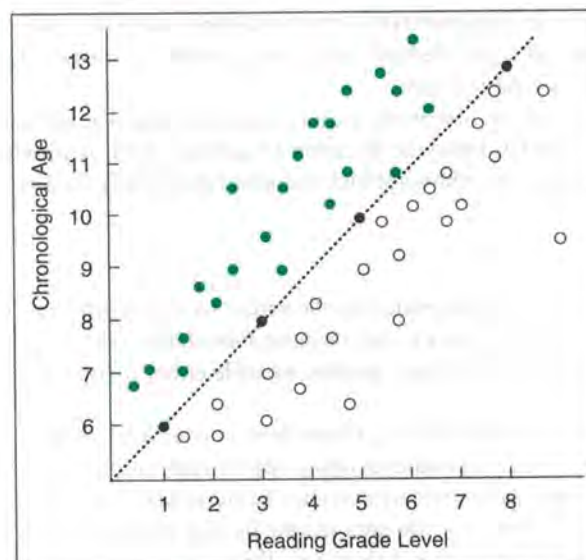
An example of a correlational study may be helpful. As you well know, as children grow older, they become better readers. In other words, there is a *correlation* between age and reading ability. Imagine that a researcher has a sample of 50 children and knows two things about these children: their age and their scores on a reading achievement test (the test scores indicate the approximate “grade level” at which the children are reading). The researcher might plot the data on a **scatter plot** (also known as a *scattergram*) to allow a visual inspection of the relationship between the two variables. Figure 8.1 presents this hypothetical scatter plot. Chronological age is on the vertical axis (the *ordinate*) of the graph, and reading level is on the horizontal axis (the *abscissa*). Each dot represents a particular child; its placement on the scatter plot indicates both the child’s age and his or her reading level.

If age and reading ability were two completely unrelated characteristics, the dots would be scattered all over the graph in a seemingly random manner. When the dots instead form a rough elliptical shape (as the dots in Figure 8.1 do) or perhaps a skinnier sausage shape, then we know that the two characteristics are correlated to some degree. The diagonal line running through the middle of the dots in Figure 8.1—sometimes called the *line of regression*—reflects a hypothetical perfect correlation between age and reading level; if all the dots fell on this line, a child’s age would tell us *exactly* what the child’s reading level is. In actuality, only four dots—the solid black ones—fall on the line. Some dots lie below the line, showing children whose reading level is, relatively speaking, advanced for their age (these children are designated by hollow black dots). Other dots lie above the line, indicating children who are lagging a bit in reading (these children are designated by solid blue dots).

As we examine the scatter plot, we can say several things about it. First, we can *describe* the homogeneity or heterogeneity of the two variables—the extent to which the children are similar to or different from one another with respect to age and reading level. For instance, if the data were to include only children of ages 6 and 7, we would have greater homogeneity with respect to age than would be the case for a sample of children ages 6 through 13. Second, we can *describe* the degree to which the two variables are intercorrelated, perhaps by computing a statistic known as a *correlation coefficient* (Chapter 11 provides details). But third—and most

FIGURE 8.1

Example of a scatter plot:
Correlation between age
and reading level



importantly—we can *interpret* these data and give them meaning. For instance, the upward trend of the dots from left to right tells us that as children grow older, their reading level improves. This correlation enables us to use a child's age to estimate—that is, to *predict*—the child's reading level to some degree.

A Caution about Interpreting Correlational Results

In all correlational studies, be alert for faulty logic. When two variables are correlated, researchers sometimes conclude that one of the variables must in some way influence the other. In some instances, such an influence may indeed exist; for example, chronological age—or at least the amount of experience that one's age reflects—almost certainly has a direct bearing on children's mental development, including their reading ability. But ultimately we can never infer a cause-and-effect relationship on the basis of correlation alone. Simply put, *correlation does not, in and of itself, indicate causation*.

Take a logically absurd yet statistically demonstrable example. We could conceivably find a statistical correlation between the number of elephants in Thailand in any given year and the size of the Florida orange crop the same year. The facts may be very clear: As the size of the elephant population increases over time, the Florida orange crop also increases. Yet it is ludicrous to think that because we can show a positive correlation there must therefore be a causal bond at the root of the relationship. There is no connection whatsoever between the elephant population in Thailand and the production of oranges in Florida—at least none that we authors can think of! The correlation is simply a fluke and has no meaning.

In the extreme situation of the elephants–oranges correlation, the faulty logic is readily apparent. Yet we often see similarly faulty reasoning proposed in correlational research reports. For instance, imagine that a researcher finds a correlation between family income level and children's school performance: On average, the lower a child's family income is, the lower the child's academic grade-point-average (GPA) is likely to be. It would be all too easy to draw the conclusion that socioeconomic status directly *affects* (i.e., has a causal influence on) academic achievement. We might think that *because* family paychecks, family living conditions, and so forth are below average, the achievement of the boys and girls of such families is also below par. If a family's economic status could be improved, we conclude, the school achievement of the family's children would also improve.

No, no, no! We *cannot* make an inference about causation on the basis of correlated data alone. It's possible that paycheck size does have an impact on children's grades, but it's equally possible that it does not. Perhaps, instead, an undetermined third variable—maybe parents' education levels or maybe the degree to which family members are victims to society-wide discrimination—influences both the size of the family paycheck and the children's school performance.

If we were to infer that socioeconomic status directly affects academic achievement, not only would we be going far beyond the data we have but we would also have trouble accounting for all of the world's intellectual giants, some of whom have grown up in severe poverty. For example, Robert Burns, a great Scottish poet, had very little formal schooling as a result of his family's poverty. George Washington Carver, a famous American botanist, was born of slaves. The father of Franz Schubert, the Austrian composer, was a peasant, and his mother was a cook; Schubert himself lived in poverty most of his life. The case of Abraham Lincoln, born in a log cabin, walking miles to borrow a book, also discredits the impoverished environment-deprived child notion.

The data may not lie, but the causal conclusions we draw from the data may, at times, be extremely suspect. Nevertheless, a good researcher must not be content to stop at the point of finding a correlational relationship, because *beneath the correlation* lie some potentially quite interesting dynamics. One way to explore these dynamics is through *structural equation modeling* (SEM), a statistical procedure we describe briefly in Table 11.5 in Chapter 11. Another approach—one that can yield more solid conclusions about cause-and-effect relationships—is to follow up a correlational study with one or more experimental studies (described in Chapter 9) to test various hypotheses about what causes what.

Developmental Designs

Earlier we presented a hypothetical example of how children's ages might correlate with their reading achievement levels. Oftentimes when researchers want to study how a particular characteristic changes as people grow older, they use one of two developmental designs, either a cross-sectional study or a longitudinal study.

In a **cross-sectional study**, people from several different age-groups are sampled and compared. For instance, a developmental psychologist might study the nature of friendships for children at ages 4, 8, 12, and 16. A gerontologist might consider how retired people in their 70s, 80s, and 90s tend to spend their leisure time.

In a **longitudinal study**, a single group of people is followed over the course of several months or years, and data related to the characteristic(s) under investigation are collected at various times.¹ For example, a psycholinguist might examine how children's spoken language changes between 6 months and 5 years of age. Or an educational psychologist might get measures of academic achievement and social adjustment for a group of fourth graders and then, 10 years later, find out which students had completed high school (and what their high school GPAs were) and which ones had not. The educational psychologist might also compute correlations between the measures taken in the fourth grade and the students' high school GPAs; thus, the project would be a correlational study—in this case enabling predictions from Time 1 to Time 2—as well as a longitudinal one.

When longitudinal studies are also correlational studies, they enable researchers to identify potential mediating and moderating variables in correlational relationships. As previously explained in Chapter 2, *mediating variables*—also known as *intervening variables*—may help explain why a characteristic observed at Time 1 is correlated with a characteristic observed at Time 2. Mediating variables are typically measured at some point between Time 1 and Time 2—we might call it Time 1½. In contrast, *moderating variables* influence the nature and strength of a correlational relationship; these might be measured at either Time 1 or Time 1½. A statistical technique mentioned earlier—structural equation modeling (SEM)—can be especially helpful for identifying mediating and moderating variables in a longitudinal study (again we refer you to Table 11.5 in Chapter 11). Yet keep in mind that even with a complex statistical analysis such as SEM, *correlational studies cannot conclusively demonstrate cause-and-effect relationships*.

Obviously, cross-sectional studies are easier and more expedient to conduct than longitudinal studies, because the researcher can collect all the needed data at a single time. In contrast, a researcher who conducts a longitudinal study must collect data over a lengthy period and will almost invariably lose some participants along the way, perhaps because they move to unknown locations or perhaps because they no longer want to participate. An additional disadvantage of a longitudinal design is that when people respond repeatedly to the same measurement instrument, they are likely to improve simply because of their *practice* with the instrument, even if the characteristic being measured hasn't changed at all.

But cross-sectional designs have their disadvantages as well. For one thing, the different age-groups sampled may have been raised under different environmental conditions. For example, imagine that we want to find out whether logical thinking ability improves or declines between the ages of 20 and 70. If we take a cross-sectional approach, we might get samples of 20-year-olds and 70-year-olds and then measure their ability to think logically about various scenarios, perhaps using a standardized multiple-choice test. Now imagine that, in this study, the 20-year-olds obtain higher scores on our logical thinking test than the 70-year-olds. Does this mean that logical thinking ability declines with age? Not necessarily. At least two other possible explanations readily come to mind: The quality of education has changed in many ways over the past few decades, and the younger people have probably had a superior education to that of the older people. And the younger folks have almost certainly had more experience taking multiple-choice tests than the older folks have. Such problems pose threats to the *interval validity*

¹ Some longitudinal studies are conducted over a much shorter time period—perhaps a few minutes or a couple of hours. Such studies, often called *microgenetic studies*, can be useful in studying how children's thinking processes change as a result of short-term, targeted interventions (e.g., Kuhn, 1995).

of this cross-sectional study: We cannot eliminate other possible explanations for the results observed (recall the discussion of internal validity in Chapter 4).

A second disadvantage of a cross-sectional design is that we cannot compute correlations between characteristics at different age levels. Consider, again, the educational psychologist who wants to use students' academic achievement and social adjustment in fourth grade to predict their tendency to complete their high school education. If the educational psychologist were to use a cross-sectional study, there would be different students in each age-group—and only one set of measures for each student—making predictions across time impossible.

To address some of the weaknesses of longitudinal and cross-sectional designs, researchers occasionally combine both approaches in what is known as a **cohort-sequential study**. In particular, a researcher begins with two or more age-groups (this is the cross-sectional piece) and follows each age-group over a period of time (this is the longitudinal piece). As an example, let's return to the issue of how people's logical thinking ability changes over time. Imagine that instead of doing a simple cross-sectional study involving 20-year-olds and 70-year-olds, we begin with a group of 20-year-olds and a group of 65-year-olds. We give them a multiple-choice test designed to assess logical reasoning both at the beginning of the study and then again five years later when they are 25 and 70 years old, respectively. If both groups improve over the five-year time span, we might wonder if practice in taking multiple-choice tests or practice in taking this *particular* test may partly account for the improvement. Alternatively, if the test scores increase for the younger group but decrease for the older group, we might reasonably conclude that logical thinking ability *does* decrease somewhat in the later decades of life.

A cohort-sequential study offers a second advantage, in that, like a longitudinal study, it enables us to calculate correlations between measures taken at two different time periods and therefore to make predictions across time. For instance, we might determine whether people who score highest on the logical thinking test at Time 1 (when they are either 20 or 65 years old) are also those who score highest on the test at Time 2 (when they are either 25 or 70 years old). If we find such a correlation, we can reasonably conclude that logical thinking ability is a relatively stable characteristic—that certain people currently think and will continue to think in a more logical manner than others. We could also add other variables to the study—for instance, determining the extent of participants' postsecondary education or the frequency with which participants engage in activities that require logical reasoning—and determine whether such variables mediate or moderate the long-term stability of logical reasoning ability.

Cross-sectional, longitudinal, and cohort-sequential designs are used in a variety of disciplines, but, as you might guess, they are most commonly seen in developmental research (e.g., studies in child development or gerontology). Should you wish to conduct a developmental study, we urge you to browse in such journals as *Child Development* and *Developmental Psychology* for ideas about specific research strategies.

Survey Research

Some scholars use the term *survey research* to refer to almost *any* form of descriptive, quantitative research. We use a more restricted meaning here: **Survey research** involves acquiring information about one or more groups of people—perhaps about their characteristics, opinions, attitudes, or previous experiences—by asking them questions and tabulating their answers. The ultimate goal is to learn about a large population by surveying a sample of that population; thus, we might call this approach a *descriptive survey* or *normative survey*.

Reduced to its basic elements, a **survey** is quite simple in design: The researcher poses a series of questions to willing participants; summarizes their responses with percentages, frequency counts, or more sophisticated statistical indexes; and then draws inferences about a particular population from the responses of the sample. It is used with more or less sophistication in many areas of human activity—for instance, in a neighborhood petition in support of or against a proposed town ordinance or in a national telephone survey seeking to ascertain people's attitudes about various candidates for political office. This is not to suggest, however, that because of their frequent use, surveys are any less demanding in their design requirements or any easier for the researcher to conduct than other types of research. Quite the contrary, the survey design

makes critical demands on the researcher that, if not carefully respected, can place the entire research effort in jeopardy.

Survey research captures a fleeting moment in time, much as a camera takes a single-frame photograph of an ongoing activity. By drawing conclusions from one transitory collection of data, we may generalize about the state of affairs for a longer time period. But we must never assume that the results can be entered into the Book of Eternal Certainties as ever-abiding Truth. Remember the wisdom of the Greek philosopher Heraclitus: There is nothing permanent but change.

An additional consideration in survey research is that we are relying on *self-report* data: People are telling us what they believe to be true or, perhaps, what they think we want to hear. As noted in the discussion of interviews in Chapter 6, people's memories for events are often distortions of reality—what they think happened isn't always what *did* happen—and people aren't always insightful about their true thoughts and feelings. Furthermore, people's descriptions of their attitudes and opinions are often constructed on the spot—sometimes they haven't really thought about a certain issue until a researcher poses a question about it—and so may be colored by recent events or the current context (Schwarz, 1999). An additional problem is that some participants may intentionally misrepresent the facts—at least, the “facts” as they know them—in order to give the researcher a favorable impression. For example, if we were to ask parents the question “Have you ever abused your children?” the percentage of parents who told us *yes* would be close to zero, and so we would almost certainly underestimate the prevalence of child abuse in our society.

Survey research typically employs a face-to-face interview, a telephone interview, or a written questionnaire. We discuss these techniques briefly here and then offer practical suggestions for conducting them in “Practical Application” sections later on. We describe a fourth approach—using the Internet—in a subsequent “Practical Application” that addresses technology-based methods of data collection.

Face-to-Face and Telephone Interviews

In qualitative research studies, interviews are often quite open-ended, perhaps addressing one or a few central issues but otherwise going in different directions for different participants. In survey research, however, interviews are fairly structured. In a **structured interview**, the researcher asks a standard set of questions and nothing more. In a **semistructured interview**, the research may follow the standard questions with one or more individually tailored questions to get clarification or probe a person's reasoning.

Another difference between qualitative and quantitative studies is the general “feel” of the interview: It tends to be informal and friendly in a qualitative study but more formal and emotionally neutral in a quantitative one. Participants in a qualitative interview may feel as if they're simply engaging in a friendly chat with the researcher, who is typically someone they have come to know and trust. In contrast, participants in survey research are continually aware that, yes, this is an interview and that the temporary relationship they've formed with the researcher will end once the interview is complete. This is not to say, however, that a survey researcher shouldn't strive to establish rapport with participants. Quite the contrary, the researcher is more likely to gain participants' cooperation and encourage them to respond honestly if he or she is likable and friendly and shows a genuine interest in what people have to say.

Face-to-face interviews have the distinct advantage of enabling the researcher to establish rapport with potential participants and therefore gain their cooperation. Thus, such interviews yield the highest **response rates**—the percentages of people agreeing to participate—in survey research. However, the time and expense involved may be prohibitive if the needed interviewees reside in a variety of states, provinces, or countries.

Telephone interviews are less time-consuming and less expensive (they involve only the cost of any long-distance calls), and the researcher has potential access to virtually anyone on the planet who has a landline telephone or cell phone. Although the response rate is not as high as for a face-to-face interview—many people are apt to be busy, annoyed at being bothered, concerned about using costly cell phone minutes, or otherwise not interested in participating—it is considerably higher than for a mailed questionnaire. Unfortunately, the researcher conducting telephone interviews cannot establish the same kind of rapport that is possible in a face-to-face

situation, and the sample will be biased to the extent that people without phones are part of the population about whom the researcher wants to draw inferences.²

Whether they are conducted face-to-face or over the telephone, personal interviews allow the researcher to clarify ambiguous answers and, when appropriate, seek follow-up information. Because such interviews take time, however, they may not be practical when large sample sizes are important.

Questionnaires

Paper-and-pencil questionnaires can be sent to a large number of people, including those who live thousands of miles away. Thus, they may save the researcher travel expenses, and postage is typically cheaper than a lengthy long-distance telephone call. The social scientist who collects data with a questionnaire and the physicist who determines the presence of radioactivity with a Geiger counter are at just about the same degree of remoteness from their respective sources of data: Neither sees the source from which the data originate. From the perspective of survey participants, this distance can be an additional advantage: Participants can respond to questions with some assurance that their responses won't come back to haunt them. Thus, they may be more truthful than they would be in a personal interview, especially when addressing sensitive or controversial issues.

Yet questionnaires have their drawbacks as well. Typically, the majority of people who receive questionnaires don't return them—in other words, there may be a low **return rate**—and the people who do return them are not necessarily representative of the originally selected sample. Even when people are willing participants in a questionnaire study, their responses will reflect their reading and writing skills and, perhaps, their misinterpretation of one or more questions. Furthermore, by specifying in advance all of the questions that will be asked—and thereby eliminating other questions that *could* be asked about the issue or phenomenon in question—the researcher is apt to gain only limited, and possibly distorted, information (Dowson & McInerney, 2001).

If questionnaires are to yield useful data, they must be carefully planned, constructed, and distributed. In fact, *any* descriptive study requires careful planning, with close attention to each methodological detail. We now turn to the topic of planning.

Planning for Data Collection in a Descriptive Study

In quantitative research, a descriptive study invariably involves measuring one or more variables in some way. With this point in mind, let's return to a distinction we first made in Chapter 4—the distinction between substantial and insubstantial phenomena. When studying the nature of *substantial phenomena*—phenomena that have physical substance, an obvious basis in the physical world—a researcher can often use measurement instruments that are clearly valid for their purpose. Tape measures, balance scales, oscilloscopes, MRI machines—these instruments are indisputably valid for measuring length, weight, electrical waves, and internal body structures, respectively. Some widely accepted measurement techniques also exist for studying *insubstantial phenomena*—concepts, abilities, and other intangible entities that cannot be pinned down in terms of precise physical qualities. For example, an economist might use Gross Domestic Product statistics as measures of a nation's economic growth, and a psychologist might use the *Stanford-Binet Intelligence Scale* to measure children's general cognitive ability.

Yet many descriptive studies address complex variables—perhaps people's or animals' day-to-day behaviors, or perhaps people's opinions and attitudes about a particular topic—for which no ready-made measurement instruments exist. In their observations, interviews, questionnaires, and so on, descriptive researchers use a variety of strategies to measure complex variables, as you will see in the upcoming Practical Application sections.

²Midway between a face-to-face interview and a telephone interview is an interview conducted on the Internet using Skype (www.skype.com) or other video conferencing software. As this tenth edition of the book goes to press, however, people's access to such software is quite limited—and, in our experience, not always dependable—thereby introducing a significant bias into the sample a researcher might get.

PRACTICAL APPLICATION Using Checklists and Rating Scales

Two techniques that facilitate evaluation and quantification of complex phenomena are the checklist and the rating scale. A **checklist** is a list of behaviors or characteristics for which a researcher is looking. The researcher—or in some studies, each participant—simply indicates whether each item on the list is observed, present, or true or, in contrast, is *not* observed, present, or true.

A **rating scale** is more useful when a behavior, attitude, or other phenomenon of interest needs to be evaluated on a continuum of, say, “inadequate” to “excellent,” “never” to “always,” or “strongly disapprove” to “strongly approve.” Rating scales were developed by Rensis Likert in the 1930s to assess people’s attitudes; accordingly, they are sometimes called **Likert scales**.³

Checklists and rating scales can presumably be used in research related to a wide variety of phenomena, including those involving human beings, nonhuman animals, plants, or inanimate objects (e.g., works of art and literature, geomorphological formations). We illustrate the use of both techniques with a simple example involving human participants. In the late 1970s, park rangers at Rocky Mountain National Park in Colorado were concerned about the heavy summer-time traffic traveling up a narrow mountain road to Bear Lake, a popular destination for park visitors. So in the summer of 1978, they provided buses that would shuttle visitors to Bear Lake and back again. This being a radical innovation at the time, the rangers wondered about people’s reactions to the buses; if there were strong objections, other solutions to the traffic problem would have to be identified for the following summer.

Park officials asked a sociologist friend of ours to address their research question: How do park visitors feel about the new bus system? The sociologist decided that the best way to approach the problem was to conduct a survey. He and his research assistants waited at the parking lot to which buses returned after their trip to Bear Lake; they randomly selected people who exited the bus and administered the survey. With such a captive audience, the response rate was extremely high: 1,246 of the 1,268 people who were approached agreed to participate in the study, yielding a response rate of 98%.

We present three of the interview questions in Figure 8.2. Based on people’s responses, the sociologist concluded that people were solidly in favor of the bus system

FIGURE 8.2

Excerpts from a survey at Rocky Mountain National Park. Item 4 is a *checklist*. Items 5 and 6 are *rating scales*.

From Trahan (1978, Appendix A).

4. Why did you decide to use the bus system?

- ☐ Forced to; Bear Lake was closed to cars
- ☐ Thought it was required
- ☐ Environmental and aesthetic reasons
- ☐ To save time and/or gas
- ☐ To avoid or lessen traffic
- ☐ Easier to park
- ☐ To receive some park interpretation
- ☐ Other (specify): _____

5. In general, what is your opinion of public bus use in national parks as an effort to reduce traffic congestion and park problems and help maintain the environmental quality of the park?

Strongly approve	Approve	Neutral	Disapprove	Strongly disapprove

If “disapprove” or “strongly disapprove,” why? _____

6. What is your overall reaction to the present Bear Lake bus system?

Very satisfied	Satisfied	Neutral	Dissatisfied	Very dissatisfied

³Although we have often heard *Likert* pronounced as “lie-kert,” Likert himself pronounced it “lick-ert.”

(Trahan, 1978). As a result, it continues to be in operation today, many years after the survey was conducted.

One of us authors was once a member of a dissertation committee for a doctoral student who developed a creative way of presenting a Likert scale to children (Shaklee, 1998). The student was investigating the effects of a particular approach to teaching elementary school science and wanted to determine whether students' beliefs about the nature of school learning—especially learning science—would change as a result of the approach. Both before and after the instructional intervention, she read a series of statements and asked students either to agree or to disagree with them by pointing to one of four faces. The statements and the rating scale that students used to respond to them are presented in Figure 8.3.

Notice that in the rating scale items in the Rocky Mountain National Park survey, park visitors were given the option of responding "Neutral" to each question. In the elementary school study, however, the children always had to answer "yes" or "no." Experts have mixed views about letting respondents remain neutral in interviews and questionnaires. If you use rating scales in your own research, you should consider the implications of letting your respondents straddle the fence by including a "no opinion" or other neutral response, and design your scales accordingly.

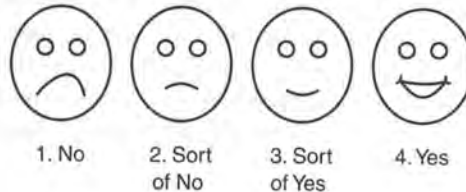
Whenever you use checklists or rating scales, you simplify and more easily quantify people's behaviors or attitudes. In the process, however, you may lose valuable information that clarifies or qualifies people's responses. Ultimately you will have to determine whether the trade-off is worth it for the particular research problem you are investigating.

FIGURE 8.3

Asking elementary school children about science and learning

From *Elementary Children's Epistemological Beliefs and Understandings of Science in the Context of Computer-Mediated Video Conferencing with Scientists* (pp. 132, 134) by J. M. Shaklee, 1998, unpublished doctoral dissertation, University of Northern Colorado, Greeley. Reprinted with permission.

Students responded to each statement by pointing to one of the faces below.



Students who were unfamiliar with Likert scales practiced the procedure using items A and B; others began with item 1.

- A. Are cats green?
- B. Is it a nice day?
- 1. The best thing about science is that most problems have one right answer.
- 2. If I can't understand something quickly, I keep trying.
- 3. When I don't understand a new idea, it is best to figure it out on my own.
- 4. I get confused when books have different information from what I already know.
- 5. An expert is someone who is born really smart.
- 6. If scientists try hard enough, they can find the truth to almost everything.
- 7. Students who do well learn quickly.
- 8. Getting ahead takes a lot of work.
- 9. The most important part about being a good student is memorizing the facts.
- 10. I can believe what I read.
- 11. Truth never changes.
- 12. Learning takes a long time.
- 13. Really smart students don't have to work hard to do well in school.
- 14. Kids who disagree with teachers are show-offs.
- 15. Scientists can get to the truth.
- 16. I try to use information from books and many other places.
- 17. It is annoying to listen to people who can't make up their minds.
- 18. Everyone needs to learn how to learn.
- 19. If I try too hard to understand a problem, I just get confused.
- 20. Sometimes I just have to accept answers from a teacher even if they don't make sense to me.

PRACTICAL APPLICATION Computerizing Observations



In quantitative research studies, most observations are, of course, quantified in some way, sometimes with simple frequency counts—perhaps using a checklist for guidance—and at other times with rating scales. Either way, you might record your observations on a computer laptop or tablet as you are making them. For example, when using a checklist, you might create a spreadsheet with a small number of columns—one for each item on the checklist—and a row for every entity you will observe. Then, as you conduct your observations, you can enter an “X” or other symbol into the appropriate cell whenever you see an item in the checklist.

For more complex observations, you might create a general *template* document in either a spreadsheet or word processing software and then electronically “save” a separate version of the document for each person, situation, or other entity you are observing. You can either print out these entity-specific documents for handwritten coding during your observations, or, if time and your keyboarding skills allow, you can fill in each document while on-site in the research setting.

You might also look for *peripheral devices*—devices you can hook to your computer to enhance its capabilities—to aid your data collection. For example, Acheson and Gall (2003) have described an approach in which observers use a bar code reader and a special list of bar codes to record different types of behaviors. Each code is associated with a specific type of behavior. When a particular behavior is observed, the bar code reader is swept over the appropriate bar code, and immediately the type and time of the behavior are recorded. In this way, all observations are electronically entered and categorized in a single motion.

PRACTICAL APPLICATION Planning and Conducting Interviews in a Quantitative Study

Interviewing involves much more than just asking questions. The questions for an interview should be carefully planned and precisely worded to yield the kinds of data the researcher needs to answer his or her research question. In the following two sections, we offer guidelines for conducting interviews in a quantitative study and describe how one student successfully planned and conducted the interviews he needed for a research project.

GUIDELINES Conducting Interviews in a Quantitative Study

In Chapter 6 we presented guidelines for conducting interviews in qualitative research. Most of those guidelines are equally applicable in quantitative research, and so we list them again here:

1. Identify questions in advance.
2. Consider how participants' cultural backgrounds might influence their responses.
3. Make sure your interviewees are representative of the group.
4. Find a suitable location.
5. Get written permission.
6. Establish and maintain rapport.
7. Focus on the actual rather than on the abstract or hypothetical.
8. Don't put words in people's mouths.
9. Record responses verbatim.
10. Keep your reactions to yourself.
11. Remember that you are not necessarily getting the facts.

But interviews are typically more structured in quantitative studies than they are in qualitative studies. Therefore, we add several additional guidelines for conducting interviews in quantitative research.

12. *As you write the questions, consider how you can quantify the responses, and modify the questions accordingly.* Remember, you are conducting a *quantitative* study. Thus you will, to some extent,

be coding people's responses as numbers and, quite possibly, conducting statistical analyses on those numbers. You will be able to assign numerical codes to responses more easily if you identify an appropriate coding scheme ahead of time.

13. *Consider asking questions that will elicit qualitative information as well.* You don't necessarily have to quantify *everything*. People's responses to a few open-ended questions may support or provide additional insights into the numerical data you obtain from more structured questions (for example, see the open-ended question in Item 5 of Figure 8.2). By combining quantitative and qualitative data in this manner, you are essentially employing a *mixed-methods design*.

14. *Pilot-test the questions.* When you plan your interview, you will, of course, be trying to develop questions that elicit the kinds of information you are seeking. But despite your best intentions, you may create questions that are ambiguous or misleading or that yield uninterpretable or otherwise useless responses. You can save yourself a great deal of time over the long run if you fine-tune your questions before you begin data collection. You can easily find the weak spots in your questions by asking a few volunteers to answer them in a pilot study.

15. *Restrict each question to a single idea.* Don't try to get too much information in any single question; in doing so, you may get multiple kinds of data—"mixed messages," so to speak—that are difficult to interpret (Gall, Gall, & Borg, 2007).

16. *Save controversial questions for the latter part of the interview.* If you will be touching on sensitive topics (e.g., opinions about gun control, attitudes toward people with diverse sexual orientations), put them near the end of the interview, after you have established rapport and gained the person's trust (Gall et al., 2007).

17. *Seek clarifying information when necessary.* Be alert for responses that are vague or otherwise difficult to interpret. Simple, nonleading questions—for instance, "Can you tell me more about that?"—may yield the additional information you need (Gall et al., 2007, p. 254).

18. *Consider how you might use a computer to streamline the process.* As mentioned in Chapter 6, some computer software programs allow you to record interviews directly onto a laptop computer and then transform these conversations into written text. Alternatively, if interviewees' responses are likely to be short, you might type them directly into a spreadsheet or word processing program.



An Example in International Relations

A student wanted to interview certain United Nations personnel to get their opinions concerning issues related to his study. He planned to go to New York City for a series of interviews and, to conserve both time and expense, wanted to schedule them as tightly as possible. His procedure was organized and logical.

Approximately 6 weeks before going for the interviews, the student wrote the United Nations representatives with whom he wished to confer; he told them when he would be in New York and requested an interview that would last 30 minutes at most. He asked each prospective interviewee for an indication of several time slots when the interview might be scheduled. In his letter, he clearly explained what information he was seeking and why he was seeking it. His reasons were mature and meaningful and were phrased to pique the interest for those he wanted to interview. (*Not* among his reasons was the fact that he was writing a thesis! If you must reveal that you are collecting data for a thesis, use the word *study* instead of *thesis*. Aside from the student and the graduate advisor, theses hold very little glamour in the everyday world. "Studies" are much more acceptable.)

With the letter, the student enclosed a separate sheet containing the questions he intended to ask during the interview, arranged in the order he would ask them. He also suggested that if the interviewee had no objections, he would tape the interview in order to conserve time and lessen the distraction of handwritten notes. He provided a check box on a return postcard for the interviewee to indicate whether he or she had any objection to recording the interview.

After receiving potential interviewees' replies, he set up a master chart of appointments and, by letter, immediately confirmed each interviewee's appointment time and thanked the

interviewee for his or her cooperation. When a time conflict arose, he sought to resolve it by suggesting alternative times that were still open.

Ten days before the scheduled interviews, the student mailed reminders, together with another copy of the interview questions, just in case any interviewees had misplaced the previously sent copies. He also enclosed his full interview schedule so that the interviewees might appreciate the time constraints under which he was working.

The student arrived promptly for each scheduled interview, introduced himself, asked whether the interviewee wanted a copy of the questions he had previously sent, and began with the first question. He tried to guide the interview, always keeping to his agenda of questions and seeking to preserve a relaxed, friendly, yet also professional atmosphere. He wrapped up each interview by thanking the interviewee for the courtesy of giving his or her time. In 3½ days, he interviewed 35 United Nations representatives and had more than four-fifths of his data on tape.

The student transcribed the substance of the interviews and, within 10 days of his visit, sent each interviewee a typed, double-spaced transcript accompanied by a thank-you letter for granting the interview. He asked each individual to read the transcript carefully and, if it was correct, to sign a statement that it was a correct record of the interview. If the person found it inexact or incorrect in any place, he or she could correct the script as desired. In the same mailing, the researcher included a request for permission to use any quotations from the interview in his final report, with the understanding that he would again send the interview content for the interviewee's approval. In his final thesis, the researcher acknowledged his interviewees and noted that they had inspected and approved all of their quoted statements. With the use of such strategies, the researcher and the readers of his report could all be confident that the participants' thoughts and opinions were accurately represented.

In summary, the researcher's use of the following steps led to a highly productive research effort:

1. Set up the interview well in advance.
2. Send the agenda of questions to ask the interviewee.
3. Ask for permission to tape the conference.
4. Confirm the date immediately in writing.
5. Send a reminder, together with another copy of the questions, 10 days before the interview.
6. Be prompt; follow the agenda; offer a copy of the questions in case the original copy has been mislaid.
7. After the interview, submit a transcript of the interview, and get from the interviewee either a written acknowledgment of its accuracy or a corrected copy.
8. After incorporating the material into a semifinal draft of the research report, send that section of the report to the interviewee for final approval and written permission to use the data in the report.

PRACTICAL APPLICATION Constructing and Administering a Questionnaire

Questionnaires seem so simple, yet in our experience they can be tricky to construct and administer. One false step can lead to uninterpretable data or an abysmally low return rate. We have numerous suggestions that can help you make your use of a questionnaire both fruitful and efficient. We have divided our suggestions into three categories: constructing your questionnaire, using technology to facilitate questionnaire administration and data analysis, and maximizing your return rate. In a subsequent Practical Application, we discuss how you might solicit survey participants and collect your data online.

GUIDELINES Constructing Your Questionnaire

Following are 12 guidelines for developing a questionnaire that encourages people to be cooperative and yields responses you can use and interpret. We apologize for the length of the list, but, as we just said, questionnaire construction is a tricky business.

1. *Keep it short.* Your questionnaire should be as brief as possible and solicit only information that is essential to the research effort. You should test every item by two criteria: (a) What do I intend to do with the information I am requesting? and (b) Is it absolutely essential to have this information to solve part of the research problem?

2. *Keep the respondent's task simple and concrete.* Make the instrument as simple to read and respond to as possible. Remember, you are asking for people's *time*, a precious commodity for many people these days. People are more likely to respond to a questionnaire—and to do so quickly—if they perceive it to be quick and easy to complete (McCrea et al., 2008).

Discussion items—those that present open-ended questions and ask people to respond with lengthy answers—are time-consuming and mentally exhausting for both the participants and the researcher. Don't forget that you will have to wrestle with the participants' words to try to determine exactly what their answers mean. The usefulness of responses to discussion items rests entirely on participants' skill to express their thoughts in writing. Those who write in the "Yes/no, and I'll tell you exactly why" style are few and far between. Some respondents may ramble, engaging in discussions that aren't focused or don't answer the questions. Furthermore, after answering 15 to 20 discussion questions, your respondents will think you are demanding a book! Such a major compositional exercise is unfair to those from whom you are requesting a favor.

3. *Provide straightforward, specific instructions.* Communicate exactly how you want people to respond. For instance, don't assume that they are familiar with Likert scales. Some of them may never have seen such scales before.

4. *Use simple, clear, unambiguous language.* Write questions that communicate exactly what you want to know. Avoid terms that your respondents may not understand, such as obscure words or technical jargon. Also avoid words that have imprecise meanings, such as *several* and *usually*.

5. *Give a rationale for any items whose purpose may be unclear.* We cannot say this enough: You are asking people to do you a favor by responding to your questionnaire. Give them a reason to *want* to do the favor. Each question should have a purpose, and in one way or another, you should make that purpose clear.

6. *Check for unwarranted assumptions implicit in your questions.* Consider a very simple question: "How many cigarettes do you smoke each day?" It seems to be a clear and unambiguous question, especially if it is accompanied with certain choices so that all the respondent has to do is to check one of them:

How many cigarettes do you smoke each day? Check one of the following:

☐ More than 25 ☐ 25–16 ☐ 15–11 ☐ 10–6 ☐ 5–1 ☐ none

One underlying assumption here is that a person is likely to be a smoker rather than nonsmoker, which isn't necessarily the case. A second assumption is that a person smokes the same number of cigarettes each day, but for many smokers this assumption isn't viable. At work, when the pressure is on, people may be chain smokers. But on weekends and holidays, they may relax and smoke only one or two cigarettes a day or go without smoking at all. How are the people in this group supposed to answer the question? What box does this type of smoker check?

Had the author of the question considered the assumptions on which the question was predicated, he or she might first have asked questions such as these:

Do you smoke cigarettes?

☐ Yes

☐ No (If you mark "no," skip the next two questions.)

Are your daily smoking habits reasonably consistent; that is, do you smoke about the same number of cigarettes each day?

☐ Yes

☐ No (If you mark "no," skip the next question.)

7. *Word your questions in ways that don't give clues about preferred or more desirable responses.* Take another question: "What strategies have you used to try to quit smoking?" By implying that the

respondent has, in fact, tried to quit, it may lead him or her to describe strategies that have never been seriously tried at all.

8. *Determine in advance how you will code the responses.* As you write your questions—perhaps even *before* you write them—develop a plan for recoding participants' responses into numerical data you can statistically analyze. Data processing procedures may also dictate the form a questionnaire should take. If, for example, people's response sheets will be fed into a computer scanner, the questionnaire must be structured differently than if the responses will be tabulated using paper and pencil (we'll say more about computer scanning in the subsequent set of guidelines).

9. *Check for consistency.* When a questionnaire asks questions about a potentially controversial topic, some respondents might give answers that are socially acceptable rather than accurate in order to present a favorable impression. To allow for this possibility, you may wish to ask the same question two or more times—using different words each time—at various points in your questionnaire. For example, take the following two items, appearing in a questionnaire as Items 2 and 30. (Their distance from each other increases the likelihood that a person will answer the second without recalling how he or she answered the first.) Notice how one individual has answered them:

2. Check one of the following choices:

☒ In my thinking, I am a liberal.

☐ In my thinking, I am a conservative.

30. Check one of the following choices:

☐ I find new ideas stimulating and attractive, and I would find it challenging to be among the first to try them.

☒ I subscribe to the position of Alexander Pope:

"Be not the first by whom the new is tried,
nor yet the last to lay the old aside."

The two responses are inconsistent. In the first, the respondent claims to be a liberal thinker but later, when given liberal and conservative positions in other forms, indicates a position generally thought to be more conservative than liberal. Such an inconsistency might lead you to question whether the respondent is truly the liberal thinker he or she claims to be.

When developing a questionnaire, researchers sometimes include several items designed to assess essentially the same characteristic. This approach is especially common in studies that involve personality characteristics, motivation, attitudes, and other complex psychological traits. For example, one of us authors once worked with two colleagues to explore factors that might influence the teaching effectiveness of college education majors who were completing their teaching internship year (Middleton, Ormrod, & Abrams, 2007). The research team speculated that one factor potentially affecting teaching effectiveness was willingness to try new teaching techniques and in other ways take reasonable risks in the classroom. The team developed eight items to assess risk taking. Following are four examples, which were interspersed among items designed to assess other characteristics:

	Not at All True		Somewhat True		Very True
11. I would prefer to teach in a way that is familiar to me rather than trying a teaching strategy that I would have to learn how to do.	1	2	3	4	5
16. I like trying new approaches to teaching, even if I occasionally find they don't work very well.	1	2	3	4	5
39. I would choose to teach something I knew I could do, rather than a topic I haven't taught before.	1	2	3	4	5
51. I sometimes change my plan in the middle of a lesson if I see an opportunity to practice teaching skills I haven't yet mastered.	1	2	3	4	5

Notice how a response of “Very true” to Items 16 and 51 would be indicative of a *high* risk taker, whereas a response of “Very true” to Items 11 and 39 would be indicative of a *low* risk taker. Such counterbalancing of items—some reflecting a high level of a characteristic and others reflecting a low level of the characteristic—can help address some people’s general tendency to agree or disagree with a great many statements, including contradictory ones (Nicholls, Orr, Okubo, & Loftus, 2006). A researcher who uses this counterbalancing approach cannot, of course, simply add up a participant’s numerical responses for a particular characteristic. For example, for the four risk-taking items just presented, a researcher who wanted high risk takers to have higher scores than low risk takers might give 5 points each for “Very true” responses to the high-risk-taking items (16 and 51) and 5 points each for “Not at all true” responses to the low-risk-taking items (11 and 39). In general, the values of the low-risk-taking items would, during scoring, be opposite to what they are on the questionnaire, with 1s being worth 5 points each, 2s being worth 4 points, 3s being worth 3, 4s being worth 2, and 5s being worth 1. In Appendix A, we describe how to recode participants’ responses in precisely this way.

Especially when multiple items are created to assess a single characteristic, a good researcher mathematically determines the degree to which, overall, participants’ responses to those items yield consistent results—for instance, the extent to which each person’s responses to all “risk-taking” items yield similar results. Essentially, the researcher is determining the *internal consistency reliability* of the set of items. Most statistical software packages can easily compute internal consistency reliability coefficients for you.⁴

Ideally, preliminary data on internal consistency reliability is collected in advance of full-fledged data collection. This point leads us to our next suggestion: Conduct at least one pilot test.

10. *Conduct one or more pilot tests to determine the validity of your questionnaire.* Even experienced researchers conduct test runs of newly designed questionnaires to make sure that the questions are clear and will effectively solicit the desired information. At a minimum, you should give your questionnaire to several friends or colleagues to see whether they have difficulty understanding any items. Have them actually fill out the questionnaire. Better still, ask your pilot-test participants what thoughts run through their minds as they read a question:

Please read this question out loud. . . . What is this question trying to find out from you? . . . Which answer would you choose as the right answer for you? . . . Can you explain to me why you chose that answer? (Karabenick et al., 2007, p. 143)

Through such strategies you can see the kinds of responses you are likely to get and make sure that in your actual study, the responses you obtain will be of sufficient quality to help you answer your research question.

If your research project will include participants of both genders and various cultural backgrounds, be sure to include a diverse sample in your pilot test(s) as well. Gender and culture *do* play a role in people’s responses to certain types of questionnaire items. For instance, some researchers have found a tendency for males to play up their strengths and overrate their abilities, whereas females are apt to ruminate on their weaknesses and *underrate* their abilities (Chipman, 2005; Lundeborg & Mohan, 2009). And people from East Asian cultures are more likely to downplay their abilities than people from Western cultures (Heine, 2007). Keep such differences in mind when asking people to rate themselves on their strengths and weaknesses, and experiment with different wordings that might minimize the effects of gender and culture on participants’ responses.

Conducting a pilot study for a questionnaire—and especially asking participants what they are thinking as they read and respond to particular items—is one step toward determining whether a questionnaire has *validity* for its purpose—in other words, whether it truly measures what it is intended to measure. Some academic disciplines (e.g., psychology and related fields) insist that a researcher use more formal and objective strategies to determine a questionnaire’s validity, especially when the questionnaire is intended to measure complex psychological traits

⁴Two common reliability coefficients, known by the researchers who originated them, are the Kuder-Richardson Formula 20 (for either-or responses such as *yes* vs. *no* or *true* vs. *false*) and Cronbach’s alpha coefficient (for multirating scales such as the 5-point scale for the risk-taking items).

TABLE 8.1 Guide for the construction of a questionnaire

Write the question in the space below.	Why are you asking the question? How does it relate to the research problem?

(personality, motivation, attitudes, etc.). We refer you to the section “Determining the Validity of a Measurement Instrument” in Chapter 4 for a refresher on strategies you might use.

11. *Scrutinize the almost-final product one more time to make sure it addresses your needs.* Item by item, a questionnaire should be quality tested again and again for precision of expression, objectivity, relevance, and probability of favorable reception and return. Have you concentrated on the recipient of the questionnaire, putting yourself in the place of someone who is asked to invest time on your behalf? If you received such a questionnaire from a stranger, what would your honest reaction be? These questions are important and should be answered impartially.

Above all, you should make sure that *every question is essential for you to address the research problem*. Table 8.1 can help you examine your items with this criterion in mind. Using either paper and pencil or appropriate software (e.g., a spreadsheet or the *table* feature in a word processing program), insert each item in the left-hand column and then, in the right-hand column, explain why you need to include it. If you can’t explain how an item relates to your research problem, eliminate it!

12. *Make the questionnaire attractive and professional looking.* Your final instrument should have clean lines, crystal-clear printing (and certainly no typos!), and perhaps two or more colors. It should ultimately communicate that its author is a careful, well-organized professional who takes his or her work seriously and has high regard for the research participants.

GUIDELINES Using Technology to Facilitate Questionnaire Administration and Data Analysis



USING
TECHNOLOGY

Throughout most of the 20th century, questionnaire-based surveys were almost exclusively paper-and-pencil in nature. But with continuing technological advances and people’s increasing computer literacy in recent years, many survey researchers are now turning to technology to share some of the burden of data collection and analysis. Following are several suggestions for using technology to make the use of a questionnaire more efficient and cost-effective:

1. *Use e-mail to request participation and obtain participants’ responses.* If the people you wish to survey are a computer-savvy group, have easily obtainable e-mail addresses, and are regularly online, then an e-mail request to participate can be quite appropriate. Furthermore, you can send the survey either within the body of your e-mail message or as an attachment. Participants can respond in a return e-mail message or electronically fill out and return your attachment.

2. *If you use paper mail delivery rather than e-mail, use a word processing program to personalize your correspondence.* Inquiry letters, thank-you letters, and other correspondence can be personalized by using the *merge* function of most word processing programs. This function allows you to combine the information in your database with the documents you wish to send out. For example, when printing the final version of your cover letter, you can include the person's name immediately after the greeting (e.g., "Dear Carlos" or "Dear Mr. Asay")—a simple touch that is likely to yield a higher return rate than letters addressed to "Potential Respondent" or "To whom it may concern." The computer inserts the names for you; you need only tell it where to find the names in your database. The same process is used on every item of junk mail you receive that addresses you personally or that shows your name on a \$10,000,000 check (that you may already have won!) and asks you to order magazines. Advertisers have realized the importance of personalization for quite some time. When individuals are addressed by name, they feel special and are more inclined to invest a little of their time.

3. *Use a computer database to keep track of who has responded and who has not.* Computer databases can be readily adapted for use with research involving questionnaires. Information regarding participants' names and addresses, as well as information regarding which materials have been sent and received, can be incorporated into such a database. A search through the database can quickly identify people who have and have not received your request for participation, those who have or have not yet responded, and those who need a first or second reminder letter or e-mail message.

Combining the database with a word processing program can also be helpful. For example, mailing labels can be quickly produced, thus decreasing the amount of time needed to address various mailings to potential participants. Many word processing and spreadsheet programs come with prepared templates for a variety of mailing labels. By identifying the type of label you have for your printer, you can print any or all of the names from your database directly onto the labels.

4. *Use a scanner to facilitate data tabulation.* When you need a large sample to address your research problem adequately, you should consider in advance how you will tabulate the responses after the questionnaires are returned to you. One widely used strategy is to have a computer scan preformatted answer sheets and automatically sort and organize the results. To use this strategy, your questions must each involve a small set of possible answers; for instance, they might be multiple-choice, have yes-or-no answers, or incorporate a 5-point rating scale. You will want the participants to respond using a pencil or dark-colored ink. Enclosing a small number 2 pencil with the questionnaire you send is common courtesy. Furthermore, anything you can do to make participants' task easier—even something as simple as providing the writing implement—will increase your response rate.

5. *When participants are in the same location that you are, administer the questionnaire directly on a computer.* Electronic questionnaires can be highly effective if participants feel comfortable with computers. When participants enter their responses directly onto a computer, you obviously save a great deal of time. Furthermore, when appropriately programmed to do so, a computer can record how *quickly* people respond—information that may in some situations be relevant to your research question.

GUIDELINES Maximizing Your Return Rate for a Questionnaire

As university professors, we authors have sometimes been asked to distribute questionnaires in our classes that relate to some aspect of the university's student services or to students' preferences for the university calendar. The end-of-semester teacher evaluation forms you often fill out are questionnaires as well. Even though participation in such surveys is voluntary, the response rate when one has such a captive audience is typically quite high, often 100%.

Mailing or e-mailing questionnaires to people one doesn't know is quite another matter. Potential respondents have little or nothing to gain by answering and returning the questionnaire, and so many of them don't return it. As a result, the typical return rate for a mailed

questionnaire is 50% or less, and that for an e-mailed questionnaire is even lower (Rogelberg & Luong, 1998; Sheehan, 2001).

We think of one doctoral student who conducted dissertation research in the area of reading. As part of her study, she sent a questionnaire to reading teachers to inquire about their beliefs and attitudes regarding a certain kind of children's literature. Initially, the student sent out 103 questionnaires; 14 teachers completed and returned them (a return rate of 13%). In a second attempt, she sent out 72 questionnaires to a different group of teachers; 12 responded (a return rate of 15%). In one final effort, she sought volunteers on the Internet by using two lists of teachers' e-mail addresses; 57 teachers indicated that they were willing to fill out her questionnaire, and 20 of them actually did so (a return rate of 35%).

Was the student frustrated? Absolutely! Yet she had made a couple of mistakes that undoubtedly thwarted her efforts from the beginning. First, the questionnaire had 36 questions, 18 of which were open-ended ones requiring lengthy written responses. A quick glance would tell any discerning teacher that the questionnaire would take an entire evening to complete. Second, the questionnaires were sent out in the middle of the school year, when teachers were probably already quite busy planning lessons, grading papers, and so on. Even teachers who truly wanted to help this struggling doctoral student (who was a former teacher herself) may simply not have found the time to do it. Fortunately for the student, the questionnaire was only one small part of her study, and she was able to complete her dissertation successfully with the limited (and almost certainly nonrepresentative) sample of responses she received.

Should you decide that a mailed or e-mailed questionnaire is the most suitable approach for answering your research question, the following guidelines can help you increase your return rate:

1. *Consider the timing.* The student just described mailed her questionnaires in the winter and early spring because she wanted to graduate at the end of the summer. The timing of her mailing was convenient for her; however, it was *not* convenient for the people to whom she sent the questionnaire, and her response rate—and her study!—suffered as a result. Consider the characteristics of the sample you are surveying, and try to anticipate when respondents will be most likely to have time to answer a questionnaire. And as a general rule, stay away from peak holiday and vacation times, such as mid-December through early January.

2. *Make a good first impression.* Put yourself in the place of a potential respondent. Imagine a stranger sending you the questionnaire you propose to send. What is your initial impression as you open the envelope or e-mail message? Is the questionnaire inordinately long and time-consuming? Is it cleanly and neatly written? Does it give the impression of relaxation and uncluttered ease? Are the areas for response adequate and clearly indicated? Is the tone courteous, and are the requests reasonable?

3. *Motivate potential respondents.* Give people a reason to *want* to respond. Occasionally, researchers may actually have the resources to pay people for their time or offer other concrete inducements. But more often than not, you will have to rely on the power of persuasion to gain cooperation. Probably the best mechanism for doing so is the cover letter you include with your questionnaire.

One potentially effective strategy is to send a letter soliciting people's cooperation *before* actually sending them the questionnaire. For example, Figure 8.4 shows an example of a letter that a researcher might use to gain people's cooperation in responding to a questionnaire about the quality of a particular academic program. Several aspects of the letter are important to note:

- The letter begins with the name of the sponsoring institution. Ideally, the letter is written on the institution's official letterhead stationery.
- Rather than saying "Dear Sir or Madam," the letter is personalized. Such personalization is easily accomplished with the *mail merge* function of a word processing program.
- The letter describes the potential value of the study, both for the individual and for alumni in general, hence giving the potential responder a reason to *want* to respond.
- The letter assures the individual that his or her cooperation will not place any unreasonable burden—in particular, that the questionnaire will take a maximum of 15 minutes to complete.
- By filling out and sending a simple enclosed postcard (for example, see Figure 8.5)—a quick and easy first step—the researcher gains the individual's commitment to completing

FIGURE 8.4

A letter of inquiry

A B C University
Address

Date _____

Dear [person's name],

Your alma mater is appealing to you for help. We are not asking for funds, merely for a few minutes of your time.

We know you are proud of your accomplishments at A B C University, and your degree has almost certainly helped you advance your professional aspirations. You can help us maintain—and ideally also improve—your program's reputation by giving us your honest opinion of its strengths and weaknesses while you were here. We have a questionnaire that, with your permission, we would like to send you. It should take at most only fifteen minutes of your time.

Our program is growing, and with your help it can increase not only in size but also in excellence and national prominence. We are confident that you can help us make it the best that it can possibly be.

Enclosed with this letter is a return postcard on which you can indicate your willingness to respond to our questionnaire. Thank you in advance for your kind assistance. And please don't hesitate to contact me at [telephone number] or [e-mail address] if you have any questions or concerns.

Respectfully yours,

Your Signature

Your Name _____

a lengthier, more complex task in the near future. The postcard should be addressed and stamped for easy return.

- The letter includes two means of communicating with the researcher in case the individual has any reservations about participating in the study.
- The overall tone of the letter is, from beginning to end, courteous and respectful.

Compare the letter in Figure 8.4 with the brief note in Figure 8.6 that was sent to one of us authors and that, unfortunately, is all too typical of students' first attempts at drafting such a letter. A focus only on the researcher's needs in letters of this sort may be another reason for the poor return of questionnaires in some research projects.

The cover letter is extremely important. It should be carefully and thoughtfully composed and should stress the concerns of the recipient rather than any selfish interests of the sender. Some students forget this and, in doing so, unintentionally reveal their own self-centeredness.

FIGURE 8.5

Questionnaire response card

Dear [your name]

☐ Please send the questionnaire; I will be happy to cooperate.

☐ I am sorry, but I do not wish to answer the questionnaire.

Comments: _____

Date: _____ Name: _____

FIGURE 8.6

A poorly worded request
for cooperation



4. *If mailing your questionnaire, include a self-addressed envelope with return postage.* To impose on a person's time and spirit of cooperation and then to expect that person also to supply the envelope and pay the postage is unreasonable.

5. *Offer the results of your study.* In return for the investment of time and the courtesy of replying to your questions, offer to send your respondent a summary of your study's results. You might provide a check box, either at the beginning or at the end of your instrument, where a respondent can indicate the desire to have such a summary, together with a place for name and either mailing or e-mailing address. If anonymity is important, a mailed questionnaire might include a separate postcard on which the respondent can request the summary; this postcard should, of course, have a place for the respondent's name and address, along with the suggestion that the card be mailed separately from the questionnaire. For e-mailed questionnaires, a respondent can simply hit the "reply" button twice, once to return the completed questionnaire and a second time (perhaps a few hours later) to request the study's results.

6. *Be gently persistent.* Experts often suggest that when people don't initially respond to a questionnaire, you can increase your response rate by sending two follow-up reminders, perhaps sending each one out a week or two after the previous mailing (e.g., Neuman, 2011; Rogelberg & Luong, 1998). But if the questionnaire is meant to be anonymous, how do you know who has returned it and who has not?

To address this problem, many researchers put a different code number on each copy they send out and keep a list of which number they have sent to each person in their sample. When a questionnaire is returned, they remove the number and person's name from the list. When it is time to send a follow-up letter, they send it only to the people who are still on the list. Researchers should use the list of names and code numbers *only* for this purpose. At no point should they use it to determine who responded in what way to each question—a practice that violates the right to privacy discussed in Chapter 4.

Let's return to the solicitation letter and postcard in Figures 8.4 and 8.5. We have modeled them after a letter and postcard that an American University faculty member successfully used to get alumni feedback about the university's nursing program. After receiving a card that indicated willingness to cooperate, the faculty member immediately mailed the questionnaire. She kept a log of questionnaires mailed, the names and addresses of people to whom they were mailed, and the date of mailing. If a reply was not received within three weeks' time, she sent a reminder letter. The reminder was written in the same tone as the initial letter. An example of such a reminder letter appears in Figure 8.7.

FIGURE 8.7

A follow-up letter

A B C University
Address

Date

Dear [person's name],

We are all very busy these days, and sometimes we have trouble staying on top of our many commitments. Despite our best intentions, we may sometimes overlook something we have said we would do.

Three weeks ago I sent you a questionnaire asking for your input regarding your program at A B C University. To date I have not yet received your completed questionnaire. Perhaps you have simply mislaid it, or perhaps it has been lost in the mail—any one of several reasons might account for its delay in reaching me.

In any event, I am enclosing another copy of the questionnaire, along with another self-addressed, stamped envelope. I am hoping you can find fifteen minutes somewhere in your busy schedule to complete and return the questionnaire. I would really appreciate your personal insights and suggestions regarding your experiences in our program.

Thank you once again for your assistance and generosity in helping us enhance our program. And remember that if you have any questions, you can easily reach me at [telephone number] or [e-mail address].

Respectfully yours,

Your Signature

Your Name

The faculty member's follow-up letter brought results. She was being firm and persuasive, but with considerable skill and tact. Courtesy, understanding, and respect for others pay large dividends in a situation in which a researcher needs others' cooperation. This is especially true in questionnaire studies.

PRACTICAL APPLICATION Using the Internet to Collect Data for a Descriptive Study



In recent years, some researchers have collected descriptive-study data directly on the Internet. For instance, they may put a questionnaire on a website and ask people who visit the site to respond. One site providing links to a wide variety of online research projects is "Psychological Research on the Net."⁵ As we write this tenth edition of the book, the site is hosting descriptive research projects on such diverse topics as television commercial effectiveness, gender roles in marriage, cigarette smoking, eating disorders, attitudes and behaviors in the workplace, the psychological impact of automobile commuting, and coping strategies for dealing with stressful life circumstances. The website is maintained by John Krantz, Professor of Psychology at Hanover College, who checks to be sure that each project has been approved by the appropriate internal review board and incorporates informed consent procedures. There is no fee for using the site (J. Krantz, personal communication, May 2003).

Commercial websites for data collection are available as well. Two widely used websites for online surveys are SurveyMonkey (www.surveymonkey.com) and Zoomerang (zoomerang.com).

⁵ You can reach the site by going to the website of the Association for Psychological Science (www.psychologicalscience.org); click on "Psychology links," scroll down to "Other Sites of Interest," and then click on "Online Psychology Experiments." Alternatively, you can go directly to the site, which, as this book goes to press, is located at psych.hanover.edu/research/exponnet.html.

Both websites provide templates that make questionnaire design easy and enable a researcher to present a variety of item types (e.g., multiple-choice items, rating scales). They also include features for communicating with a preselected sample of participants (e.g., through e-mail invitations), as well as features through which the researcher can tabulate, statistically analyze, and download the results. For most research projects these websites charge a modest monthly fee.

Conducting a survey online has several advantages (Kraut et al., 2004). When the desired sample size is quite large, an online questionnaire is far more cost-effective than a mailed questionnaire. Often a questionnaire can be adapted based on a participant's previous responses; for instance, if a person responds *no* to the question "Do you smoke cigarettes?" the questionnaire software will subsequently skip questions related to smoking habits. Furthermore, some evidence indicates that online surveys yield data comparable to those obtained through face-to-face contact (Gosling, Vazire, Srivastava, & John, 2004).

If you choose to collect data on the Internet, keep in mind that your ethical standards must be just as rigorous as they would be if you were collecting data through face-to-face contacts or the postal service. Participants must be informed about and agree to the general nature of a study, perhaps by means of a website page that serves as an informed consent letter and a virtual "click to accept" button with which participants can indicate consent (Kraut et al., 2004). Also, participants' responses must remain as confidential as they would in any study. The *protection from harm* ethical standard can be especially troublesome in an online study, as it may be virtually impossible to determine that a participant has found a task or question extremely stressful or upsetting and needs some sort of follow-up intervention. Your research advisor and university's internal review board can help you work through ethical issues and develop appropriate precautions for any study that might potentially cause even minor harm or distress to participants.

Sampling, too, must be a source of concern in an online study. SurveyMonkey and Zoomerang enable a researcher to zero in on a defined set of participants—for example, by uploading a list of e-mail addresses to which the participation request will be sent. Other online research projects, such as those on the "Psychological Research on the Net" website mentioned earlier, are open to anyone who wants to participate. But in virtually any online study, the people who participate will not be representative either of a particular group of people or of the overall population of human beings (Gosling et al., 2004; McGraw, Tew, & Williams, 2000). After all, participants will be limited to people who (a) are comfortable with computers, (b) spend a fair amount of time on the Internet, (c) enjoy partaking in research studies, and (d) have been sufficiently enticed by your research topic to do what you ask of them. In cases where a questionnaire can be completed by anyone who has access to the Internet, many responders are apt to be college students who are earning course credit for their participation. In short, *your sample will be biased to some degree.*

Sampling is a concern for any researcher, but it is especially so for the researcher who wants to draw inferences about a large population. In the following section, we look at strategies for selecting an appropriate sample and consider how bias in sampling procedures can distort the data obtained.

Choosing a Sample in a Descriptive Study

The researcher who conducts a descriptive study wants to determine *the nature of how things are*. Especially when conducting survey research, the researcher may want to describe one or more characteristics of a fairly large population—perhaps the television viewing habits of 10-year-olds, the teaching philosophies of elementary school teachers, or the attitudes that visitors to Rocky Mountain National Park have about a shuttle bus system. Whether the population is 10-year-olds, elementary school teachers, or park visitors, we are talking about *very large* groups of people; for instance, more than 3 million people visit Rocky Mountain National Park every year.

In such situations, the researcher will, of course, usually not study the entire population of interest. Instead, he or she will select a subset, or **sample**, of the population. But the researcher can use the results obtained from the sample to make generalizations about the entire population

only if the sample is truly representative of the population. Here we are talking about a research study's *external validity*, a concept introduced in Chapter 4.

When phrasing their research problems, many novice researchers forget that they will be studying a sample rather than a population. They announce, for example, that their goal is

to survey the legal philosophies of the attorneys of the United States and to analyze the relationship of these several philosophical positions with respect to the recent decisions of the Supreme Court of the United States.

Someone who words a problem in this way has simply not thought through the meaning of the words themselves: "The attorneys of the United States"! The American Bar Association consists of approximately 400,000 attorneys distributed over more than 3.5 million square miles in the 50 states. As we look at the problem more closely, we begin to discern other, more serious difficulties. What are "legal philosophies"? How does one isolate these philosophies to study them? How can one show a "relationship of philosophical positions" to "recent decisions of the Supreme Court"? How will this relationship be expressed? Will it be expressed statistically? If so, how would one quantify "philosophical positions" and "decisions"? If not, then how will the relationship be shown?

The difficulty in the stated research problem basically arises out of the way the problem is worded. If, on the one hand, the researcher has said what he or she means, he or she proposes to survey "the attorneys"—all of them! If, on the other hand, the researcher intends to survey only a subset of the nation's attorneys, the statement of the problem should have said so with such qualifying words as *selected*, *representative*, *typical*, *certain*, *a random sample of*, and so on. Careful researchers say precisely what they mean. Notice the difference in the meaning between "The purpose of this research is to survey the representative legal philosophies of a random sample of attorneys . . ." and the original wording, "The purpose of this research is to survey the legal philosophies of the attorneys of the United States. . . ."

The specific sampling procedure used depends on the purpose of the sampling and a careful consideration of the parameters of the population. But in general, *the sample should be so carefully chosen that, through it, the researcher is able to see characteristics of the total population in the same proportions and relationships that they would be seen if the researcher were, in fact, to examine the total population.*

When you look through the wrong end of a set of binoculars, you see the world in miniature. If the lenses aren't precision-made and accurately ground, you get a distorted view of what you are looking at. In the same way, a sample should, ideally, be a population microcosm. If the sampling procedure isn't carefully planned, any conclusions the researcher draws from the data are likely to be distorted. Such distortion is known as *bias*. We discuss various sources of bias a bit later in the chapter. For the moment, however, we concern ourselves with various sampling designs.

Sampling Designs

Different sampling designs may be more or less appropriate in different situations. Here we consider eight approaches to sampling, which fall into two major categories: probability sampling and nonprobability sampling.

Probability Sampling

In **probability sampling**, every part of the population has the potential to be represented in the sample. The sample is chosen from the overall population by **random selection**—that is, it is chosen in such a way that each member of the population has an equal chance of being selected. When such a *random sample* is selected, the researcher can assume that the characteristics of the sample approximate the characteristics of the total population.

An analogy might help. Suppose we have a beaker containing 100 ml of water. Another beaker holds 10 ml of a concentrated acid. We combine the water and acid in proportions

of 10:1. After thoroughly mixing the water and acid, we should be able to extract 1 ml from any part of the solution and find that the sample contains 10 parts water for every 1 part acid. In the same way, if we have a population with considerable variability in race, wealth, education level, social standing, and other factors, and if we have a perfectly selected random sample—a situation usually more theoretical than logistically feasible—we will find in the sample those same characteristics that exist in the larger population, and we will find them in roughly the same proportions.

A sample is no more representative of the total population than the degree to which it has been randomly selected. There are, of course, many methods of random selection. For example, we could assign each person in the population a different number and then use an arbitrary method of picking certain numbers, perhaps by using a roulette wheel (if the entire population consists of 36 or fewer members) or drawing numbers out of a hat. Many computer spreadsheet programs (e.g., Excel) also provide a means of generating random numbers.

A widely used paper-and-pencil method of selecting a random sample is to use a **table of random numbers**, such as that presented in Table 8.2. The researcher typically does not start at the beginning of the table; instead, he or she identifies a starting point randomly. One fundamental principle must be kept in mind: *The purpose of randomness is to let blind chance determine the outcomes of the selection process to as great a degree as possible.* Hence, in determining a starting point for the selection of random numbers, *pure chance* must always initiate the process.

Consider the table of random numbers presented in Table 8.2. It includes 100 blocks of numbers, arranged in 10 rows and 10 columns. The rows and columns are numbered only to assist us in choosing a starting point for using the table. Any block within the table will be at the intersection of two of these “guide” numbers. To enter the table, we need an entry number of two digits; one digit will be used to designate the row, and the other the column, for the block at which we will begin.

But how do we find an entry number? Pull a dollar bill from your wallet. The one we have just pulled as we write this book has the serial number C 45 391827A. We choose the first two digits of the serial number, which makes the entry number 45. But which is the row and which is the column? We flip a coin. If it comes down heads, the first digit will designate the row; otherwise, the digit will designate the column. The coin comes down tails. This means that we will begin in the fourth column and the fifth row. The block where the two intersect is the block where we begin within the table (see Figure 8.8).

We don’t have to use a dollar bill to determine the entry point, of course. We could use any source of numbers, such as a telephone directory, a license plate, a friend’s social security number, or the stock quotations page in a newspaper. Only one rule governs the final determination of an entry point: *Pure chance dictates the choice.*⁶

Having determined the starting block, we must now consider the size of the proposed sample. If it is to be fewer than 100 individuals, we will need only two-digit numbers. If it is to be more than 99 but fewer than 1,000, we will need three digits to accommodate the sample size.

At this point, let’s go back to the total population to consider the group from which the sample is to be drawn. It will be necessary to designate individuals in some manner. A reasonable approach is to arrange the members of the population in a logical order—for instance, alphabetically by surname—and assign each member a serial number for identification purposes.

We are now ready for the random selection. We start with the upper left-hand digits in the designated starting block and work downward through the two-digit column in the rest of the table. If we need additional numbers, we proceed to the top of the next column, work our way down, and so on, until we have selected the sample we need. For purposes of illustration, we will assume that the total population consists of 90 individuals from which we will select a sample of 40. We will need random numbers of two digits each. Beginning in the upper left-hand corner of the designated block and remembering that only 90 individuals are in the total population,

⁶One of our readers has correctly pointed out that not all of the sources just suggested reflect strictly random numbers; rather, they may show a predictable pattern, with some numbers appearing more frequently than others. Nevertheless, using such a source ensures that the entry point into the table is chosen *arbitrarily*, eliminating any chance that the researcher might, either intentionally or unintentionally, tilt the sample selection in one direction or another.

TABLE 8.2 Random numbers table

1	2	3	4	5	6	7	8	9	0
38 01 08 18 62	82 52 01 82 29	02 56 28 19 24	88 42 92 63 07	23 99 90 93 57	78 10 48 55 21	29 84 46 25 60	50 24 21 62 40	56 88 09 95 57	01 78 78 86 35
51 10 40 21 24	04 69 90 71 43	04 78 84 81 84	41 31 82 31 79	40 79 15 65 18	28 86 32 62 70	55 33 27 42 50	35 64 53 48 31	15 24 21 63 23	17 59 94 69 95
92 12 24 41 22	72 73 42 19 31	84 53 15 16 78	98 77 86 76 75	66 51 70 90 93	87 83 76 49 08	32 14 91 93 50	47 49 89 17 52	05 46 40 33 05	77 27 15 16 31
94 72 67 55 42	52 52 26 41 89	32 38 14 58 97	71 94 93 90 49	66 42 05 69 12	28 08 82 64 25	29 79 96 81 66	29 86 00 94 07	84 54 57 62 22	70 80 69 62 87
77 75 72 87 20	86 70 64 02 44	89 24 08 35 53	32 96 00 80 78	48 68 39 83 83	56 47 81 77 42	39 59 64 69 02	10 19 94 50 51	40 34 19 81 31	38 54 97 22 17
92 44 11 50 85	05 70 08 70 64	91 81 58 48 16	61 87 43 52 08	60 42 80 59 20	34 28 49 91 25	19 73 63 65 72	86 97 13 99 06	08 03 07 62 52	33 68 42 59 38
60 04 91 78 89	71 40 77 32 66	11 30 10 01 21	49 12 88 73 47	68 54 94 32 12	20 77 44 49 64	06 44 46 32 33	29 70 98 08 59	64 16 94 48 45	32 22 40 94 41
88 39 28 16 75	92 57 77 21 95	56 93 73 19 17	94 62 18 76 31	00 85 74 86 15	67 31 42 98 65	39 93 31 64 30	75 28 25 82 38	25 99 97 68 63	98 45 70 95 46
28 49 94 80 45	16 20 72 31 64	74 04 31 00 86	79 79 33 98 04	55 26 34 15 70	76 70 92 40 97	79 70 00 64 33	24 29 27 30 25	86 96 27 96 51	67 14 98 05 49
71 23 62 84 00	35 01 41 52 70	05 91 02 35 24	53 74 60 11 41	36 34 18 08 46	80 67 41 60 61	29 06 20 72 42	53 96 06 98 42	81 13 97 04 71	27 27 63 85 31
96 96 31 54 02	00 91 92 76 35	15 68 62 95 24	32 12 73 38 93	77 48 20 37 47	38 54 96 41 02	71 55 92 52 72	83 88 14 67 44	60 20 86 72 56	55 39 83 52 20
42 24 86 51 17	60 92 31 00 55	68 99 02 84 40	43 90 67 66 07	93 58 14 66 19	24 57 83 41 75	81 17 09 22 06	62 51 63 29 07	48 78 30 39 28	07 12 36 20 41
48 04 03 20 10	64 51 11 11 69	31 07 84 90 36	84 56 50 31 14	58 67 15 93 17	36 07 58 97 44	23 49 17 02 59	56 11 41 67 04	46 66 29 36 60	25 24 38 04 18
24 82 46 95 57	73 54 42 99 51	33 72 12 89 86	63 44 34 78 78	62 23 04 30 78	31 36 77 96 03	07 00 17 98 53	40 50 52 29 34	71 77 22 68 97	18 04 81 99 29
81 67 50 87 94	68 85 73 36 83	04 80 31 52 66	70 04 32 61 56	87 67 45 06 85	39 67 65 57 22	79 93 95 80 66	97 60 51 95 17	94 12 85 08 16	20 28 59 36 47
13 45 91 94 98	03 88 43 86 42	98 65 79 38 10	91 12 81 98 30	31 10 49 95 83	14 60 19 90 87	95 99 76 05 42	73 99 91 91 12	62 89 71 43 96	57 23 16 02 17
72 24 96 81 87	52 68 73 61 17	51 94 47 58 01	13 88 40 38 70	51 11 02 00 63	66 16 46 92 60	98 74 78 50 78	17 07 25 83 52	18 65 27 13 89	55 91 08 06 12
55 05 71 44 11	66 04 57 07 14	92 20 82 92 33	30 08 96 22 15	50 11 40 49 53	07 91 17 27 71	46 03 95 14 94	47 94 10 17 21	23 03 23 69 95	40 24 39 29 92
92 36 97 30 14	88 41 90 80 35	07 75 80 26 05	94 14 31 80 07	55 41 14 57 90	52 12 46 94 70	44 39 99 50 47	19 52 92 92 66	26 20 82 10 07	95 75 73 66 54
89 92 58 84 08	73 41 65 61 95	43 97 81 33 05	74 67 22 23 00	86 26 66 99 63	08 89 21 14 83	40 34 50 36 28	86 84 82 22 34	25 82 86 47 85	06 38 07 97 44
26 80 83 98 13	77 10 83 11 03	00 44 16 60 42	30 88 02 35 74	26 31 51 32 71	73 00 22 00 49	22 32 76 37 24	26 97 70 87 34	04 87 06 52 84	90 74 09 74 06
10 71 47 27 12	75 45 51 26 23	19 59 86 21 70	98 76 96 40 12	97 70 77 57 74	44 22 44 97 62	42 28 17 09 55	76 04 82 14 40	76 81 74 96 79	38 81 96 85 99
41 04 81 62 78	06 77 53 27 14	52 71 25 82 93	52 02 54 04 07	51 23 05 30 59	52 40 03 90 79	48 44 83 66 71	57 21 01 45 68	50 25 03 32 31	96 92 82 16 62
33 85 26 45 29	22 81 84 43 83	11 60 71 38 45	93 07 22 30 42	99 30 52 21 40	89 18 59 29 95	23 43 14 58 06	56 71 31 48 25	27 57 85 25 76	72 63 83 95 00
45 50 56 50 40	26 05 25 93 64	78 17 59 58 83	80 47 43 71 41	03 06 18 79 54	16 23 18 55 60	08 65 50 82 07	72 25 01 31 09	53 44 59 31 94	16 87 80 32 14
92 53 64 22 75	68 24 20 99 94	21 95 33 19 10	23 01 49 45 26	34 34 81 38 89	60 81 53 29 72	13 93 48 56 37	67 99 80 12 23	12 06 07 80 23	39 20 96 85 52
99 41 50 17 32	32 35 95 10 22	34 50 81 80 34	12 13 53 83 62	86 07 50 83 86	82 21 93 36 48	56 58 87 36 67	55 45 12 45 69	88 80 97 65 58	28 90 74 84 74
82 40 93 92 43	88 84 79 42 86	15 16 07 30 59	92 54 78 72 92	34 17 73 57 56	31 46 75 31 06	65 20 76 65 17	68 67 24 41 65	83 90 16 72 41	62 28 75 63 34
86 48 92 62 86	04 86 51 39 73	61 17 22 69 09	93 39 10 59 24	06 79 60 36 84	93 97 68 61 23	15 40 89 22 72	49 06 40 32 58	59 56 50 32 05	39 01 75 86 33
89 52 47 59 01	52 00 88 05 98	80 62 64 78 59	33 74 08 06 67	41 77 42 65 24	77 95 26 51 27	38 18 74 92 72	33 78 77 53 41	88 04 45 86 61	44 98 66 68 80
22 21 03 21 90	20 37 19 57 62	99 37 27 35 26	12 68 43 81 53	71 92 33 99 26	78 69 45 98 85	87 39 43 19 80	24 88 10 64 90	78 18 95 35 38	72 70 45 72 58
90 53 60 07 99	17 18 66 37 53	74 41 09 90 62	44 56 94 44 36	31 74 10 57 63	79 62 84 26 54	88 56 10 22 55	03 90 19 50 76	07 50 95 86 69	90 00 55 69 31
26 37 79 96 33	88 52 34 17 95	31 23 24 58 77	75 88 64 08 53	05 81 86 00 75	87 59 13 45 98	63 38 70 31 41	52 50 16 84 22	73 57 74 10 88	98 63 77 24 70
43 02 16 18 53	51 79 03 90 34	30 34 88 89 36	85 70 92 05 82	01 57 58 98 83	81 19 21 57 98	43 34 00 57 50	20 67 30 37 14	39 84 53 12 90	78 65 10 23 09
20 95 76 51 15	97 32 97 58 43	39 53 93 17 32	03 16 16 65 24	34 21 10 91 88	58 22 98 68 41	54 15 20 20 14	46 70 52 78 60	06 90 74 56 50	76 31 58 65 83
37 42 58 13 23	44 82 67 97 98	54 38 79 75 11	75 72 16 32 68	74 59 37 13 10	34 52 36 07 27	98 37 33 09 60	58 31 70 14 75	49 07 71 35 11	95 06 72 41 00
31 51 93 66 48	35 56 04 55 20	29 19 24 55 97	32 79 45 92 58	76 37 11 35 79	05 68 08 06 87	64 66 88 25 52	14 82 30 95 52	01 07 97 96 23	64 99 36 57 51
26 19 62 67 67	70 61 22 01 70	18 84 36 99 94	00 57 54 28 92	45 72 40 45 06	60 23 19 24 02	03 63 69 83 81	80 85 48 37 86	47 97 87 19 83	66 94 71 49 11
42 55 29 27 78	76 81 16 63 13	04 93 73 21 04	84 91 77 67 10	50 00 21 92 50	96 91 83 66 19	00 59 90 66 01	65 48 83 28 97	72 80 43 01 65	55 64 58 40 56
90 27 74 26 17	61 75 82 44 09	42 16 28 32 77	98 36 41 39 73	97 45 85 85 08	15 25 93 29 16	19 84 12 57 24	61 33 44 76 24	54 70 39 17 47	08 99 19 55 89
91 78 11 76 19	19 09 29 73 14	94 09 50 52 42	00 57 41 82 55	34 52 10 83 68	17 58 82 71 68	72 43 90 64 99	10 49 51 53 27	52 04 26 28 62	10 70 98 42 50
53 42 75 49 95	48 74 48 88 94	52 58 34 96 85	31 34 96 82 31	47 55 84 13 48	86 83 54 45 38	32 46 69 85 86	24 47 32 60 28	29 45 36 58 21	30 84 54 36 35
09 48 74 77 55	89 30 05 25 52	93 01 06 07 66	98 09 08 65 43	51 91 90 05 91	79 45 73 93 80	89 71 17 57 63	79 51 34 10 11	12 97 06 05 00	96 77 28 11 35
23 43 73 69 25	89 55 18 26 26	87 69 88 38 12	40 83 94 06 23	96 55 93 14 80	49 89 86 62 40	81 15 86 62 40	81 15 69 78 10	18 40 92 93 88	13 01 90 70 82
59 47 86 32 53	50 76 46 09 82	07 81 75 38 35	53 87 71 45 94	64 24 20 27 23	45 39 48 71 66	76 34 85 81 21	04 93 90 43 54	92 77 48 04 07	70 52 35 99 32
80 73 71 10 81	52 63 88 33 59	35 11 58 95 64	10 50 63 30 30	07 31 89 22 05	92 53 64 22 75	68 24 20 99 94	21 95 33 19 10	23 01 49 45 26	34 34 81 38 89
67 44 30 85 08	18 69 78 50 15	88 24 02 62 46	36 89 70 02 89	78 50 98 86 01	99 41 50 17 32	32 35 95 10 22	34 50 81 80 34	12 13 53 83 62	86 07 50 83 86
35 15 98 11 06	18 64 51 44 49	37 38 41 40 68	14 65 00 37 23	87 63 85 50 66	87 40 93 92 43	88 84 79 42 86	15 16 07 30 59	92 54 78 72 92	34 17 73 57 56
66 17 63 96 90	44 57 17 04 05	85 84 64 44 05	34 06 31 13 81	92 30 29 21 61	86 48 92 62 86	04 86 51 39 73	61 17 22 69 09	03 39 10 59 24	06 79 60 36 84
53 52 48 13 33	85 33 38 32 53	77 67 16 86 82	50 81 83 25 84	53 65 33 12 88	89 52 47 59 61	52 00 88 05 98	80 62 64 78 59	33 74 08 06 67	41 77 42 65 24

FIGURE 8.8

Choosing the starting point in a random numbers table

	1	2	3	4	5
	38 01 08 18 62	82 52 01 82 29	02 56 28 19 24	88 42 92 63 07	23 99 90 93 57
	51 10 40 21 24	04 69 90 71 43	04 78 84 81 84	41 31 82 31 79	40 79 15 65 18
1	92 12 24 41 22	72 73 42 19 31	84 53 15 16 78	98 77 86 76 75	66 51 70 90 93
	94 72 67 55 42	52 52 26 41 89	32 38 14 58 97	71 94 93 90 49	66 42 05 69 12
	77 75 72 87 20	86 70 64 02 44	89 24 08 35 53	32 96 00 84 78	48 68 39 83 83
	92 44 11 50 85	05 70 08 70 64	91 81 58 48 16	61 87 48 52 08	60 42 80 59 20
	60 04 91 78 89	71 40 77 32 66	11 30 10 01 21	49 12 88 73 47	68 54 94 32 12
2	28 39 28 16 75	92 57 77 21 95	56 93 73 19 17	94 62 18 76 31	00 85 74 86 15
	88 49 94 80 45	16 20 72 31 64	74 04 31 00 86	97 79 33 98 04	55 26 34 15 70
	71 23 62 84 00	35 01 41 52 70	05 91 02 35 24	53 74 60 11 41	36 34 18 08 46
	96 96 31 54 02	00 91 92 76 35	15 68 62 95 24	32 12 73 38 93	77 48 20 37 37
	42 24 86 51 17	60 92 31 00 55	68 99 02 84 40	43 90 67 66 07	93 58 14 66 19
3	48 04 03 20 10	64 51 11 11 69	31 07 84 90 36	84 56 50 31 14	58 67 15 93 17
	24 82 46 95 57	73 54 42 99 51	33 72 12 89 86	63 44 34 78 78	62 23 04 30 78
	81 67 50 87 94	68 85 73 36 83	04 80 31 52 66	70 04 32 61 56	87 67 45 06 85
	13 45 91 94 98	03 88 43 86 42	98 65 79 38 10	91 12 81 98 30	31 10 49 95 83
	72 24 96 81 87	52 68 73 61 17	51 94 47 58 01	13 88 40 38 70	51 11 02 00 63
4	55 05 71 44 11	66 04 57 07 14	92 20 82 92 33	30 08 96 22 15	50 11 40 49 63
	92 36 97 30 14	88 41 90 80 35	07 75 80 26 05	94 14 31 80 07	55 41 14 57 90
	89 92 58 84 08	73 41 65 61 95	43 97 81 33 05	74 67 22 23 00	86 26 66 99 63
	26 80 83 98 13	77 10 83 11 03	00 44 16 60 42	30 88 02 35 74	26 31 51 32 71
	10 71 47 27 12	75 45 51 26 23	19 59 86 21 70	98 76 96 40 12	97 70 77 57 74
5	44 94 81 62 78	00 77 55 27 14	52 71 25 82 30	52 02 54 04 07	51 23 05 30 59
	33 85 26 45 29	22 81 84 43 83	11 60 71 38 45	93 07 22 30 42	99 30 52 21 40
	45 50 56 50 40	26 05 25 93 64	78 17 59 58 83	80 47 43 71 41	03 06 18 79 54
	92 53 64 22 75	68 24 20 99 94	21 95 33 19 10	23 01 49 45 26	34 34 81 38 89
	99 41 50 17 32	32 35 95 10 22	34 50 81 80 34	12 13 53 83 62	86 07 50 83 86
6	82 40 93 92 43	88 84 79 42 86	15 16 07 30 59	92 54 78 72 92	34 17 73 57 56
	86 48 92 62 86	04 86 51 39 73	61 17 22 69 09	03 39 10 59 24	06 79 60 36 84
	89 52 47 59 01	52 00 88 05 98	80 62 64 78 59	33 74 08 06 67	41 77 42 65 24
	22 21 03 21 90	20 37 19 57 62	99 37 27 35 26	12 68 43 81 53	71 92 33 99 26
	90 53 60 07 99	17 18 66 37 53	74 41 09 90 62	44 56 94 44 36	31 74 10 57 63
7	26 37 79 96 33	88 52 34 17 95	31 23 24 58 77	75 88 64 08 53	05 81 86 00 75
	43 02 16 18 53	51 79 03 90 34	30 34 88 89 36	85 70 92 05 82	01 57 58 98 83
	20 95 76 51 15	97 32 97 58 43	39 53 93 17 32	03 16 16 65 24	34 21 10 91 88

we see that the first number in the leftmost column is 30, so we choose individual number 30 in the population. The next number (98) does not apply because only 90 persons are in the population. Our next choice is 52, we ignore 93, and then we choose 80. Proceeding to the next block down, we choose 23 and 12, ignore 92, choose 3 and 33. We continue down the column and proceed to any additional columns we need, ignoring the numbers 91–99 and 00 and any numbers we have already selected, until we get a sample of 40.

We have probably said enough about the use of a random numbers table. We turn now to specific probability sampling techniques.

Simple random sampling. Simple random sampling is exactly the process just described: Every member of the population has an equal chance of being selected. Such an approach is easy when the population is small and all of its members are known. For example, one of us authors once used it in a study to evaluate the quality of certain teacher training institutes during the summer of 1992 (Cole & Ormrod, 1995). Fewer than 300 people had attended the institutes,

and we knew who and where they all were. But for very large populations—for instance, all 10-year-olds or all lawyers—simple random sampling is neither practical nor, in many cases, possible.

Stratified random sampling. Think of grades 4, 5, and 6 in a public school. This is a *stratified population*. It has three different layers (*strata*) of distinctly different types of individuals. In stratified random sampling, the researcher samples equally from each one of the layers in the overall population.

If we were to sample a population of fourth-, fifth-, and sixth-grade children in a particular school, we would assume that the three strata are roughly equal in size (i.e., there are similar numbers of children at each grade level), and so we would take equal samples from each of the three grades. Our sampling method would look like that in Figure 8.9.

Stratified random sampling has the advantage of guaranteeing equal representation of each of the identified strata. It is, of course, most appropriate when those strata are roughly equal in size in the overall population as well.

Proportional stratified sampling. In the simple stratified random sampling design just described, all strata of the population are essentially equal in size. But now we come to a different situation. Consider a community that has, for example, 1,000 Jewish people, 2,000 Catholics, and 3,000 Protestants. Let's imagine a survey situation. A local newspaper publishes a section dealing with interfaith church news, religious events, and syndicated articles of interest to the religious community in general. The editor wants to obtain certain information and opinions from the paper's readers and so decides to conduct a survey.

In this situation, the editor chooses his sample in accordance with the proportions of each religious group in the paper's readership. For every Jewish person, there should be two Catholics and three Protestants. In this situation, the people are not obviously segregated into the different strata, so the first step is to identify the members of each stratum and then select a random sample from each one. Figure 8.10 schematically represents this type of sampling.

Cluster sampling. Sometimes the population of interest is spread out over a large area. It may not be feasible to make up a list of every person living within the area and, from the list,

FIGURE 8.9

Stratified random sampling design

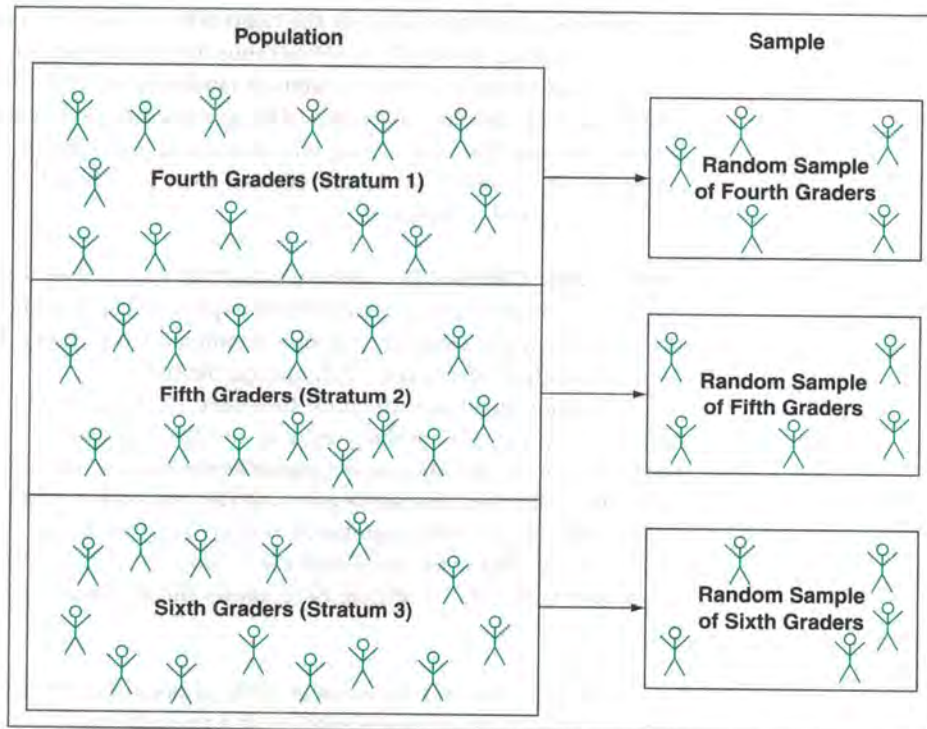
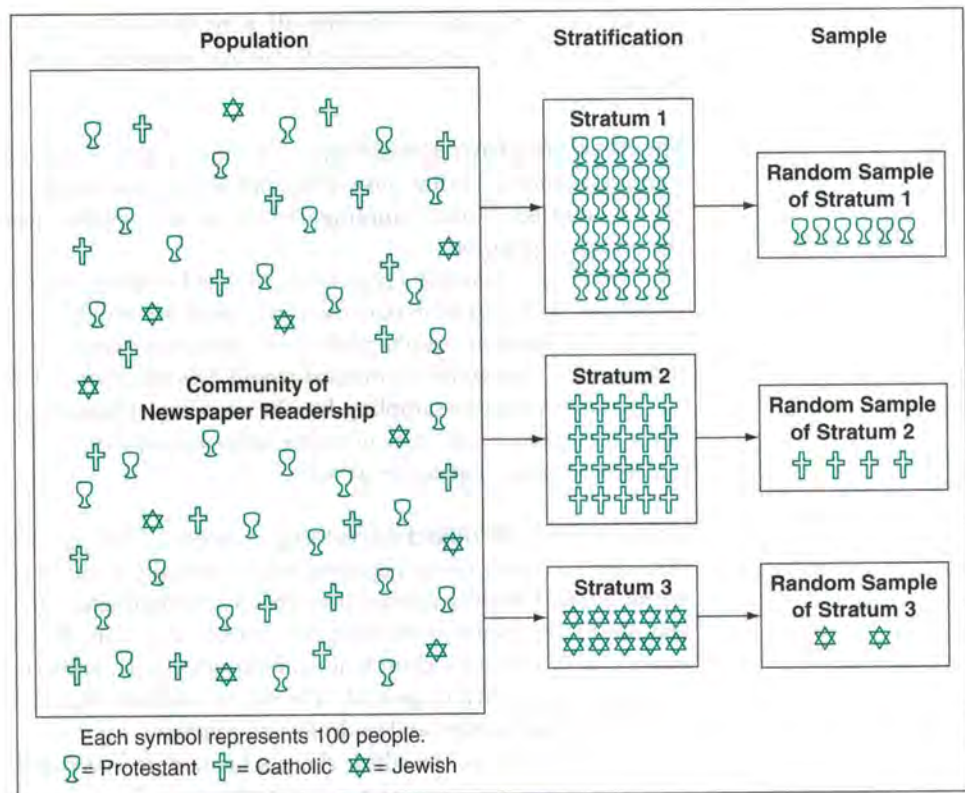


FIGURE 8.10

Proportional stratified sampling design



select a sample for study through normal randomization procedures. Instead, we might obtain a map of the area showing political boundaries or other subdivisions. We can then subdivide an expansive area into smaller *units*. For example, a city can be subdivided into precincts, clusters of city blocks, or school boundary areas; a state can be divided into counties or townships. In cluster sampling, it is important that the clusters be as similar to one another as possible, with each cluster containing an equally heterogeneous mix of individuals.

A subset of the identified clusters is randomly selected. The sample consists of the people within each of the chosen clusters. Using our example of community religious groups, let's assume the community is a large city that we have divided into 12 areas, or clusters. We randomly select clusters 1, 4, 9, and 10, and their members become our sample. This sampling design is depicted in Figure 8.11.

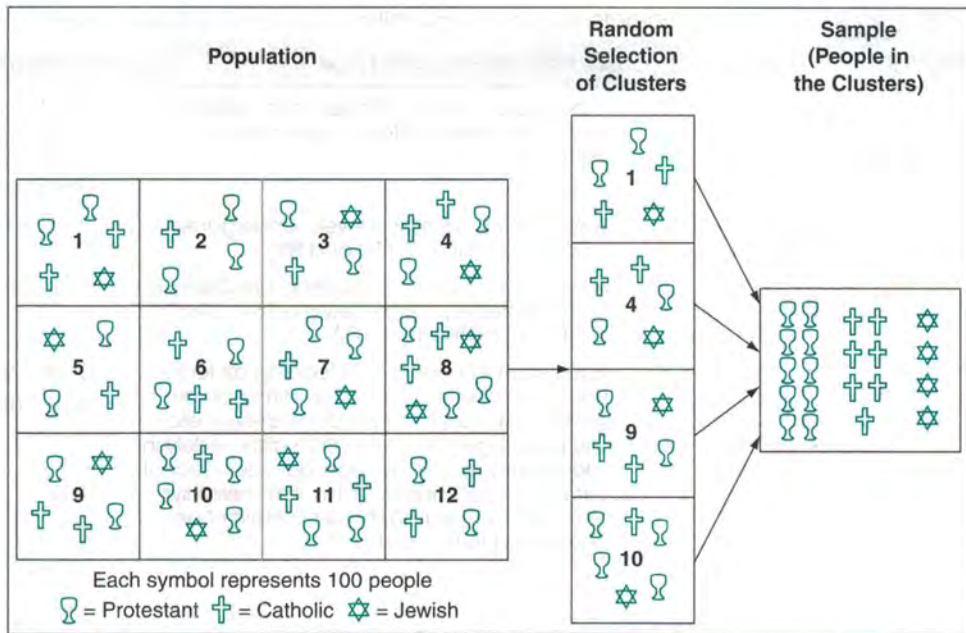
Systematic sampling. Systematic sampling involves selecting individuals—or perhaps clusters—according to a predetermined sequence. The sequence must originate by chance. For instance, we might create a randomly scrambled list of units that lie within the population of interest and then select every 10th unit on the list.

Let's take the cluster diagram presented in Figure 8.11. The population has 12 cells, or clusters. Half the cell numbers are odd, and the other half are even. Using the systematic sampling technique, we choose, by *predetermined sequence*, the clusters for sampling. Let's toss a coin. Heads dictates that we begin with the first odd-numbered cluster; tails dictates that we begin with the first even-numbered cluster. The coin comes down tails, which means that we start with the first even-numbered digit, which is 2, and select the systematically sequential clusters 4, 6, 8, 10, 12. Figure 8.12 shows the systematic sampling design as used for sampling clusters.

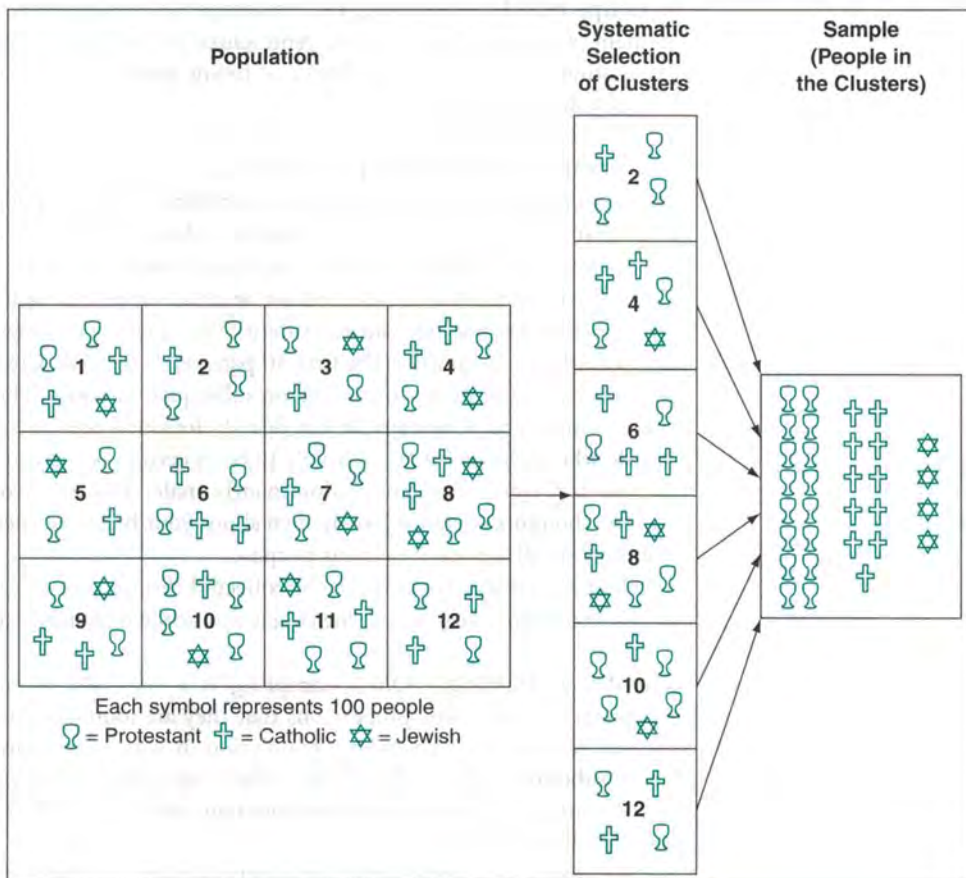
Table 8.3 identifies the various kinds of populations for which different probability sampling techniques might be appropriate in a research study.

FIGURE 8.11

Cluster sampling design

**FIGURE 8.12**

Systematic sampling design



A sampling design should not be chosen blindly or willy-nilly. Each of the designs just described is uniquely suited to a particular kind of population, and you should therefore consider the nature of your population when selecting your sampling technique. The design diagrams in Figures 8.9 through 8.12 can assist you in your choice.

TABLE 8.3 Population characteristics and probability sampling techniques appropriate for each population type

Population Characteristic	Example of Population Type	Appropriate Sampling Technique(s)
1. Population is generally a homogeneous group of individual units.	A particular variety of flower seeds, which a researcher wants to test for germination potential.	• Simple random sampling • Systematic sampling of individual units (when large populations of human beings are involved)
2. Population contains definite strata that are approximately equal in size.	A school with six grade levels: kindergarten, first, second, third, fourth, and fifth.	• Stratified random sampling
3. Population contains definite strata that appear in different proportions within the population.	A community in which residents are Catholic (25%), Protestant (45%), Jewish (15%), Muslim (5%), or nonaffiliated (10%).	• Proportional stratified sampling
4. Population consists of discrete clusters with similar characteristics. The units within each cluster are as heterogeneous as units in the overall population.	Travelers in the nation's 20 leading air terminals. (It is assumed that all air terminals are similar in atmosphere, purpose, design, etc. The passengers who use them differ widely in such characteristics as age, gender, national origin, socioeconomic status, and belief system, with such variability being similar from one airport to the next.)	• Cluster sampling • Systematic sampling (of clusters)

Nonprobability Sampling

In **nonprobability sampling**, the researcher has no way of predicting or guaranteeing that each element of the population will be represented in the sample. Furthermore, some members of the population have little or no chance of being sampled. Following are three common forms of nonprobability sampling.

Convenience sampling. Convenience sampling—also known as *accidental sampling*—makes no pretense of identifying a representative subset of a population. It takes people or other units that are readily available—for instance, those that arrive on the scene by mere happenstance.

Convenience sampling may be quite appropriate for some research problems. For example, suppose you own a small restaurant and want to sample the opinions of your patrons on the quality of food and service at your restaurant. You open for breakfast at 6 a.m., and on five consecutive weekdays you question the first 40 patrons who arrive. Customers who have on one occasion expressed an opinion are eliminated on subsequent arrivals. The opinions you eventually get are from 36 men and 4 women. It is a heavily lopsided poll in favor of men, perhaps because the people who arrive at 6 a.m. are likely to be in certain occupations—laborers, construction workers, truck drivers—that are predominantly male. The data from this convenience sample give you the thoughts of robust, hardy men about your breakfast menu—that's all. Yet such information may be all you need for your purpose.

Not all research data need to be collected through careful, thoughtful sampling procedures. But without such safeguards, the conclusions drawn from the research may not be trustworthy.

Quota sampling. Quota sampling is a variation of convenience sampling. It selects respondents in the same proportions that they are found in the general population, but not in a random fashion. Let's consider a population in which the number of African Americans equals the number of European Americans. Quota sampling would choose, say, 20 African Americans and 20 European Americans, but without any attempt to select these individuals randomly from the overall population.

For example, suppose you are a reporter for a television station. At noon, you position yourself with microphone and television camera beside Main Street in the center of a particular city. As people pass, you interview them. The fact that people in the two categories may come in clusters of two, three, or four is no problem. All you need are the opinions of 20 people from each category. This type of sampling regulates only the size of each category within the sample; in every other respect, the selection of the sample is nonrandom and, in most cases, convenient.

Purposive sampling. In purposive sampling, people or other units are chosen, as the name implies, for a particular *purpose*. For instance, we might choose people who we have decided are “typical” of a group or those who represent diverse perspectives on an issue.

Pollsters who forecast elections frequently use purposive sampling: They may choose a combination of voting districts that, in past elections, has been quite useful in predicting the final outcomes.

Purposive sampling may be very appropriate for certain research problems. However, the researcher should always provide a rationale explaining why he or she selected the particular sample of participants.

Sampling in Surveys of Very Large Populations

Nowhere is sampling more critical than in survey research. Frequently the researcher reports that x percent of people believe such-and-such, that y percent do so-and-so, or that z percent are in favor of a particular political candidate. *Such percentages are meaningless unless the sample is representative of the population about which inferences are to be drawn.*

A basic rule governs survey research: Nothing emerges from a long, involved study that is any better than the care, precision, and thought that went into the basic planning of the research design and the selection of the population. The results of a survey are no more trustworthy than the representativeness of the sample. Population parameters and sampling procedures are of paramount importance and become critical factors in the success of the study.

But now imagine that a researcher wants to conduct a survey of the country's *entire adult population*. How can the researcher possibly hope to get a random, representative sample of such a large group of people? The Survey Research Center of the University of Michigan's Institute for Social Research (1976) has used what it calls a *multistage sampling of areas*:

1. *Primary area selection.* The country is divided into small “primary areas,” each consisting of a specific county, a small group of counties, or a large metropolitan area. A predetermined number of these areas are randomly selected.
2. *Sample location selection.* Each of the selected primary areas is divided into smaller sections (“sample locations”), such as specific towns. A small number of these locations is randomly selected.
3. *Chunk selection.* The sample locations are divided into even smaller “chunks” that have identifiable boundaries such as roads, streams, or the edges of a city block. Most chunks have 16 to 50 dwellings, although the number may be larger in large cities. Once again, a random sample is selected.
4. *Segment selection.* Chunks are subdivided into areas containing a relatively small number of dwellings, and some of these “segments” are, again, chosen randomly.
5. *Housing unit selection.* Approximately four dwellings are selected (randomly, of course) from each segment, and the residents of those dwellings are asked to participate in the survey. If a doorbell is unanswered, the researcher returns at a later date and tries again.

As you may have deduced, the approach just described is a multistage version of cluster sampling (see Figure 8.13). At each stage of the game, units are selected randomly. “Randomly” does *not* mean haphazardly or capriciously. Instead, a mathematical procedure is employed to ensure that selection is entirely random and the result of blind chance. This process should yield a sample that is, in all important respects, representative of the country's population.

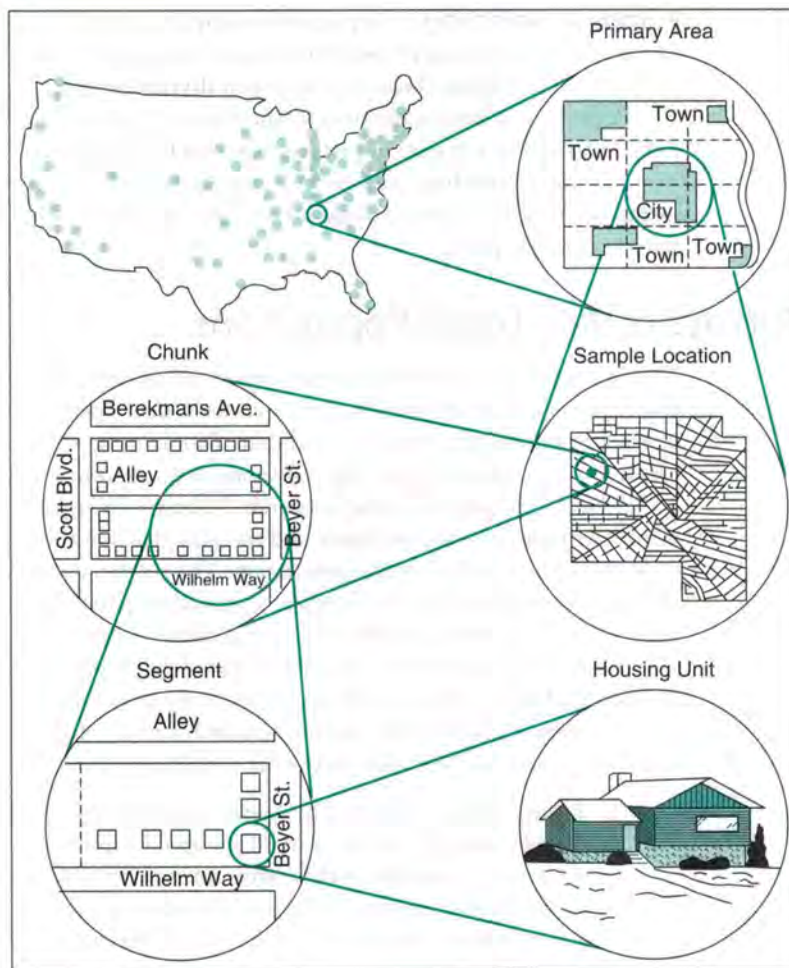
Identifying a Sufficient Sample Size

A basic rule in sampling is: *The larger the sample, the better.* But such a generalized rule is not very helpful to a researcher who must make a practical decision about a specific research situation. Obviously, we need to provide more guidance here. Gay, Mills, and Airasian (2009, p. 133) have offered the following guidelines for selecting a sample size, which we'll refer to by the symbol N :

- For smaller populations, say, $N = 100$ or fewer, there is little point in sampling; survey the entire population.

FIGURE 8.13**Multistage sampling**

From the *Interviewer's Manual* (Rev. ed., p. 36) by the Survey Research Center, Institute for Social Research, 1976, Ann Arbor: University of Michigan. Reprinted with permission.



- If the population size is around 500 (give or take 100), 50% should be sampled.
- If the population size is around 1,500, 20% should be sampled.
- Beyond a certain point (about $N = 5,000$), the population size is almost irrelevant and a sample size of 400 will be adequate.

Generally speaking, then, the larger the population, the smaller the percentage—but *not* the smaller the number!—one needs to get a representative sample.

To some extent, the size of an adequate sample depends on how homogeneous or heterogeneous the population is—how alike or different its members are with respect to the characteristics of research interest. If the population is markedly heterogeneous, a larger sample will be necessary than if the population is fairly homogeneous. Important, too, is the degree of precision with which the researcher wishes to draw conclusions or make predictions about the population under study.

Statisticians have developed formulas for determining the desired sample size for a given population. Such formulas are beyond the scope of this book, but you can find them in many introductory statistics books (e.g., Lind, 2006; Triola, 2008) and on many Internet websites (e.g., search “calculating sample size”).

Bias in Research Sampling

Look once again at the five steps in the University of Michigan’s Survey Research Center procedure for obtaining a sample in a national survey. Notice the last sentence in the fifth step: “If a doorbell is unanswered, the researcher returns at a later date and tries again.” The researcher does *not* substitute one housing unit for another; doing so would introduce *bias* into the sampling design. The center’s *Interviewer’s Manual* describes such bias well:

The house on the muddy back road, the apartment at the top of a long flight of stairs, the house with the growling dog outside must each have an opportunity to be included in the sample. People who live on back roads can be very different from people who live on well paved streets, and people who stay at home are not the same as those who tend to be away from home. If you make substitutions, such important groups as young men, people with small families, employed women, farmers who regularly trade in town, and so on, may not have proportionate representation in the sample. (Survey Research Center, 1976, p. 37)

In research, **bias** is any influence, condition, or set of conditions that singly or in combination distort the data. Data are, in many respects, delicate and sensitive to unintended influences. We talk about the hard facts, the solid truth, and yet every researcher soon learns that data are neither so hard nor so solid as the phrases might suggest. Data are highly susceptible to distortion.

Bias can creep into a research project in a variety of subtle and undetected ways. It can be easily overlooked by even the most careful and conscientious researcher. For instance, when conducting an interview, the researcher's personality may affect the interviewee's responses. In asking questions, the researcher's tone of voice or emphasis within a sentence may influence how a respondent replies.

Bias attacks the integrity of the facts. It is especially vicious when it enters surreptitiously into the research system and goes undetected. It can render suspect even the most carefully planned research effort.

Among the conditions that lead to bias is any influence that may have disturbed the randomness with which a sample population has been selected. Here we are talking about **sampling bias**.

The best way to appreciate sampling bias is to see it at work. Suppose that a researcher wants to conduct a survey of a certain city's population and decides to use the city telephone book as a source for selecting a random sample. She opens to a page at random, closes her eyes, puts her pencil down on the page, and selects the name that comes closest to the pencil point. You can't get much more random than that, she reasons. But the demon of bias is there. Her possible selections are limited to people who are listed in the phone book. People with very low income levels won't be adequately represented because some of them cannot afford telephone service. Nor will wealthy individuals be proportionally represented because many of them have unlisted numbers. And, of course, people who use only cell phones—people who, on average, are fairly young—aren't included in the phone book. Hence, the sample will consist of a greater percentage of people at middle-income levels and in older age-groups than exists in the city's overall population (e.g., Keeter, Dimock, Christian, & Kennedy, 2008).

As noted earlier in the chapter, studies involving Internet-based questionnaires are apt to be biased in favor of computer-literate individuals with easy access to the Internet. Studies involving mailed questionnaires frequently fall victim to bias as well, often without the researcher's awareness. Let's take a simple situation. Suppose a questionnaire is sent to 100 citizens, asking, "Have you ever been audited by the Internal Revenue Service (IRS) to justify your income tax return?" Of the 70 questionnaires returned, 35 are from people who indicate they have been audited, whereas 35 are from people who indicate they have never been audited. The researcher might therefore conclude that 50% of American citizens are likely to be audited by the IRS at one time or another.

The researcher's generalization may not be accurate. We need to consider how the nonrespondents—30% of the original sample—might be different from those who responded to the questionnaire. Many people consider an IRS audit to be a reflection of their integrity. Perhaps for this reason, some individuals in the researcher's sample may not have wanted to admit that they had been audited and so tossed the questionnaire into the wastebasket. If previously audited people were less likely to return the questionnaire than nonaudited people, the sample was biased, and thus the results didn't accurately represent the truth of the matter. Perhaps, instead of a 50-50 split, an estimate of 65% (people audited) versus 35% (people not audited) is more accurate. The data the researcher obtained do not enable the researcher to make such an estimate, however.

The examples just presented illustrate two different ways in which bias may creep into the research sample. In the case of the telephone survey, *sample selection* itself was biased because not

everyone in the population had an equal chance of being selected. In fact, people not listed in the telephone directory had *zero* chance of being selected. Here we see the primary disadvantage of nonprobability sampling, and especially of convenience sampling: People who happen to be readily available for a research project—those who are in the right place at the right time—are almost certainly *not* a random sample of the overall population.

In the example concerning IRS audits, *response rate*—and, in particular, potential differences between respondents and nonrespondents—was the source of bias. In that situation, the researcher's return rate of 70% was quite high. More often, however, the return rate in a questionnaire study is 50% or less, and the more nonrespondents there are, the greater the likelihood of bias. Likewise, in telephone surveys, a researcher won't necessarily reach certain people even with 10 or more attempts, and those who *are* eventually reached won't all agree to an interview (Witt & Best, 2008).

Nonrespondents to *mailed questionnaires* may be different from respondents in one or more ways (Rogelberg & Luong, 1998). They may have less interest in the topic being studied. They may have illnesses, disabilities, or language barriers that prevent them from responding. And on average, they have lower educational levels. In contrast, people who are hard to reach *by telephone* are apt to be young working adults who are *more* educated than the average individual (Witt & Best, 2008). To the extent that such characteristics affect how people respond to a survey, bias will exist in the data the survey yields.

Acknowledging the Probable Presence of Bias

It is almost impossible for people to live in this world without coming into contact with disease-bearing germs and other microorganisms. Likewise, in the research environment, the researcher cannot avoid having data contaminated by bias of one sort or another. What is unprofessional, however, is for the researcher to fail either to recognize or to acknowledge the likelihood of biased data in the study. When formulating conclusions about the data, a researcher must be sure to consider the effect that bias may have had in distorting the data.

In survey research, you should *always* report the percentages of people who have and have not consented to participate, such as those who have agreed and refused to be interviewed or those who have and have not returned questionnaires. Furthermore, you should be candid about possible sources of bias that result from differences between participants and nonparticipants. Rogelberg and Luong (1998) have suggested several strategies for identifying possible bias in questionnaire research; we list three especially useful ones here:

1. Carefully scrutinize the questionnaire for items that might be influenced by one's education level, interest in the topic, or other factors that frequently distinguish respondents from nonrespondents.
2. Compare the responses on questionnaires that were returned quickly with responses on those that were returned later, perhaps after a second reminder letter or after the deadline you imposed. The late ones may, to some extent, reflect the kinds of responses that nonrespondents would have given. Significant differences between the early and late questionnaires probably indicate bias in your results.
3. Randomly select a small number of nonrespondents and try to contact them by mail or telephone. Present an abridged version of your survey, and, if some people reply, match their answers against those in your original set of respondents.

One of us authors once used a variation on the third strategy in the study of summer training institutes mentioned earlier in the chapter (Cole & Ormrod, 1995). A research assistant had sent questionnaires to all attendees at one summer's institutes so that the institutes' leaders could improve the training sessions the following year, and she had gotten a return rate of 50%. She placed telephone calls to small random samples of both respondents and nonrespondents and asked a few of the questions that had been on the questionnaire. She obtained similar responses from both groups, leading the research team to conclude that the responses to the questionnaire were probably fairly representative of the entire population of institute participants.

Good researchers demonstrate their integrity by admitting, without reservation, that bias is omnipresent and may well have influenced their findings. Ideally, they point out precisely how bias may have infiltrated the research design. With this knowledge, other scholars can realistically appraise the research and judge its merits.

PRACTICAL APPLICATION Population Analysis for a Descriptive Study

Select a particular population and conduct an analysis of its structure and characteristics. Analyze the population you have chosen by completing the following checklist.

✓ CHECKLIST

Analyzing a Population

_____ 1. On the following line, identify the particular population you have chosen:

_____ 2. Now answer the following questions with respect to the *structure of the population*:

	Yes	No
a. Is the population a relatively homogeneous group of individuals or other units?	_____	_____
b. Could the population be considered to consist generally of equal "layers," each of which is fairly homogeneous in composition?	_____	_____
c. Could the population be considered to be composed of separate homogeneous layers differing in size and number of units comprising them?	_____	_____
d. Could the population be envisioned as isolated islands or clusters of individual units, with the clusters being similar to one another in composition?	_____	_____

_____ 3. Through what means would you extract a representative sample from the total population? Describe your procedure on the following lines:

_____ 4. Refer to Table 8.3. Is your sampling procedure appropriate for the characteristics of the population? _____ Yes _____ No

_____ 5. Have you guaranteed that your sample will be chosen by chance and yet will be representative of the population? _____ Yes _____ No

_____ 6. If the preceding answer is yes, explain how this will be done.

_____ 7. Indicate what means will be employed to obtain the information you need from the sample.

8. What are the weaknesses inherent in this method of obtaining the data?

9. What safeguards have you established to counteract any potential bias in your approach to data collection? Be specific.

Interpreting Data in Descriptive Research

Data are of little or no value merely as data. In our discussion of descriptive research methods, we have primarily discussed the acquisition of data: how to obtain the data from the general population with appropriate techniques (observations, interviews, questionnaires, sampling) and how to protect those data against distortion of bias. We have been thinking of the process of data collection only.

At this juncture, we remind you of two basic principles of research:

1. The purpose of research is to seek the answer to a problem in the light of the data that relate to the problem.
2. Although collecting data for study and organizing it for inspection require care and precision, extracting meaning from the data—the interpretation of the data—is all-important.

A descriptive study is often a very “busy” research method: The researcher must decide on a population; choose a technique for sampling it; develop a valid means of collecting the desired information; minimize the potential entrance of bias into the study; and then actually collect, record, organize, and analyze all of the necessary data. The activities connected with descriptive research are complex, time-consuming, and occasionally distracting. Therein lies an element of danger. With all this action going on, it would not be surprising if the researcher lost sight of the problem and subproblems. But the problem and its subproblems are precisely the reason for the entire endeavor.

All research activity is subordinate to the research problem itself. Sooner or later, the entire effort must result in an interpretation of the data and a setting forth of conclusions, drawn from the data, to resolve the problem under investigation. Inexperienced researchers sometimes forget this fact. Activity for activity’s sake is seductive. Amassing great quantities of data can provide a sense of well-being. Like Midas looking at his hoard of gold, researchers might lose sight of the ultimate demands that the problem itself makes on those data. Presenting the data in displays and summaries—graphs, charts, tables—does nothing more than demonstrate the researcher’s acquisitive skills and consummate ability to present the same data in various ways. Descriptive research ultimately aims to solve problems through the *interpretation* of the data that have been gathered.

Some Final Suggestions

As we approach the end of the chapter, it is important to reflect on several issues related to descriptive research. Consider each of the following questions within the context of the research project you have in mind:

- Why is a description of this population and/or phenomenon valuable?
- What specific data will I need to solve my research problem and its subproblems?

- What procedures should I follow to obtain the necessary information? How should I implement those procedures?
- How do I get a sample that will truly be reflective of the entire population about which I am concerned?
- How can I collect my data in a way that minimizes misrepresentations and misunderstandings?
- How can I control for possible bias in the collection and description of the data?
- What do I do with the data once I have collected them? How do I organize and prepare them for analysis?
- Above all, in what ways might I reasonably interpret the data? What conclusions might I reach from my investigation?

A Sample Dissertation

We conclude the chapter by illustrating how questionnaires might be used in a correlational study to address the topic of violence in intimate relationships (e.g., husband and wife, boyfriend and girlfriend) in American society. The excerpts we present are from Luis Ramirez's doctoral dissertation in sociology completed at the University of New Hampshire (Ramirez, 2001).

Ramirez hypothesizes that violence between intimate partners—in particular, assault by one partner on the other—is, in part, a function of ethnicity, acculturation (e.g., adoption of mainstream American behaviors and values), criminal history, and social integration (e.g., feelings of connectedness with family and friends). He further hypothesizes that as a result of such factors, differences in intimate partner violence might be observed in Mexican Americans and non-Mexican Americans.

Ramirez begins Chapter 1 by discussing the prevalence of violence (especially assault) in intimate relationships. We pick up Chapter 1 at the point where he identifies his research questions and hypotheses. We then move into Chapter 2, where he describes his methodology. As has been true for earlier proposal and dissertation samples, the research report appears on the left-hand side, and our commentary appears on the right.

Dissertation ANALYSIS

5

RESEARCH QUESTIONS

[T]he following questions will be addressed: What role does acculturation into American society have on intimate partner violence for Mexican Americans? What are the effects of a person's criminal history on intimate partner violence? What is the extent and the relation of criminal history to intimate partner violence, and is criminal history restricted to one type of crime or is it a more general tendency (violent versus property crimes)? Are crimes that are committed early in life more indicative of a pattern of crime as compared to crimes that begin later in life? Do people who assault their partners possess weak social bonds with the society they live in? Finally, this study will ask the question, "Are there differences between criminal history and bond to society for Mexican Americans and Non-Mexican Whites, and how do these factors affect intimate partner violence?"

Comments

To understand factors underlying violence in intimate partner relationships—his main research problem—the author identifies a number of subproblems, which he expresses here as research questions.

If relations are found between these characteristics, it suggests that social agencies that deal with intimate partner violence need to adjust their policies and intervention procedures to better meet the characteristics of their clients. The focus of primary prevention could be put on the social bonding process, the criminal history of the individual, or the acculturation process in order to help solve future problems. Furthermore, a comparative study of intimate partner assault among ethnic groups could provide further clarification to a body of literature and research that has produced mixed results.

[The author briefly reviews theoretical frameworks related to ethnicity and acculturation, criminal history, and control theory, which he then uses as a basis for his hypotheses.]

HYPOTHESES

The theoretical frameworks reviewed led to the following hypotheses:

Ethnicity and Acculturation

1. The rate of intimate partner violence is lower for Mexican Americans than Non-Mexicans.
2. The higher the acculturation into American Society, the higher the probability of assaulting a partner for Mexican Americans.

Criminal History

3. Criminal history is more prevalent for Mexican Americans than for Non-Mexicans.
4. The more crimes committed in the past, the higher the probability of physically assaulting a partner.
5. Criminal history is more associated with an increased risk of intimate partner violence for Mexican Americans than Non-Mexicans.
6. Early onset crime is more associated with an increased risk of intimate partner violence than criminal behavior beginning later in life.
7. Previous violent crime is more associated with an increased risk of intimate partner violence than property crime.

Social Integration

8. Mexican Americans are more socially integrated than Non-Mexican Whites.
9. The more socially integrated an individual is, the lower the probability of physically assaulting a partner.
10. Social integration is more associated with a decreased risk of intimate partner violence for Mexican Americans than Non-Mexicans.

A more detailed review of the literature will be presented in . . . following chapters. Literature for all hypotheses will be reviewed in their respective chapters.

Here the author addresses the importance of the study, both pragmatic (results have potential implications for social policy and practice) and theoretical (results may shed light on inconsistencies in previous research studies).

The hypotheses are organized by the theoretical frameworks from which they have been derived, helping the reader connect them to rationales the author has previously provided.

Notice how the hypotheses are single-spaced. Single-spaced hypotheses often appear in theses and dissertations, but check the guidelines at your own institution to see whether such formatting is desired.

An in-depth review of the literature is postponed until Chapters 3 through 5, where the author also relates his own results to previous research findings. Although this is an unusual organizational structure, it works well in this situation, allowing the reader to connect results relative to each hypothesis to the appropriate body of literature.

Figure 1.1 is a diagrammed representation of what I believe is the causal process that could affect intimate partner violence. It includes demographic and control variables, the main independent variables (acculturation, criminal history, social integration), and intimate partner violence. These variables will be described in detail in the next chapter.

Note the transition to the next chapter, which immediately follows.

FIGURE 1.1 Model of Intimate Partner Violence

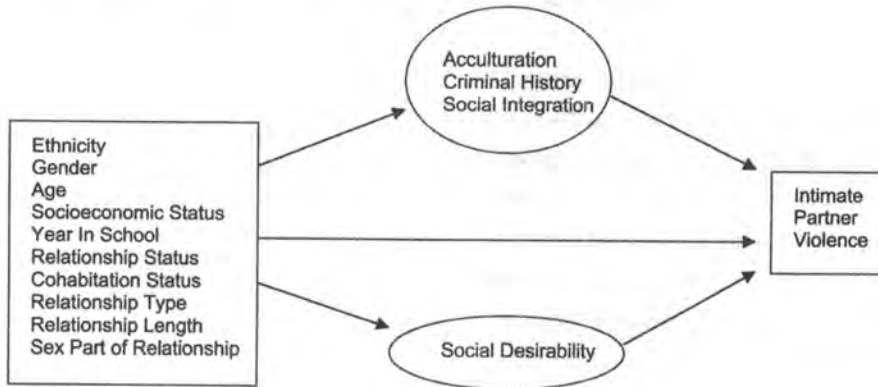


Figure 1.1 effectively condenses and summarizes the researcher's hypotheses. Also, it graphically demonstrates that four variables—acculturation, criminal history, social integration, and social desirability—are hypothesized to be mediating variables in the relationship between demographics and violence.

CHAPTER 2

Methods

Sample

The issues discussed in the previous chapter will be investigated using data from a sample of college students who have been or are currently in a dating or married relationship. A sample of college students is appropriate for this study for the following reasons: (1) The National Crime Victimization Survey found that the rates of non-lethal intimate partner violence was greatest for the 20 to 24 year age group, followed by the 16 to 19 age group, and then the 25 to 34 age group (Renison & Welchans, 2000). The majority of college students fall into the high-risk age categories. Sugarman and Hotaling (1989) identified eleven studies that provided rates for physical assault of dating partners and concluded the rates of assaulting a partner range from 20% to 59%. (2) College students make up about a third of the 18 to 22 year old population. College students are a sizable population in reference to the general population (about 15 million). (3) College students are in a formative period of their lives in relation to the habits that they develop with an intimate partner. These habits could surface in other intimate relations (O'Leary, Malone, & Tyree, 1994; Pan, Neidig, & O'Leary, 1994).

Some style manuals suggest that an author include at least a small amount of text between two headings of different levels. For example, before beginning the "Sample" section, the author might provide an advance organizer, describing the topics he will discuss in the chapter and in what order.

It is important to mention that a sample of college students is not a representative sample of the general population in the United States. This group generally has lower levels of criminal behavior, substance abuse, and marriage rates. Additionally, college students may be more socially integrated into society and are engaged in education as a tool for upward mobility. In short, this is a segment of society that plays by the rules.

Data Collection

Six hundred and fifty questionnaires were passed out to students at The University of Texas at El Paso and Texas Tech University during the fall 1999, spring 2000, and summer 2000 semesters. Students who were enrolled in Sociology, Anthropology, and History classes [were] the respondents.

Respondents filled out the questionnaire (Appendix A) in a classroom setting. Each respondent received a booklet consisting of: (1) a cover sheet explaining the purpose of the study, the participant's rights, and the name of a contact person and telephone number for those who might have questions after the test session was over; (2) the demographic questions; (3) the instruments described in this section. The purpose, task demands, and rights were explained orally as well as in printed form at the beginning of each session. Respondents were told that the questionnaire would include questions concerning attitudes, beliefs, and experiences they may have had. They were guaranteed anonymity and confidentiality of their responses and they were told that the session would take an hour or slightly more. In actuality, the range of time that it took students to finish was between 30 minutes to 1 hour. All students were asked to sign a written consent form before completing their questionnaires. Students were also given instructions on how to properly fill out three scantron sheets before they were left to fill out the questionnaire at their own pace.

A debriefing form was given to each participant as they turned in their questionnaire. It explained the study in more detail and provided names and telephone numbers of local mental health services and community resources, such as services for battered women. Students that voluntarily participated in the study were offered extra credit points by their professors.

The initial sample consisted of 650 respondents of which 576 chose to complete the questionnaire. Of these, 33 questionnaires were omitted because they were illegible or partially completed. Finally, of the 543 remaining questionnaires, 348 were selected for this study because they met the criteria of having no missing data for any specific question, were either Mexican American/Mexican National or Non-Mexican White, and had been in a heterosexual romantic relationship for a month or longer during the previous 12 months.

The author clearly realizes that his sample (college students) is not representative of the entire U.S. population. He presents a good case that the sample is quite appropriate for his research questions. At the same time, he acknowledges that his sample has some shortcomings.

The author, whose home town is El Paso, has numerous acquaintances at both institutions and so can easily gain access to these students. He must, of course, seek approval from the internal review boards at the two institutions, as well as at the institution where he is completing his doctorate.

The author has combined his informed consent forms and questionnaires into a single booklet that he can easily distribute. Doing so is quite common in descriptive research, especially with adult samples, and increases the efficiency of data collection.

The author is using computer technology (scantron sheets) in his data collection. Given the nature of his sample (college students) and his sample size (576), this approach is reasonable.

Given the sensitive nature of some questionnaire items, the debriefing that follows data collection appropriately includes information about community resources for individuals who have been victims of partner violence.

Here the author describes his criteria for including completed questionnaires in his data set. In essence, he is addressing the issue of admissibility of the data (see Chapter 4).

NOTE: Excerpt is from *The Relation of Acculturation, Criminal History, and Social Integration of Mexican American and Non-Mexican Students to Assaults on Intimate Partners* (pp. 3–4, 14–20) by I. L. Ramirez, 2001, unpublished doctoral dissertation, University of New Hampshire, Durham. Reprinted with permission.

For Further Reading

- Alreck, P. L., & Settle, R. B. (2003). *The survey research handbook* (3rd ed.). New York: McGraw-Hill.
- Berdie, D. R., Anderson, J. F., & Niebuhr, M. A. (1986). *Questionnaires: Design and use* (2nd ed.). Metuchen, NJ: Scarecrow Press.
- Bourque, L. B., & Fielder, E. P. (2002). *How to conduct telephone surveys* (2nd ed.). Thousand Oaks, CA: Sage.
- Bourque, L. B., & Fielder, E. P. (Eds.). (2002). *How to conduct self-administered and mail surveys* (2nd ed.). Thousand Oaks, CA: Sage.
- Carpenter, J. (2002). Measuring social capital: Adding field experimental methods to the analytical toolbox. In J. Isham, T. Kelly, & S. Ramaswamy (Eds.), *Social capital and economic development: Well-being in developing countries* (pp. 119–135). Cheltenham, England: Edward Elgar Publishing.
- Daniel, J. (2011). *Sampling essentials: Practical guidelines for making sampling choices*. Thousand Oaks, CA: Sage.
- Delandshere, G., & Petrosky, A. R. (1998). Assessment of complex performances: Limitations of key measurement assumptions. *Educational Researcher*, 27, 14–24.
- Fink, A. (2009). *How to conduct surveys: A step-by-step guide* (4th ed.). Thousand Oaks, CA: Sage.
- Fowler, F. J., Jr. (2008). *Survey research methods* (4th ed.). Thousand Oaks, CA: Sage.
- Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust Web-based studies? A comparative analysis of six preconceptions about Internet questionnaires. *American Psychologist*, 59, 93–104.
- Gubrium, J. F., & Holstein, J. A. (Eds.). (2001). *Handbook of interview research: Context and method*. Thousand Oaks, CA: Sage.
- Henry, G. T. (1990). *Practical sampling*. Thousand Oaks, CA: Sage.
- Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., & Couper, M. (2004). Psychological research online: Report of Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet. *American Psychologist*, 59, 105–117.
- Litwin, M. S. (1995). *How to measure survey reliability and validity*. Thousand Oaks, CA: Sage.
- Magnusson, D., Bergman, L. R., Rudinger, G., & Torestad, B. (Eds.). (1991). *Problems and methods in longitudinal research: Stability and change*. New York: Cambridge University Press.
- Menard, S. (Ed.). (2007). *Handbook of longitudinal research: Design, measurement, and analysis*. Burlington, MA: Academic Press.
- Neuman, W. L. (2011). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Boston: Allyn & Bacon. (See Chapter 10, "Survey Research.")
- Oppenheim, A. N. (1992). *Questionnaire design, interviewing, and attitude measurement*. New York: St. Martin's Press.
- Panther, A. T., & Sterba, S. K. (Eds.). *Handbook of ethics in quantitative methodology*. New York: Routledge.
- Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, 38, 109–119.
- Rogelberg, S. G., & Luong, A. (1998). Nonresponse to mailed surveys: A review and guide. *Current Directions in Psychological Science*, 7, 60–65.
- Rosenfeld, P. (1993). Computer-administered surveys in organizational settings: Alternatives, advantages, and applications. *American Behavioral Scientist*, 36, 485–511.
- Schuman, H., & Presser, S. (1982). *Questions and answers: Experiments in the form, wording, and context of survey questions*. San Diego, CA: Academic Press.
- Schwarz, P. N., & Sudman, S. (1996). *Answering questions: Methodology for determining cognitive and communicative processes in survey research*. San Francisco: Jossey-Bass.
- Wright, J., & Marsden, P. (2010). *Handbook of survey research* (2nd ed.). Bingley, England: Emerald Group Publishing.

Now go to MyEducationLab at www.myeducationlab.com to take assessments to evaluate your mastery of chapter concepts. Assignments and Activities exercises and Building Skills in Reading Research exercises are also available to help you master concepts and skills. Feedback for many of these exercises is provided so that you can see why your answers are correct or incorrect.

Experimental, Quasi-Experimental, and Ex Post Facto Designs

Progress is relative: We measure it by noting the degree of change between what was and what is. And we attempt to account for the change by identifying the dynamics that have caused it. Ideally, we must manipulate one possible causal factor while controlling all other possible causal factors; only in this way can we determine whether the manipulated factor has a causal effect on the phenomenon we are studying. To the extent that many potentially causal factors all vary at once, in a confounded manner, we learn little or nothing about what causes what.

In the designs described in preceding chapters, we have made no systematic attempt to determine the causes of the phenomena being studied. But ultimately we often do want to know what leads to what; in other words, we want to identify *cause-and-effect relationships*.

A researcher can most convincingly identify cause-and-effect relationships by using an **experimental design**. In such a design, the researcher considers many possible factors that might cause or influence a particular condition or phenomenon. The researcher then attempts to control for all influential factors *except* those whose possible effects are the focus of investigation.

An example can help clarify the point. Imagine that we have two groups of people. We take steps to make sure that, on average, the two groups are so similar that we can, for all intents and purposes, call them equivalent. We give members of both groups a pretest to measure a particular characteristic in which we are interested—for instance, this might be blood pressure, academic achievement, or purchasing habits. Then we expose only one of the groups to a **treatment** or intervention of some sort—perhaps a new pharmaceutical drug, an instructional method, or an advertising campaign—that we think may have an effect on the characteristic we are studying. Afterward, we give members of both groups a posttest to measure the characteristic once again. If the characteristic changes for the group that received the intervention but does *not* change for the other group, and if everything about the two groups has been the same *except for the intervention*, we can reasonably conclude that the treatment or intervention brought about the change we observed. Because we have not only observed but also *manipulated* the situation, we have used an experimental design.

Some of the research designs we describe in this chapter are true experimental designs; as such, they allow us to identify cause-and-effect relationships. Other designs in this chapter eliminate some—but not all—alternative explanations of an observed change. Yet all of the designs in the chapter have one thing in common: clearly identifiable independent and dependent variables.

We introduced you to independent and dependent variables in Chapter 2, but because these concepts will guide so much of our discussion in this chapter, a brief refresher might be in order here. An **independent variable** is one that the researcher studies as a possible cause of something else. In many of the designs described in this chapter, the researcher directly manipulates the independent variable. In contrast, a **dependent variable** is a variable that is potentially influenced by the independent variable—it is the “something else” the independent variable

possibly influences, and thus it *depends* on the independent variable. In other words, the hypothesized relationship is this:

Independent variable → Dependent variable

As an example, let's look at a dissertation in educational psychology (Thraillkill, 1996). The researcher wanted to study the effects of three different kinds of lecture material on people's ability to remember information contained in the lecture. Working with undergraduate students, she presented different parts of a lecture on an obscure American Civil War battle in one of three ways: (1) she described certain historical figures and events in such a manner that they were easy to imagine and visualize (*imagery* condition), (2) she included attention-grabbing phrases in the lecture (*attention* condition), or (c) she did neither of these things (*control* condition). In the following examples, the underscored phrases illustrate the modifications made for each of the three conditions; other variations in wording made the three lectures equivalent in length:

Imagery: Lincoln also created the Army of Virginia, incorporating several forces which had been under different commanders. Lincoln set the dimpled, baby-faced young blond Major General John Pope in charge of this new combined force. Being put under his command was objectionable to some of the former commanders. . . .

Attention: Lincoln also created the Army of Virginia, incorporating several forces which had been under different commanders. LISTEN TO ME NOW. Lincoln set the less experienced Major General John Pope in charge of this new combined force. Being put under the command of Pope was objectionable to some of the former commanders. . . .

Control: Lincoln also created the Army of Virginia, incorporating several forces which had been under different commanders. Lincoln the less experienced junior officer Major General John Pope in charge of this new combined force. Being put under the command of Pope was objectionable to some of the former commanders. (Thraillkill, 1996, p. 62, some underscoring added)

After presenting different parts of the lecture under the three different conditions, the researcher measured the students' recall for the lecture in two ways. She first gave students blank sheets of paper and asked them to write down as much of the lecture as they could remember (a "free recall" task). When they had completed the task, she gave them a multiple-choice test that assessed their memory for specific facts within the lecture. In this study, the independent variable was the nature of the lecture material: easily visualized, attention-getting, or neutral. There were two dependent variables, both of which reflected students' ability to recall facts within the lecture: students' performance on the free recall task and their scores on the multiple-choice test. Thraillkill's hypothesis was confirmed: The students' ability to recall lecture content *depended*, to some extent, on the way in which the content was presented.

The Importance of Control

A particular concern in any experimental study is its **internal validity**, the extent to which its design and the data it yields allow the researcher to draw legitimate conclusions about cause-and-effect and other relationships (see Chapter 4). In experimental designs, internal validity is essential. Without it, a researcher cannot draw firm conclusions about cause and effect—and that is, after all, the whole point of conducting an experimental study.

As an example, suppose we have just learned about a new method of teaching science in elementary school. We want to conduct an experiment to investigate the method's effect on students' science achievement test scores. We find two fifth-grade teachers who are willing to participate in the study. One teacher agrees to use the new method in the coming school year; in fact, she is quite eager to try it. The other teacher wants to continue using the same approach he has always used. Both teachers agree that at the end of the school year we can give their students a science achievement test.

Are the two classes the same in every respect *except for the experimental intervention*? If the students taught with the new method obtain higher science achievement test scores at the end of the year, will we know that the method was the *cause* of the higher scores? The answer to both questions is a resounding *no*! The teachers are different: One is female and the other male, and they almost certainly have different personalities, educational backgrounds, teaching styles, and so on. In addition, the two groups of students may be different; perhaps the students instructed by the new method are, on average, more intelligent or motivated than the other, or perhaps they live in a more affluent school district. Other, more subtle differences may be at work as well, including the interpersonal dynamics in the two classes, and the light, temperature, and noise levels within each classroom. Any of these factors—and perhaps others we haven't thought of—might be reasons for any group differences in achievement test scores we obtain.

Whenever we compare two or more groups that are or might be different in ways *in addition to* the particular treatment or intervention we are studying, we have **confounding variables** in our study. The presence of such variables makes it extremely difficult to draw conclusions about cause-and-effect relationships, because we cannot pin down *what* is the cause of any pattern in the data observed after the intervention. In other words, confounding variables threaten a study's internal validity. In a classic book chapter, Campbell and Stanley (1963) identified several potential threats to the internal validity of an experimental study; we describe them in Figure 9.1.

Controlling for Confounding Variables

To maximize internal validity when a researcher wants to identify cause-and-effect relationships, the researcher needs to control confounding variables in order to rule them out as explanations for any effects observed. Researchers use a variety of strategies to control for confounding variables. Following are several common ones:

1. *Keep some things constant.* When a factor is the *same* for everyone, it cannot possibly account for any differences observed. Oftentimes researchers ensure that different treatments are imposed in the same or very similar environments. They may also seek research participants who share a certain characteristic, such as sex, age, grade level, or socioeconomic status. Keep in mind, however, that restricting the nature of one's sample may lower the *external validity*, or generalizability, of any findings obtained (see Chapter 4's discussion of this concept).

2. *Include a control group.* In Chapter 4 we described a study in which an industrial psychologist begins playing classical music as employees in a typing pool go about their daily task of typing documents. At the end of the month, the psychologist finds that the typists' productivity is 30% higher than it was during the preceding month. The increase in productivity may or may not be due to the classical music. There are too many possible confounding variables—personnel changes, nature of the documents being typed, numbers of people out sick or on vacation during the two-month period, even just the knowledge that an experiment is being conducted—that may also account for the typists' increased productivity.

To better control for such extraneous variables, researchers frequently include a **control group**, a group that receives either no intervention or a "neutral" intervention that should have little or no effect on the dependent variable. The researchers then compare the performance of this group to an **experimental group**—also known as a **treatment group**—that participates in an intervention.

As you should recall from Chapter 4, people sometimes show improved performance simply because they know they are participating in a research study—a phenomenon known as *reactivity* and, more specifically, the *Hawthorne effect*. To take this fact into account, a researcher sometimes gives the people in a control group a **placebo** that has the appearance of having an effect but in reality *shouldn't* have an effect. For instance, a researcher studying the effects of a new arthritis medication might give some participants a particular dosage of the medicine and give others a

FIGURE 9.1

Potential threats to the internal validity in an experimental study

When a researcher studies the possible effects of an intervention on some other (dependent) variable, a number of confounding variables can come into play that threaten the study's internal validity and thereby also jeopardize any cause-and-effect conclusions the researcher might hope to draw. Campbell and Stanley (1963) have identified the following potential threats to internal validity, which can be present either singly or in combination:

1. **History:** An uncontrolled outside event occurring between two measurements of the dependent variable brings about a change in the dependent variable. For example, a noteworthy event in the local community might change participants' knowledge, abilities, or emotional states in ways that affect the second measurement of the dependent variable.
2. **Maturation:** A change in participants' characteristics or abilities might simply be the result of the passage of time. For example, children might make normal developmental gains in eye-hand coordination or intellectual ability.
3. **Testing:** Taking a test at one time influences participants' performance during a subsequent administration of the test, perhaps simply as a result of practice in taking the test. For example, people who take a multiple-choice test at one time may gain general test-taking skills that enhance their performance on a subsequent multiple-choice test.
4. **Instrumentation:** A change occurs in how a measurement instrument is used from one time to the next. For example, a researcher might have one research assistant rate participants' performance on the first occasion but have a different research assistant judge their performance on the subsequent occasion. Any observed change might be the result of the two assistants' differing standards for rating the performance. (This threat to internal validity reflects a problem with interrater reliability; see Chapter 4).
5. **Statistical regression:** People who score extremely high or low on a measure at one time are likely to score in a less extreme manner on the same measure at a later time; that is, extreme scorers tend to "drift" toward more average performance during a subsequent measure. For example, a researcher might assign people to one of two groups—"high-anxiety" or "low-anxiety"—based on their extremely high or low scores on a self-report questionnaire designed to measure general anxiety level. Especially if the initially extreme scores were the result of people's temporary circumstances—circumstances that might make them feel either exceptionally anxious or, instead, quite "mellow" on the first testing—the supposedly high-anxiety people would become less anxious and the supposedly low-anxiety people would become more anxious regardless of any experimental interventions the two groups might undergo.
6. **Selection:** A bias exists in how members of different groups in a study are chosen. For example, when recruiting college students for a study, a researcher might put all students enrolled in an 8:00 a.m. class in one treatment group and all students enrolled in a 2:00 p.m. class in another treatment group. Students taking the early-morning class might be different in some significant way from those taking the afternoon class (e.g., the sleeping habits of the two groups might be different).
7. **Attrition:*** Members of different groups drop out of the study at proportionally different rates. For example, one group in a study might lose 25% of its members before the final measurement, whereas another group might lose only 5% of its members. Thus, even if the two groups were equivalent with regard to important characteristics at the beginning of the study, they might be different in some significant way later in the study simply as a result of the differential dropout rate.

Campbell and Stanley listed an eighth threat to internal validity as well: an *interaction* among two of the threats listed above. For example, if students in an 8:00 a.m. class are assigned to one treatment group and students in a 2:00 p.m. class are assigned to a different treatment group, and if students in the 8:00 a.m. group drop out of the study in greater numbers than students in the 2:00 p.m. group, any final differences observed in the dependent variable might be the result of the fact that early risers are, for some reason, more likely to drop out than students who like to sleep in a bit. In this situation, it becomes virtually impossible to disentangle possible effects of an experimental intervention from effects of (a) the selection bias, (b) the differing dropout rates, and (c) the interaction of these two confounding variables.

*Note: Campbell and Stanley use the term *experimental mortality* for this threat to internal validity, but the term *attrition* is more commonly seen in contemporary research literature.

similar-looking sugar pill. Or a researcher investigating a new approach to treating test anxiety might use the new treatment with some individuals but give other individuals general relaxation training that, although possibly beneficial in other ways, won't necessarily address their test anxiety.

We must emphasize—and we emphasize it quite strongly—that any researcher who incorporates placebos in a study must consider *three ethical issues* related to the use of placebos. First is the principle of informed consent: Participants in the study must be told that the study includes a placebo treatment as well as an experimental treatment and that they won't know which treatment they have received until the study has ended. Second, if participants in the study have actively sought help for a medical, psychological, or other significant problem, those who initially receive the placebo treatment should, at the conclusion of the study, be given the opportunity to receive more effective treatment. (This is assuming, of course, that the treatment *is* more effective than the placebo.) Third, and most important, when studying a treatment related to life-threatening situations (e.g., a new drug for terminal cancer, a new psychotherapeutic technique for suicidal teenagers), the researcher must seriously weigh (a) the benefits of the new knowledge that can be gained by a control group receiving no treatment against (b) the lives that may be saved by including all participants in the treatment group.

Our last point raises an issue we cannot possibly resolve for you here. Should you find yourself having to make a decision about the best research design to use in a life-and-death situation, you should consult with your professional colleagues, the internal review board at your institution, and, of course, your own conscience.

3. Randomly assign people to groups. In Chapter 8 we spoke at length of the value of selecting people at random to participate in a descriptive research study; such random selection enhances the probability that any results obtained for the sample also apply to the population from which the sample has been drawn. In experimental studies, researchers use random selection for a different purpose: to assign participants within their sample to various groups.

In any research study involving human beings or other living things, members of the sample are likely to be different from one another in many ways that are relevant to the variables under investigation. For example, earlier in the chapter we described a situation in which a researcher wants to compare two methods of teaching elementary school science. The students in the study will almost certainly differ from one another in intelligence, motivation, educational opportunities at home, and other factors that will affect their performance on the achievement test given at the end of the school year. It would be virtually impossible to control for such variables by having all students in the study have the *same* intelligence, the *same* motivation, the *same* kinds of outside opportunities, and so on.

As an alternative to keeping some characteristics the same for everyone, a researcher can, instead, randomly assign participants to groups. When people have been selected for one group or another on a random basis, the researcher can reasonably assume that, *on average, the groups are quite similar* and that *any differences between them are due entirely to chance*. In fact, many inferential statistical tests—especially those that allow the researcher to make comparisons among two or more groups—are based on the assumption that group membership is randomly determined and that any pretreatment differences between the groups result from chance alone.

4. Assess equivalence before the treatment with one or more pretests. Sometimes random assignment to two different groups simply isn't possible; for instance, researchers may have to study groups that already exist (e.g., students in classrooms, participants in different medical treatment programs). An alternative in this situation is to assess other variables that might influence the dependent variable and determine whether the groups are similar with respect to those variables. If the groups *are* similar, the probability that such variables could account for any future group differences is reduced considerably.

Another strategy is to identify **matched pairs**: pairs of people—one in each of two groups being compared—who are identical or very similar with respect to characteristics that may potentially have an effect on the dependent variable. For instance, a researcher comparing the achievement test scores of students in two different instructional programs might identify pairs

of students of the same sex and age who have similar IQ scores. A researcher comparing two different treatments for a particular illness might match patients according to sex, age, and duration and intensity of the illness. In either case, the researcher does not study the data collected for *all* people in the two groups, only the people who are part of “matched sets” that he or she has identified. A researcher who uses this approach will, in the final research report, explain in what way(s) the participants in the study have been matched. For example, he or she might say, “Pairs were matched on the basis of sex, age, and IQ.”

One problem with assessing before-treatment equivalence with pretests is that the researcher rules out *only the variables that he or she has actually assessed and determined to be equivalent across groups*. The design does not rule out other influential factors that the researcher has not assessed and perhaps not even considered.

5. *Expose participants to all experimental conditions.* Still another strategy for controlling for individual differences is to *use participants as their own controls*—that is, to have every participant in the study undergo all experimental and control treatments and then assess the effects of each treatment independently. Any independent variable that is systematically varied for every participant is known as a **within-subjects variable**, and an approach that includes a within-subjects variable is known as a **within-subjects design**. You may also see the term **repeated-measures design** used in reference to this approach.

As an example, let's return to the dissertation involving three different lecture methods and their possible effects on recall for lecture content (Thrailkill, 1996). The researcher's sample consisted of volunteer students who were enrolled in three sections of an undergraduate class in educational psychology, and she planned to give the lecture just three times, once to each class. The lecture was about an American Civil War battle sufficiently obscure that participants were unlikely to have had any prior knowledge about it; thus, participants' prior knowledge about the battle was a constant—they all had *zero* prior knowledge—rather than a confounding variable. The researcher divided the lecture into three parts of approximately equal length and wrote three versions of each part, one version each for the imagery, attention, and control conditions. She combined the three versions of the three lecture parts such that each class received the different treatments in a different sequence, as follows:

	PART OF LECTURE		
	<i>First Part</i>	<i>Middle Part</i>	<i>Last Part</i>
Group 1	Attention	Imagery	Control
Group 2	Control	Attention	Imagery
Group 3	Imagery	Control	Attention

In this manner, all participants in her study were exposed to the two treatments and the control condition, and each condition occurred in all possible places (first, second, and third) in the sequence.

In the study just described, the researcher used a within-subjects variable (type of intervention: imagery vs. attention vs. control) to compensate for the fact that participants had not been randomly assigned to the three class sections in her sample. Sometimes researchers use a similar strategy with just a single group, and in some cases with just a single individual. You will learn some strategies for showing causation in single-group and single-individual studies later in the chapter, when we explore *quasi-experimental designs*.

6. *Statistically control for confounding variables.* Sometimes researchers can control for known confounding variables, at least in part, through statistical techniques. Such techniques as *partial correlation*, *analysis of covariance* (ANCOVA), and *structural equation modeling* are suitable for

this purpose. We briefly describe each of these in Chapter 11. Should you choose to use one of them in your own research, we urge you to consult one or more statistics books for guidance about their use and appropriateness for various research situations.

Keep in mind, however, that statistically controlling confounding variables is no substitute for controlling for them in one's research design if at all possible. *A carefully controlled experimental design is the only approach that allows you to draw firm conclusions about cause-and-effect relationships.*

Overview of Experimental, Quasi-Experimental, and Ex Post Facto Designs

In true experimental research, the researcher manipulates the independent variable and examines its effects on another, dependent variable. A variety of research designs have emerged that differ in the extent to which the researcher manipulates the independent variable and controls for confounding variables—in other words, the designs differ in the degree to which they have *internal validity*. In the upcoming sections, we present a number of possible designs, which we have divided into five general categories: *pre-experimental designs*, *true experimental designs*, *quasi-experimental designs*, *ex post facto designs*, and *factorial designs*. Altogether we describe 16 different designs that illustrate various ways—some more effective than others—of attempting to identify cause-and-effect relationships. Some of our discussion is based on designs identified by Campbell and Stanley (1963).¹

We illustrate the designs using tables that have this general format:

Group	Time →		
Group 1			
Group 2			

Each group in a design is shown in a separate row, and the things that happen to the group over time are shown in separate cells within the row. The cells have one of four notations:

Tx: Indicates that a *treatment* (reflecting the independent variable) is presented.

Obs: Indicates that an *observation* (reflecting the dependent variable) is made.

—: Indicates that nothing occurs during a particular time period.

Exp: Indicates a previous *experience* (an independent variable) that some participants have had and others have not; the experience has *not* been one that the researcher could control.

The nature of these tables will become more apparent as we proceed.

As you read about the 16 designs, keep in mind that they are hardly an exhaustive list; researchers can modify or combine them in various ways. For example, although we will be limiting ourselves to studies with only one or two groups (perhaps one treatment group and one control group), it is entirely possible to have two or more treatment groups (each of which is exposed to a different variation of the independent variable) and, in some cases, two control groups (perhaps one getting a placebo and another getting no intervention at all). More generally, the designs we describe here should simply provide a starting point that gets you thinking about how you might best tackle your own research problem.

¹In particular, Designs 1–6 and Designs 8–11 are based on those that Campbell and Stanley described. However, when describing Design 11, we use the contemporary term *reversal time-series design* rather than Campbell and Stanley's original term *equivalent time-samples design*.

Pre-Experimental Designs

In pre-experimental designs, it isn't possible to show cause-and-effect relationships, because either (a) the independent "variable" doesn't vary or (b) experimental and control groups are not comprised of equivalent or randomly selected individuals. Such designs are helpful only for forming tentative hypotheses that should be followed up with more controlled studies.

Design 1: One-Shot Experimental Case Study

The one-shot experimental case study is probably the most primitive type of experiment that might conceivably be termed "research." An experimental treatment (Tx) is introduced, and then a measurement (Obs)—a posttest of some sort—is administered to determine the effects of the treatment. This design is shown in the following table:

Group	Time →	
Group 1	Tx	Obs

The design has low internal validity because it is impossible to determine whether participants' performance on the posttest is the result of the experimental treatment per se. Many other variables may have influenced participants' performance, such as physiological maturation or experiences elsewhere in the participants' general environment. Perhaps the characteristic or behavior observed after the treatment existed *before* the treatment as well. The reality is that with a single measurement or observation, we have no way of knowing whether the situation has changed or not, let alone whether it has changed as a result of the intervention.

One-shot experimental case studies may be at the root of many common misconceptions. For example, imagine that we see a boy sitting on the damp ground in mid-April. The next day he has a sore throat and a cold. We conclude that sitting on the damp earth caused him to catch cold. Thus, the design of our "research" thinking is something like this:

Exposure to cold, damp ground (Tx) → Child has a cold (Obs)

Such "research" may also "support" such superstitious folk beliefs as these: If you walk under a ladder, you will have bad luck; Friday the 13th is a day of catastrophes; a horseshoe above the door brings good fortune to the house. Someone observed an event, then observed a subsequent event, and linked the two together as cause and effect.

Be careful not to confuse the one-shot experimental case study method with the case study design of many qualitative studies. As described in Chapter 6, case study research involves extensive engagement in a research setting—a far cry from basing conclusions on a single observation.

Although the one-shot experimental case study is simple to carry out, its results are, for all intents and purposes, meaningless. At the very least, researchers should use the design described next.

Design 2: One-Group Pretest–Posttest Design

In a one-group pretest–posttest design, a single group (a) undergoes a pre-experimental observation or evaluation, then (b) is administered the experimental treatment, and finally (c) is observed or evaluated again after the treatment. This design can be represented as follows:

Group	Time →		
Group 1	Obs	Tx	Obs

Suppose an elementary school teacher wants to know if simultaneously reading a story and listening to it on audiotape will improve the reading skills of students in his class. He gives his

students a standardized reading test, then has them simultaneously read and listen to simple stories every day for eight weeks, and then administers an alternate form of the same standardized reading test. If the students' test scores improve over the eight-week period, the teacher might conclude—perhaps accurately, but perhaps not—that the simultaneous-reading-and-listening intervention was the cause of the improvement.

Now suppose an agronomist hybridizes two strains of corn. She finds that the hybrid strain is more disease-resistant and has a better yield than either of the two parent types. She concludes that the hybridization process has made the difference. Once again we have an Obs–Tx–Obs design: The agronomist measures the disease level of the parent strains (Obs), then develops a hybrid of the two strains (Tx), and then measures the disease level of the next generation (Obs).

In a one-group pretest–posttest design, we at least know that a change has taken place. However, we have not ruled out other possible explanations for the change. In the case of the elementary school teacher's study, improvement in reading scores may have been due to other activities within the classroom curriculum, to more practice taking the reading test, or simply to the fact that the students were eight weeks older. In the case of the agronomist's experiment, changes in rainfall, temperature, or soil conditions may have been the primary reason for the healthier corn crop.

Design 3: Static Group Comparison

The static group comparison involves both an experimental group and a control group. Its design takes the following form:

Group	Time →	
Group 1	Tx	Obs
Group 2	—	Obs

An experimental group is exposed to a particular experimental treatment; the control group is not. After the treatment, both groups are observed and their performance compared. In this design, however, no attempt is made to obtain equivalent groups or even to examine the groups to determine whether they are similar before the treatment. Thus, we have no way of knowing if the treatment actually causes any observed differences between the groups.

The three designs just described leave much to be desired in terms of drawing conclusions about what causes what. The experimental designs we describe next are far superior in this respect.

True Experimental Designs

In contrast with the three very simple designs just described, **experimental designs** offer a greater degree of control and, as a result, greater internal validity. The first three of the four designs we discuss in this section share one thing in common: People or other units of study are *randomly assigned to groups*. Such random assignment guarantees that any differences between the groups are probably quite small and, in any case, are due entirely to chance. The last design in this section involves a different strategy: presenting all treatments and any control conditions to a single group.

Design 4: Pretest–Posttest Control Group Design

In a pretest–posttest control group design, an experimental group and a control group are carefully selected through appropriate randomization procedures. The experimental group is observed, subjected to the experimental treatment, and observed again. The control group is

isolated from any influences of the experimental treatment; it is simply observed both at the beginning and at the end of the experiment. The basic format for the pretest–posttest control group design is as follows:

	Group	Time →		
Random Assignment	Group 1	Obs	Tx	Obs
	Group 2	Obs	—	Obs

Such a design, simple as it is, solves two major problems associated with pre-experimental designs. We can (a) determine whether a change takes place after the treatment, and, if so, we can (b) eliminate most other possible explanations (in the form of confounding variables) as to why the change has taken place. Thus, we have a reasonable basis on which to draw a conclusion about a cause-and-effect relationship.

Design 5: Solomon Four-Group Design

One potential problem in the preceding design is that the process of observing or assessing people before administering the experimental treatment may, in and of itself, influence how people respond to the treatment. For instance, perhaps the pretest increases people's motivation: It makes them want to benefit from the treatment they receive. Such an effect is another instance of the *reactivity* effect described in Chapter 4.

To address the question, What effect does pretesting have?, Solomon (1949) proposed an extension of the pretest–posttest control group design that involves four groups, as depicted in the following table:

	Group	Time →		
Random Assignment	Group 1	Obs	Tx	Obs
	Group 2	Obs	—	Obs
	Group 3	—	Tx	Obs
	Group 4	—	—	Obs

The addition of two groups that are not pretested provides a distinct advantage. If the researcher finds that in the final observation, Groups 3 and 4 differ in much the same way that Groups 1 and 2 do, then the researcher can more easily generalize his or her findings to situations in which no pretest has been given. In other words, the Solomon four-group design enhances the *external validity* of the study.

Compared to Design 4, this design obviously involves a larger sample and demands more of the researcher's time and energy. Its principal value is in eliminating pretest influence; when such elimination is desirable, the design is ideal.

Design 6: Posttest-Only Control Group Design

Some life situations defy pretesting. You cannot pretest the forces in a thunderstorm or a hurricane, nor can you pretest growing crops. Additionally, at times you may be unable to locate a suitable pretest, or, as just noted, the very act of pretesting can influence the results of the experimental manipulation. In such circumstances, the posttest-only control group design offers

a possible solution. The design may be thought of as the last two groups of the Solomon four-group design. The paradigm for the posttest-only approach is as follows:

	Group	Time →	
	Random Assignment		
	Group 1	Tx	Obs
	Group 2	—	Obs

Random assignment to groups is, of course, critical in the posttest-only design. Without it, the researcher has nothing more than a static group comparison (Design 3), from which, for reasons previously noted, the researcher has a difficult time drawing inferences about cause and effect.

Design 7: Within-Subjects Design

Earlier we introduced you to the nature of a within-subjects design—also known as a repeated-measures design—in which all participants receive all treatments (including any control conditions) in a research study. Note that we have switched from the term *participant* to the term *subject* here. The latter term has a broader meaning than *participants* in that it can be used to refer to a wide variety of populations—perhaps human beings, dogs, pigeons, or laboratory rats.

In a good within-subjects design, the various treatments are administered very close together in time, in some cases simultaneously. If we use the subscripts *a* and *b* to designate the different treatments and treatment-specific measures, then in its simplest form a within-subjects design is as follows:

	Group	Time →	
	Group 1	Tx _a	Obs _a
		Tx _b	Obs _b

As an example, imagine that a researcher wants to study the effects of illustrations in teaching science concepts to sixth graders. The researcher creates a short textbook that presents, say, 20 different concepts. In the text, all 20 concepts are defined and described with similar precision and depth. In addition, the text illustrates 10 of those concepts (chosen randomly) with pictures or diagrams. After students read the text, they take a quiz that assesses their understanding of the 20 concepts, and the researcher computes separate quiz scores for the illustrated and nonillustrated concepts. If the students perform better on quiz items for illustrated concepts than on items for nonillustrated ones, the researcher can reasonably conclude that, yes, illustrations help students learn science more effectively. In other words, the researcher has identified a cause-and-effect relationship: Illustrations improve science learning.

For a within-subjects design to work, the various forms of treatment must be such that their effects are fairly localized and unlikely to “spread” beyond specifically targeted behaviors. This is the case in the study just described: The illustrations help students learn the particular concepts that have been illustrated but don’t help students learn science more generally. In contrast, it would not make sense to use a within-subjects design to study the effects of two different psychotherapeutic techniques to reduce adolescents’ criminal behaviors: If the same group of adolescents receives both treatments and then shows a significant reduction in juvenile offenses, we might suspect that either treatment could have had a fairly broad impact.

Ideally, too, the two different treatments should be administered repeatedly, one after another, in a balanced but somewhat random order. For example, in the textbook that presents both illustrated and nonillustrated science concepts, we might begin with an illustrated concept, then have two nonillustrated ones, then another illustrated one, another nonillustrated

one, two illustrated ones, and so on, with the presentation of the two conditions being evenly balanced throughout the book.

With the last point in mind, let's return to the dissertation involving the American Civil War lecture described earlier. Each group received each of the three treatments: the imagery, attention, and control conditions. The logistics of the study were such that it was difficult to intermingle the three treatments throughout the lecture; instead, the researcher administered first one treatment (e.g., attention), then another (e.g., imagery), and finally the third (e.g., control). Had the researcher limited her study to a single group, she could not have ruled out an alternative explanation—*when* in the lecture the information appeared (whether it appeared near the beginning, in the middle, or at the end)—for the results she obtained. By using three different groups, each of which had any particular condition in a different part of the lecture, she was able to eliminate that alternative explanation. Strictly speaking, however, because the researcher could neither randomize assignment to groups nor randomly distribute different treatment conditions throughout the lecture, her study is probably better characterized as a quasi-experimental study than a true experimental study. We look more closely at quasi-experimental designs now.

Quasi-Experimental Designs

In the preceding discussion of true experimental designs, we emphasized the importance of *randomness*, either in the selection of group members in a multiple-groups study or in the presentation of different treatments in a single-group study. Sometimes, however, randomness is either impossible or impractical. In such situations, researchers often use **quasi-experimental designs**. When they conduct quasi-experimental studies, they don't control for all confounding variables and so cannot completely rule out some alternative explanations for the results they obtain. They must take whatever variables and explanations they have not controlled for into consideration when they interpret their data.

Design 8: Nonrandomized Control Group Pretest–Posttest Design

The nonrandomized control group pretest–posttest design can perhaps best be described as a compromise between the static group comparison (Design 3) and the pretest–posttest control group design (Design 4). Like Design 3, it involves two groups to which participants have not been randomly assigned. But it incorporates the pretreatment observations of Design 4. In sum, the nonrandomized control group pretest–posttest design can be depicted as follows:

Group	Time →		
Group 1	Obs	Tx	Obs
Group 2	Obs	—	Obs

Without random assignment, there is, of course, no guarantee that the two groups are similar in every respect prior to the experimental treatment or intervention—no guarantee that any differences between them are due entirely to chance. However, an initial observation (e.g., a pretest) can confirm that the two groups are at least similar in terms of the dependent variable under investigation. If, after one group has received the experimental treatment, we then find group differences with respect to the dependent variable, we might reasonably conclude that the posttreatment differences are probably the result of that treatment.

Identifying matched pairs in the two groups is one way of strengthening the pretest–posttest control group design. For instance, if we are studying the effect of a particular preschool program on children's IQ scores, we might find pairs of children—each pair including one child who is enrolled in the preschool program and one who is not—who are the same sex and age and have similar IQ scores before the program begins. Although we cannot rule out all other possible

explanations in this situation (e.g., it may be that the parents who enroll their children in the preschool program are, in general, more concerned about their children's cognitive development), we can at least rule out *some* alternative explanations.

Design 9: Simple Time-Series Design

In its simplest form, a time-series design consists of making a series of observations (i.e., measuring the dependent variable on several occasions), introducing an intervention or other new dynamic into the system, and then making additional observations. If a substantial change is observed in the second series of observations in comparison to the first series, we might reasonably conclude that the cause of the change was the factor introduced into the system. This design thus looks something like the following:

Group Time →

Group 1	Obs	Obs	Obs	Obs	Tx	Obs	Obs	Obs	Obs
---------	-----	-----	-----	-----	----	-----	-----	-----	-----

In such studies, the sequence of observations made prior to the treatment is typically referred to as **baseline** data.

Such a design has been widely used in the physical and biological sciences. Sir Alexander Fleming's discovery that *Penicillium notatum* (a mold) could inhibit staphylococci (a type of bacteria) is an example of this type of design. Fleming had been observing the growth of staphylococci on a culture plate. Then, unexpectedly, a culture plate containing well-developed colonies of staphylococci was contaminated with the spores of *Penicillium notatum*. Fleming observed that the bacteria near the mold seemed to disappear. He intentionally repeated the situation: After periodically observing the bacteria, he introduced the mold. Each time he used this procedure, his subsequent observations were the same: no staph germs near the mold.

The major weakness of this design is the possibility that some other, unrecognized event in the laboratory or outside world may occur at approximately the same time that the experimental treatment does, reflecting the *history* factor described in Figure 9.1. If this other event is actually the cause of the change, any conclusion that the treatment has brought about the change will, of course, be an erroneous one.

Design 10: Control Group, Time-Series Design

In a variation of the time-series design, two groups are observed over a period of time, but one group (a control) doesn't receive the experimental treatment. The design is configured as follows:

Group Time →

Group 1	Obs	Obs	Obs	Obs	Tx	Obs	Obs	Obs	Obs
Group 2	Obs	Obs	Obs	Obs	—	Obs	Obs	Obs	Obs

This design has greater internal validity than the simple time-series design (Design 8). If an outside event is the cause of any changes we observe, then presumably the performance of *both* groups will be altered after the experimental treatment takes place. If, instead, the experimental treatment is the factor that affects performance, then we should see a change only for Group 1.

Design 11: Reversal Time-Series Design

The reversal time-series design uses a within-subjects approach as a way of minimizing—though not entirely eliminating—the probability that outside effects might bring about any changes observed. The intervening experimental treatment is sometimes present, sometimes

absent, and we measure the dependent variable at regular intervals. Thus, we have the following design:

Group	Time →							
Group 1	Tx	Obs	—	Obs	Tx	Obs	—	Obs

To illustrate, suppose we are interested in whether audiovisual materials help students learn astronomy. On some days we might include audiovisual materials in a lesson, and on other days we might omit them. We can then measure how effectively students learn under both conditions. If the audiovisual materials do, in fact, promote student learning, we should see consistently better student performance on those days.

Design 12: Alternating Treatments Design

A variation on the reversal time-series design involves including two or more different forms of experimental treatment in the design. Referring to the two different forms of treatment with the notations Tx_a and Tx_b , we can depict this design in the following way:

Group	Time →													
Group 1	Tx_a	Obs	—	Obs	Tx_b	Obs	—	Obs	Tx_a	Obs	—	Obs	Tx_b	Obs

If such a sequence were pursued over a long enough time span, we would hope to see different effects for the two different treatments.

Design 13: Multiple Baseline Design

Designs 11 and 12 are based on the assumption that the effects of any single treatment are temporary and limited to the immediate circumstances. Thus, these designs won't work if a treatment is apt to have long-lasting and perhaps fairly general effects. Furthermore, if an experimental treatment is likely to be quite beneficial for all participants, then ethical considerations may discourage us from including an untreated control group. In such instances, a multiple baseline design provides a good alternative. This design requires at least two groups. Prior to the treatment, baseline data are collected for all groups, and then the treatment itself is introduced at a different time for each group. In its simplest form, a multiple baseline design might be configured as follows:

Group	Time →					
	Baseline →			Treatment →		
Group 1	—	Obs	—	Obs	Tx	Obs
	Baseline →			Treatment →		
Group 2	—	Obs	—	Obs	Tx	Obs

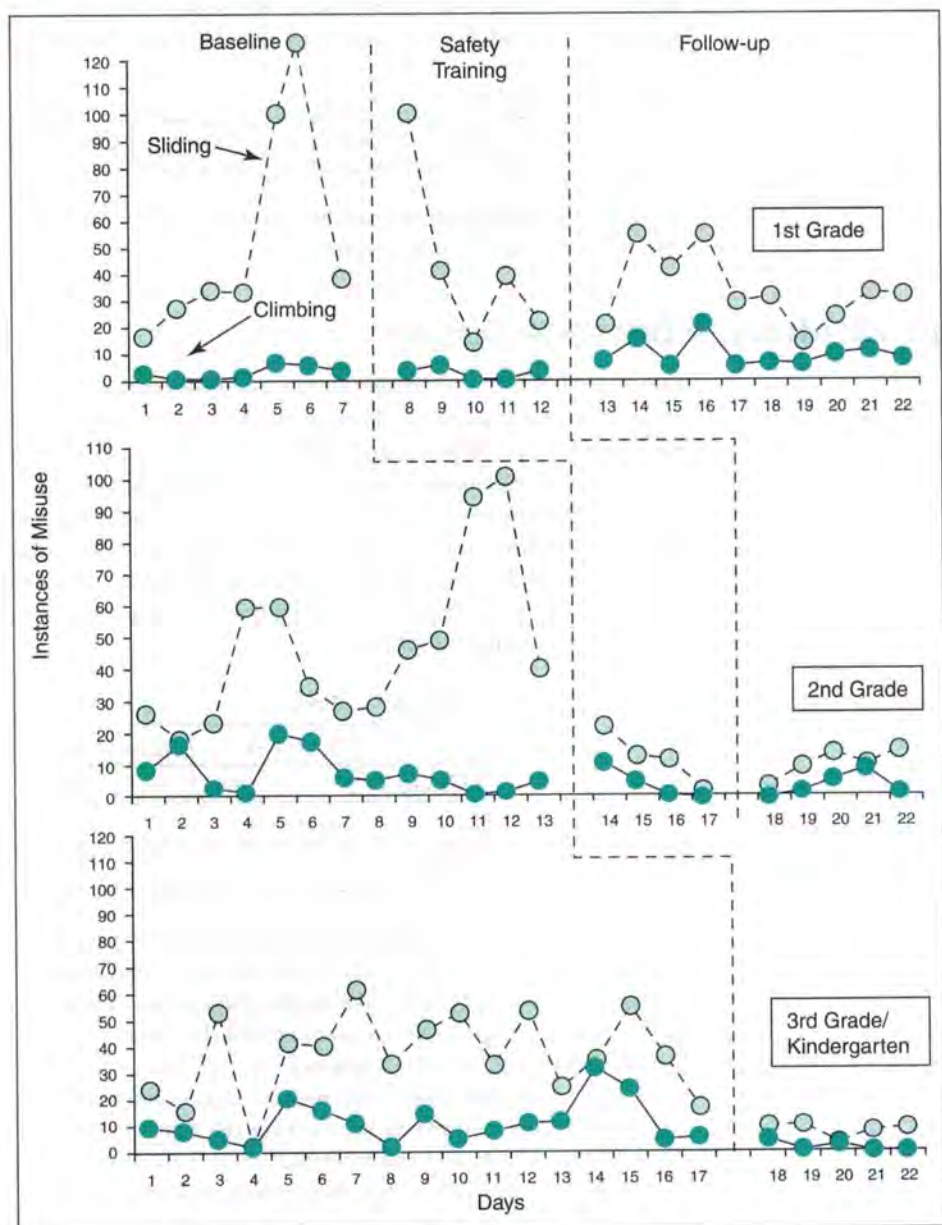
A study by Heck, Collins, and Peterson (2001) provides an example of this approach. The researchers wanted to determine if instruction in playground safety would decrease elementary school children's risky behaviors on the playground. The treatment in this case involved a five-day intervention in which a woman visited the classroom to talk about potentially risky behaviors on slides and climbing equipment, as well as about the unpleasant consequences that might result from such behaviors. The woman visited four classrooms on different weeks; a random selection process resulted in her visiting the first-grade class one week, the second-grade class the following week, and the kindergarten and third-grade classes (which went to recess at the same time) the week after that. Meanwhile, two independent observers simultaneously counted the number of risky behaviors on the playground before, during, and after the interventions in

the four classrooms. The data they collected are depicted in Figure 9.2; numbers of risky behaviors on the slide are shown with the lighter dots, whereas those on the climbing equipment are shown with the darker dots. Notice that each group has data for three time periods: a pre-intervention baseline period, the five-day safety-training period, and a posttraining follow-up period. As you can see, once safety training began, the children in the second-grade and third-grade/kindergarten groups showed noticeable declines in risky behaviors on the slide and, to a lesser extent, on the climbing equipment (where risky behavior was relatively infrequently to begin with). Because the behavior changes occurred at different times for these two groups, and in particular because the changes for each group occurred at about the time that the group began its safety training, the researchers reasonably concluded that the training itself (rather than some other factor) was probably the reason for the changes. The first graders, who received the training first, showed little or no benefit from it, especially for the climbing equipment. Perhaps the trainer was still perfecting her training procedures that first week; however, we have no way of knowing for sure why the training appeared to be relatively ineffective for the first group.

FIGURE 9.2

Instances of risky behavior on slides and climbers by grade level; third graders and kindergartners shared a single recess

Reprinted from "Decreasing Children's Risk Taking on the Playground" by A. Heck, J. Collins, and L. Peterson, 2001, *Journal of Applied Behavior Analysis*, 34, p. 351. Reprinted with permission of the Society for the Experimental Analysis of Behavior, Inc.



Ex Post Facto Designs

In many situations, it is either unethical or impossible to manipulate certain variables in order to investigate their potential influence on other variables. For example, a researcher cannot intentionally infect people with a potentially deadly new virus, withhold instruction, ask parents to abuse their children, or modify a person's personality to compare the effects of these factors on the dependent variables in one's research problem.

Ex post facto designs² (the term *ex post facto* literally means "after the fact") provide an alternative means by which a researcher can investigate the extent to which specific independent variables—perhaps involving a virus, lack of schooling, a history of family violence, or a personality trait—may possibly affect the dependent variable(s) of interest. In an ex post facto study, a researcher identifies *events that have already occurred or conditions that are already present* and then collects data to investigate a possible relationship between these factors and subsequent characteristics or behaviors. In particular, after observing that differing circumstances have prevailed for two or more different groups—such circumstances comprise the independent variable—the researcher tries to determine whether the groups differ on some other, dependent variable. For example, a researcher might identify two groups of adults with different immunization records—those who, as children, were vaccinated against measles and those who were not—and then calculate the percentage of reported cases of measles in each group. Similarly, a researcher might identify two groups of 10-year-olds—those who had extensive musical training in preschool and those whose preschools provided no such training—and compare the musical skills of the two groups of children.

Ex post facto designs are often confused with correlational or experimental designs because they share certain characteristics with each of these other design types. Like correlational research, ex post facto research involves looking at existing circumstances. But like experimental research, it has clearly identifiable independent and dependent variables.

Unlike experimental studies, however, ex post facto designs involve no direct manipulation of the independent variable: The presumed "cause" has already occurred. To the extent that such manipulation is not possible, the researcher cannot draw firm conclusions about cause and effect. The problem here is that the experimenter cannot control for confounding variables that may provide alternative explanations for any group differences that are observed.

Although an ex post facto study lacks the control element—and so does not allow us to draw definite conclusions about cause and effect—it is nevertheless a legitimate research method that pursues truth and seeks the solution of a problem through the analysis of data. Medicine uses it widely in its research activities. Physicians discover an illness and then initiate their search *after the fact*. They delve into antecedent events and conditions to discover a possible cause for the illness. Such was the approach of medical researchers when the AIDS virus emerged in the 1980s.

Like experimental designs, ex post facto designs can take a variety of forms. Here we present one possible design for illustrative purposes. We present a second ex post facto design in the subsequent section on factorial designs.

Design 14: Simple Ex Post Facto Design

Design 14 is similar to the static group comparison (Design 3), which we included in our discussion of pre-experimental designs. The important difference is one of timing: In this case, the "treatment" in question occurred long before the study began; hence, we call it an *experience* rather than a treatment because the researcher has not been responsible for imposing it. A simple

² Ex post facto designs are also known as *causal-comparative* designs. However, as Johnson (2001) has pointed out, the latter term may mislead novice researchers to believe that such designs show cause and effect as clearly and definitively as true experimental designs. In reality, such designs never eliminate all other possible explanations for an observed effect; thus, they cannot truly show cause and effect.

ex post facto design can be depicted as follows, where *Exp* refers to a prior experience that one group has had and another has not:

Group	Time →	
	Prior event(s)	Investigation period
Group 1	Exp	Obs
Group 2	—	Obs

An obvious variation on this design is one in which Group 2 has an experience as well, albeit a different experience from that of Group 1.

Such designs are common in studying the possible effects of previously occurring environmental variables such as television viewing habits, child abuse, and malnutrition. They are also used in studying the potential influences of pre-existing characteristics—perhaps those that are inherited or congenital—such as gender, mental illness, and physical disability. (In the latter instances, we might want to replace the term *experience* with a term such as *characteristic*.) The most we can conclude from these studies is that certain behaviors or other variables tend to be *associated* with certain pre-existing conditions; we can never determine that those other variables were actually caused by those conditions.

Factorial Designs

Thus far, we have been describing designs in which only one independent variable is studied. Yet in many situations, a researcher examines the effects of two or more independent variables in a single study; this approach is known as a **factorial design**.

Design 15: Two-Factor Experimental Design

In its simplest form—one involving two independent variables, which we call *Variable 1* and *Variable 2*—such a design might look something like the following:

Group		Time →		
		Treatments related to the two variables may occur simultaneously or sequentially		
		Treatment related to Variable 1	Treatment related to Variable 2	
Random Assignment	Group 1	Tx ₁	Tx ₂	Obs
	Group 2	Tx ₁	—	Obs
	Group 3	—	Tx ₂	Obs
	Group 4	—	—	Obs

We can determine the effects of the first independent variable by comparing the performance of Groups 1 and 2 with that of Groups 3 and 4. We can determine the effects of the second independent variable by comparing Groups 1 and 3 with Groups 2 and 4. If you think you've seen this design before, in a way you have. This is simply a more generalized form of the Solomon four-group design (Design 5), but we are no longer limiting ourselves to having the presence or absence of a pretest be one of our independent variables.

Such a design allows us to examine not only the possible effects of two independent variables but also the possible *interaction* of the variables as they influence the dependent variable. For example, imagine that, after presenting both treatments, we find that Groups 2, 3, and 4 show similar performance but that Group 1 outperforms the other three. Such a result may indicate that neither independent variable produces a particular effect on its own—that *both* variables are necessary to bring about the effect.

Design 16: Combined Experimental and Ex Post Facto Design

In the factorial design just presented, participants are randomly assigned to groups in a true experimental study. But it is also possible to combine elements of experimental research and ex post facto research into a single factorial design. In its simplest form, such a design looks like the following:

Group	Time →				
	Prior event(s)	Investigation period →			
Group 1	Exp _a	Random Assignment	Group 1a	Tx _a	Obs
			Group 1b	Tx _b	Obs
Group 2	Exp _b	Random Assignment	Group 2a	Tx _a	Obs
			Group 2b	Tx _b	Obs

In this case, the researcher initially divides the sample into two groups based on the participants' previous experiences or pre-existing conditions; this is the *ex post facto* part of the study. Then the researcher randomly assigns members of each group to one of two treatment groups (or perhaps a treatment group and a control group); this is the *experimental* part of the study. The result is four groups that represent all four possible combinations of the previous experience/pre-existing characteristic and the treatment variable. Such a design enables the researcher to study how an experimental manipulation may influence a particular dependent variable *and* how a previous experience or pre-existing characteristic may possibly interact with that manipulation.

In a variation of such a design, the experimental manipulation might be a within-subjects variable rather than a between-groups variable. As an example, one of us authors once joined forces with two colleagues and a graduate student to test the hypothesis that people with different educational backgrounds interpret and remember maps differently and, more specifically, that only people with a background in geography apply general principles of geography when they interpret maps (J. E. Ormrod, Ormrod, Wagner, & McCallin, 1988). We constructed two maps to test our hypothesis. One map was arranged in accordance with the patterns of a typical city; for instance, a downtown business district was located at a point where it could be easily reached from different directions (this is typical), and factories, a lumberyard, and low-income housing were situated near railroad tracks (also typical). The second map was less "logical" in the sense that it violated basic geographic principles; for instance, a

river originated in the plains and ran *up* into a mountain range, and various transportation networks did not interconnect in ways that they normally do. The two different maps reflected one of our independent variables: logic (or lack thereof) of the spatial arrangement of features within a map.

Three groups of college professors—geographers, sociologists, and educational psychologists—provided the basis for our second independent variable: educational background. We asked each professor to study each of the two maps aloud for three two-minute intervals (we tape-recorded what they said during the study sessions) and then, after each interval, to draw as much of the map as he or she could remember.

Thus, if we call the two maps Tx_a (logical map) and Tx_b (illogical map), our design looked like the following:

Group	Time →							
Geographers	Tx_a	Obs	Obs	Obs	Tx_b	Obs	Obs	Obs
Sociologists	Tx_a	Obs	Obs	Obs	Tx_b	Obs	Obs	Obs
Educational psychologists	Tx_a	Obs	Obs	Obs	Tx_b	Obs	Obs	Obs

In this situation, one independent variable—the logic or illogic of the map presented—was a variable we directly manipulated, and we presented it to all participants in a *within-subjects* (repeated-measures) manner. The second independent variable, educational background, was a pre-existing condition and therefore something we could *not* control; this was the *ex post facto* part of the design.

The upshot of the study was that there was an *interaction* between the two independent variables, map logic and educational background. In particular, the geographers remembered more of the logical map than they did of the illogical map; in contrast, the sociologists and educational psychologists remembered each map with equal accuracy. We interpreted this result to indicate that only the geographers were applying geographic principles to study the maps and that they could use such principles effectively only with the geographically logical one. We supported our conclusion with a qualitative element in our study; that is, we used a *mixed-methods design*. In particular, we conducted content analyses of the professors' study sessions. Indeed, the content analyses revealed that the geographers had applied many geographic principles to the logical map but had trouble applying them to the illogical one. Meanwhile, the sociologists and educational psychologists studied both maps in a haphazard fashion, with few attempts to make sense of what they saw on the maps.

Table 9.1 provides a summary of the pre-experimental, experimental, quasi-experimental, ex post facto, and factorial designs described in the preceding sections. Keep in mind that, as stated earlier, this is not an exhaustive list of experimental and ex post facto designs. You can combine and expand on these designs in a number of ways—and perhaps incorporate elements of qualitative or descriptive-quantitative designs (e.g., content analysis or longitudinal data collection) as well—to more effectively address your own research question.

TABLE 9.1 Summary of Experimental and Ex Post Facto Designs

Name of the Design	Aim of the Research	Graphic Depiction	Comments on the Design							
Pre-Experimental Designs										
1. One-shot experimental case study	To show that one event (a treatment) precedes another event (the observation)	<table><tr><th>Group</th><th>Time →</th></tr><tr><td>Group 1</td><td>Tx Obs</td></tr></table>	Group	Time →	Group 1	Tx Obs	Shows a before-and-after sequence but cannot substantiate that this is a cause-and-effect relationship.			
Group	Time →									
Group 1	Tx Obs									
2. One group pretest-posttest design	To show that change occurs after a treatment	<table><tr><th>Group</th><th>Time →</th></tr><tr><td>Group 1</td><td>Obs Tx Obs</td></tr></table>	Group	Time →	Group 1	Obs Tx Obs	Provides a measure of change but yields no conclusive results about the cause of the change.			
Group	Time →									
Group 1	Obs Tx Obs									
3. Static group comparison	To show that a group receiving a treatment behaves differently than one receiving no treatment	<table><tr><th>Group</th><th>Time →</th></tr><tr><td>Group 1</td><td>Tx Obs</td></tr><tr><td>Group 2</td><td>— Obs</td></tr></table>	Group	Time →	Group 1	Tx Obs	Group 2	— Obs	Fails to determine pretreatment equivalence of groups.	
Group	Time →									
Group 1	Tx Obs									
Group 2	— Obs									
True Experimental Designs										
4. Pretest-posttest control group design	To show that change occurs following, but only following, a particular treatment	<table><tr><th>Group</th><th>Time →</th></tr><tr><td rowspan="2">Random Assignment</td><td>Group 1 Obs Tx Obs</td></tr><tr><td>Group 2 Obs — Obs</td></tr></table>	Group	Time →	Random Assignment	Group 1 Obs Tx Obs	Group 2 Obs — Obs	Controls for many potential threats to internal validity.		
Group	Time →									
Random Assignment	Group 1 Obs Tx Obs									
	Group 2 Obs — Obs									
5. Solomon four-group design	To investigate the possible effect of pretesting	<table><tr><th>Group</th><th>Time →</th></tr><tr><td rowspan="4">Random Assignment</td><td>Group 1 Obs Tx Obs</td></tr><tr><td>Group 2 Obs — Obs</td></tr><tr><td>Group 3 — Tx Obs</td></tr><tr><td>Group 4 — — Obs</td></tr></table>	Group	Time →	Random Assignment	Group 1 Obs Tx Obs	Group 2 Obs — Obs	Group 3 — Tx Obs	Group 4 — — Obs	Enables the researcher to determine how pretesting may affect the final outcome observed.
Group	Time →									
Random Assignment	Group 1 Obs Tx Obs									
	Group 2 Obs — Obs									
	Group 3 — Tx Obs									
	Group 4 — — Obs									
6. Posttest-only control group design	To determine the effects of a treatment when pretesting cannot or should not occur	<table><tr><th>Group</th><th>Time →</th></tr><tr><td rowspan="2">Random Assignment</td><td>Group 1 Tx Obs</td></tr><tr><td>Group 2 — Obs</td></tr></table>	Group	Time →	Random Assignment	Group 1 Tx Obs	Group 2 — Obs	Uses the last two groups in the Solomon four-group design; random assignment to groups is critical for maximizing group equivalence.		
Group	Time →									
Random Assignment	Group 1 Tx Obs									
	Group 2 — Obs									
7. Within-subjects design	To compare the relative effects of different treatments for the same participants	<table><tr><th>Group</th><th>Time →</th></tr><tr><td rowspan="2">Group 1</td><td>Tx_a Obs_a</td></tr><tr><td>Tx_b Obs_b</td></tr></table>	Group	Time →	Group 1	Tx _a Obs _a	Tx _b Obs _b	Useful only when effects of each treatment are temporary and localized.		
Group	Time →									
Group 1	Tx _a Obs _a									
	Tx _b Obs _b									
Quasi-Experimental Designs										
8. Nonrandomized control group pretest-posttest design	To show that two groups are equivalent with respect to the dependent variable prior to the treatment, thus eliminating initial group differences as an explanation for posttreatment differences	<table><tr><th>Group</th><th>Time →</th></tr><tr><td>Group 1</td><td>Obs Tx Obs</td></tr><tr><td>Group 2</td><td>Obs — Obs</td></tr></table>	Group	Time →	Group 1	Obs Tx Obs	Group 2	Obs — Obs	Differs from experimental designs because test and control groups are not totally equivalent; equivalence on the pretest ensures equivalence only for variables that have specifically been measured.	
Group	Time →									
Group 1	Obs Tx Obs									
Group 2	Obs — Obs									

TABLE 9.1 Summary of Experimental and Ex Post Facto Designs (*continued*)

Name of the Design	Aim of the Research	Graphic Depiction	Comments on the Design																											
Quasi-Experimental Designs (continued)																														
9. Simple time-series experiment	To show that, for a single group, change occurs during a lengthy period only after the treatment has been administered	<table><tr><th>Group</th><th colspan="5">Time →</th></tr><tr><td>Group 1</td><td>Obs</td><td>Obs</td><td>Tx</td><td>Obs</td><td>Obs</td></tr></table>	Group	Time →					Group 1	Obs	Obs	Tx	Obs	Obs	Provides a stronger alternative to Design 2; external validity can be increased by repeating the experiment in different places under different conditions.															
Group	Time →																													
Group 1	Obs	Obs	Tx	Obs	Obs																									
10. Control group, time-series design	To bolster the internal validity of the preceding design with the addition of a control group	<table><tr><th>Group</th><th colspan="5">Time →</th></tr><tr><td>Group 1</td><td>Obs</td><td>Obs</td><td>Tx</td><td>Obs</td><td>Obs</td></tr><tr><td>Group 2</td><td>Obs</td><td>Obs</td><td>—</td><td>Obs</td><td>Obs</td></tr></table>	Group	Time →					Group 1	Obs	Obs	Tx	Obs	Obs	Group 2	Obs	Obs	—	Obs	Obs	Involves conducting parallel series of observations for experimental and control groups.									
Group	Time →																													
Group 1	Obs	Obs	Tx	Obs	Obs																									
Group 2	Obs	Obs	—	Obs	Obs																									
11. Reversal time-series design	To show, in a single group or individual, that a treatment consistently leads to a particular effect	<table><tr><th>Group</th><th colspan="6">Time →</th></tr><tr><td>Group 1</td><td>Tx</td><td>Obs</td><td>—</td><td>Obs</td><td>Tx</td><td>Obs</td></tr></table>	Group	Time →						Group 1	Tx	Obs	—	Obs	Tx	Obs	Is an on-again, off-again design in which the experimental treatment is sometimes present, sometimes absent.													
Group	Time →																													
Group 1	Tx	Obs	—	Obs	Tx	Obs																								
12. Alternating treatments design	To show, in a single group or individual, that different treatments have different effects	<table><tr><th>Group</th><th colspan="6">Time →</th></tr><tr><td>Group 1</td><td>Tx_a</td><td>Obs</td><td>—</td><td>Obs</td><td>Tx_b</td><td>Obs</td></tr></table>	Group	Time →						Group 1	Tx _a	Obs	—	Obs	Tx _b	Obs	Involves sequentially administering different treatments at different times and comparing their effects against the possible consequences of nontreatment.													
Group	Time →																													
Group 1	Tx _a	Obs	—	Obs	Tx _b	Obs																								
13. Multiple baseline design	To show the effect of a treatment by initiating it at different times for different groups or individuals, or perhaps in different settings for a single individual	<table><tr><th>Group</th><th colspan="6">Time →</th></tr><tr><td>Group 1</td><td>—</td><td>Obs</td><td>Tx</td><td>Obs</td><td>Tx</td><td>Obs</td></tr><tr><td>Group 2</td><td>—</td><td>Obs</td><td>—</td><td>Obs</td><td>Tx</td><td>Obs</td></tr></table>	Group	Time →						Group 1	—	Obs	Tx	Obs	Tx	Obs	Group 2	—	Obs	—	Obs	Tx	Obs	Involves tracking two or more groups or individuals over time, or tracking a single individual in two or more settings, for a lengthy period of time, as well as initiating the treatment at different times for different groups, individuals, or settings.						
Group	Time →																													
Group 1	—	Obs	Tx	Obs	Tx	Obs																								
Group 2	—	Obs	—	Obs	Tx	Obs																								
Ex Post Facto Designs																														
14. Simple ex post facto design	To show the possible effects of an experience that occurred, or a condition that was present, prior to the investigation	<table><tr><th>Group</th><th colspan="2">Time →</th></tr><tr><td>Group 1</td><td>Exp</td><td>Obs</td></tr><tr><td>Group 2</td><td>—</td><td>Obs</td></tr></table>	Group	Time →		Group 1	Exp	Obs	Group 2	—	Obs	May show a difference between groups but does not conclusively demonstrate that the difference is due to the prior experience/condition in question.																		
Group	Time →																													
Group 1	Exp	Obs																												
Group 2	—	Obs																												
Factorial Designs																														
15. Two-factor experimental design	To study the effects of two experimenter-manipulated variables and their possible interaction	<table><tr><th colspan="2"></th><th>Group</th><th colspan="3">Time →</th></tr><tr><td rowspan="4">Random Assignment</td><td>Group 1</td><td>Tx₁</td><td>Tx₂</td><td>Obs</td><td></td></tr><tr><td>Group 2</td><td>Tx₁</td><td>—</td><td>Obs</td><td></td></tr><tr><td>Group 3</td><td>—</td><td>Tx₂</td><td>Obs</td><td></td></tr><tr><td>Group 4</td><td>—</td><td>—</td><td>Obs</td><td></td></tr></table>			Group	Time →			Random Assignment	Group 1	Tx ₁	Tx ₂	Obs		Group 2	Tx ₁	—	Obs		Group 3	—	Tx ₂	Obs		Group 4	—	—	Obs		Requires a larger sample size than two-group studies; random assignment to treatments is essential.
		Group	Time →																											
Random Assignment	Group 1	Tx ₁	Tx ₂	Obs																										
	Group 2	Tx ₁	—	Obs																										
	Group 3	—	Tx ₂	Obs																										
	Group 4	—	—	Obs																										
16. Combined experimental and ex post facto design	To study the possible effects of an experimenter-manipulated variable, a previously existing condition, and the interaction between the two	<table><tr><th colspan="2"></th><th>Group</th><th colspan="3">Time →</th></tr><tr><td rowspan="2">Group 1</td><td rowspan="2">Exp_a</td><td rowspan="4">Random Assignment</td><td>Group 1a</td><td>Tx_a</td><td>Obs</td></tr><tr><td>Group 1b</td><td>Tx_b</td><td>Obs</td></tr><tr><td rowspan="2">Group 2</td><td rowspan="2">Exp_b</td><td>Group 2a</td><td>Tx_a</td><td>Obs</td></tr><tr><td>Group 2b</td><td>Tx_b</td><td>Obs</td></tr></table>			Group	Time →			Group 1	Exp _a	Random Assignment	Group 1a	Tx _a	Obs	Group 1b	Tx _b	Obs	Group 2	Exp _b	Group 2a	Tx _a	Obs	Group 2b	Tx _b	Obs	Requires a larger sample size than two-group studies; random assignment to the experimenter-manipulated variable is essential.				
		Group	Time →																											
Group 1	Exp _a	Random Assignment	Group 1a	Tx _a	Obs																									
			Group 1b	Tx _b	Obs																									
Group 2	Exp _b		Group 2a	Tx _a	Obs																									
			Group 2b	Tx _b	Obs																									

PRACTICAL APPLICATION Determining Possible Cause-and-Effect Relationships

The research designs described in this chapter vary considerably in the degree to which they control for potential confounding variables—variables that threaten a study's internal validity—and thus they also vary in terms of the degree to which they enable a researcher to draw firm conclusions about cause-and-effect relationships. The following checklist can help you evaluate a research design with respect to its internal validity.

✓ CHECKLIST

Looking for Confounding Variables

If you are planning a study in which you hope to find one or more cause-and-effect relationships—or if, instead, you are evaluating another person's research proposal or report—scrutinize the study with the following questions in mind:

- _____ 1. What are the independent and dependent variables in the study?
 Independent variable(s): _____
 Dependent variable(s): _____
- _____ 2. Is every independent variable actively manipulated by the researcher? ____ Yes ____ No
- _____ 3. If the researcher is manipulating one or more independent variables, what precautions is the researcher taking to ensure that the manipulation is minimizing or eliminating the potential effects of confounding variables? For example, is the researcher:
 - Keeping certain other variables constant? If so, which ones?

 - Including a control group or at least two treatment groups?

 - Randomizing assignment to groups?

 - Using a within-subjects (repeated-measures) design?

 - Using other appropriate strategies? If so, which ones?

- _____ 4. If the researcher is *not* manipulating one or more independent variables, what precautions is the researcher taking to control for confounding variables? For example, is the researcher:
 - Using one or more pretests to assess before-treatment group equivalence? If so, what variables are being pretested?

 - Identifying matched pairs? If so, on the basis of what variables?

 - Statistically controlling for confounding variables? If so, which ones?

- _____ 5. If the researcher is conducting a single-group or single-subject study, is the researcher:
- Conducting a series of observations both before and after the intervention (a time-series design)? _____
 - Alternating either between two or more treatments or between treatment and nontreatment, with a new observation being made after each treatment or nontreatment (a reversal design)? _____
 - Beginning an intervention at different times for different individuals or different contexts (a multiple baseline design)? _____
- _____ 6. What other variables in the study (either identified or not identified by the researcher) might potentially affect the dependent variable? _____
- _____ 7. To what extent might each of the following factors threaten the study's internal validity? If any of these factors pose a potential threat, how is the researcher minimizing or eliminating its influence? (Refer to Figure 9.1.)
- History: _____
- Maturation: _____
- Testing: _____
- Instrumentation: _____
- Statistical regression: _____
- Selection: _____
- Attrition: _____
- _____ 8. With your answers to the preceding questions in mind, explain whether the study's results justifiably demonstrate a cause-and-effect relationship: _____
- _____
- _____
- _____

Meta-Analyses

Remember, we can conclude that a cause-and-effect relationship exists between an independent variable and a dependent variable only when we have directly manipulated the independent variable and have controlled for confounding variables that might offer alternative explanations for any changes in the dependent variable. Even when we have taken such precautions, however, there is the possibility that our alleged "cause" doesn't really produce the effect we think it does—that the situation we have just observed is a one-time-in-a-million fluke.

In Chapter 4 we introduced the idea of *replication*: A research study should be repeatable. In fact, we gain greater confidence in our research findings when a study is repeated over and over again—perhaps with a different population, in a different setting, or with slight variations on the treatment implementation.

Once researchers have conducted many such replications, another researcher may come along and conduct a **meta-analysis**—that is, an analysis of the analyses. In particular, the researcher combines the results of many experimental and/or ex post facto studies to determine whether they yield consistent, predictable results. A meta-analysis is primarily a statistical technique, and so we describe this procedure in greater depth in Chapter 11.

Conducting Experiments on the Internet



In Chapter 8 we mentioned that some researchers now conduct research studies on the Internet. Although most of these studies can best be categorized as descriptive studies, we occasionally see experimental studies as well. For instance, one of us authors once visited the website “Psychological Research on the Net,” which provides links to numerous sites that host online research projects.³ To learn more about this growing approach to data collection, she became a participant in several online studies that were active at the time. Although most of the studies involved completing questionnaires and so appeared to be correlational or survey studies, one of them was clearly an experimental study. In particular, this author was asked to (a) read and study a story that was illustrated by several photographs; (b) read three additional stories, one of which was quite similar to the initial story; and (c) answer a series of questions about details in the stories. In a subsequent debriefing on the website, she learned that she had been randomly assigned to the experimental group in the second part of the study; other participants were assigned to a control group, in which all three stories were quite different from the initial story. The researcher was investigating the possible effects that a similar story in Part b might have on recall for the story in Part a.

Internet-based experimental studies don’t necessarily have to be one-shot affairs. For example, in one online study (Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008), researchers enticed people into participating in a three-session experiment with the promise that for every session they completed, their names would be entered into an end-of-study lottery that would award cash prizes. A total of 1,354 people completed all three sessions; they ranged in age from 18 to 72 and lived in various countries around the world. In Session 1 of the experiment, participants studied a list of 32 obscure trivia facts, such as the answer to “What European nation consumes the most spicy Mexican food?” (p. 1097), and they continued to study each fact until they could correctly recall it. After this first session, participants were divided into different treatment groups that varied in terms of the timing for Sessions 2 and 3, and they were sent e-mail messages when it was time to complete these subsequent sessions. In Session 2 (which might be as little as three minutes or as much as 105 days after Session 1), participants studied the trivia facts again, this time studying each one twice. Then, in Session 3 (which was 7, 35, 70, or 350 days after Session 2), participants were asked to remember as many of the facts as they could. The findings of the study are important for any conscientious student to note: Especially when the final test session was considerably delayed (e.g., by 2½ months or almost a year), people who spread out their studying more (i.e., those with a longer delay between Sessions 1 and 2) remembered more facts.⁴ (In case you’re curious, Norwegians are especially partial to spicy Mexican food.)

In some instances, an Internet-based research study might be quite suitable for your research question. Keep in mind, however, that ethical practices ensuring protection from harm, informed consent, and right to privacy are as important in online experimental research as they are in any face-to-face studies. The suggestions for ethical practices presented in Chapter 8 for online questionnaires are equally applicable to online experiments (see the Practical Application “Using the Internet to Collect Data for a Descriptive Study” on pp. 205–206).

³As noted in Chapter 8, you can reach the site by going to the website of the Association for Psychological Science (www.psychologicalscience.org); click on “Psychology links,” scroll down to “Other Sites of Interest,” and then click on “Online Psychology Experiments.” Alternatively, you can go directly to the site, which, as this book goes to press, is located at psych.hanover.edu/research/exponnet.html.

⁴As you might guess, the attrition (dropout) rate was higher for participants with longer between-session delays. To determine the extent to which the differing attrition rates for different treatment groups might jeopardize the study’s internal validity, the researchers collected basic demographic data at the beginning of Session 1. During the data analysis, the researchers found no significant differences in any demographic variables or in learning speed between participants who completed all three sessions and those who did not.

Remember, too, that the sample you get in an online study will hardly be representative of the overall population; for instance, it is likely to consist largely of college-educated, computer-literate people who enjoy participating in research studies. An additional problem is that you cannot observe your participants to determine whether they are accurately reporting demographic information (their age, gender, etc.) and whether they are truly following the instructions you present. Accordingly, unless you are interested in a topic such as very-long-term memory (as Cepeda and his colleagues were in their 2008 study) and can carefully control the conditions under which people are participating, we suggest that you use an Internet-based study primarily to formulate tentative hypotheses or to pilot-test experimental materials you plan to use in a more controlled and observable situation.

Testing Your Hypotheses, and Beyond

Experimental and ex post facto studies typically begin with specific research hypotheses, and subsequent statistical analyses should, of course, be conducted to test these hypotheses. Such analyses often take the form of a *t* test, analysis of variance, or analysis of covariance (we discuss these procedures in Chapter 11).

Yet one's analyses need not be restricted *only* to the testing of initially stated hypotheses. Oftentimes a study may yield additional results—results that are unexpected yet intriguing—that merit analysis. There is no reason why the researcher can't examine these findings as well, perhaps statistically, perhaps not.

CONCEPTUAL ANALYSIS EXERCISE Identifying Research Designs

As a way of reviewing the designs described in this chapter, we offer a brief pop quiz. Following are short summaries of five research studies. The studies don't necessarily fit exactly into one of the design categories presented, but each one is definitely *experimental*, *quasi-experimental*, or *ex post facto* in nature. Identify the type of research that each study reflects. The answers appear after the suggested readings at the end of the chapter.

1. Two researchers want to see if a particular training program is effective in teaching horses to enter a horse trailer without misbehaving in the process—that is, without rearing, trying to turn around, or in some other way resisting entry into the trailer. Five horses (Red, Penny, Shadow, Sammy, and Fancy) go through the training, with each horse beginning training on a different day. For each horse, an observer counts the number of misbehaviors every day prior to and during training, with data being collected for a time span of at least 45 days (Ferguson & Rosales-Ruiz, 2001).
2. Two researchers wonder whether an eyewitness's memory of an event is affected by questions that he or she is asked subsequent to the event. To find out, the researchers show adults a film that depicts a car accident. Each adult is then asked one of five questions (randomly selected) about the accident:
 - About how fast were the cars going when they *contacted* each other?
 - About how fast were the cars going when they *hit* each other?
 - About how fast were the cars going when they *bumped into* each other?
 - About how fast were the cars going when they *collided into* each other?
 - About how fast were the cars going when they *smashed into* each other?
 The researchers compute the average speed given in response to each of the five questions to determine whether the questions have influenced participants' "memory" for the accident (Loftus & Palmer, 1974).
3. A researcher studies the effects of two different kinds of note-taking training (one of which is a placebo) on the kinds of notes that college students take. Her sample consists of students enrolled in two sections of an undergraduate course in educational psychology; with the flip of a coin, she randomly determines which section will be the treatment group and which will be the control group. She analyzes the content of students' class notes both before and after the training, making the prediction that the two

groups' notes will be similar before the training but qualitatively different after the training (Jackson, 1996).

4. At the request of the National Park Service, two researchers at Rocky Mountain National Park investigate the degree to which signs along hiking trails might influence hikers' behaviors. Park Service officials are concerned that the heavy traffic on one particular hiking trail, the trail to Emerald Lake, may be having a negative impact on the local environment; they would like to divert some traffic to a lesser-used trail to Lake Haiyaha, which begins at the same place as the Emerald Lake trail. One day in early summer, the researchers hide battery-operated, optic counters at key locations along the two trails to record the number of hikers. The study has four phases: (1) at the spot where the two trails originate, only signs indicating the destinations of the two trails are present; (2) a "positively worded" sign is added that describes the attractive features of the Lake Haiyaha trail and encourages hikers to use it; (3) the positively worded sign is replaced by a "negatively worded" sign that describes the crowdedness of the Emerald Lake trail and discourages its use; and (4) both the positively worded and negatively worded signs are posted. The researchers compare the frequency of hikers during each of the four phases (R. K. Ormrod & Trahan, 1982).
5. A team of researchers has a sample of elementary school boys, some of whom have been identified as having attention-deficit hyperactivity disorder (ADHD) and some of whom have not. One of the researchers asks each boy to interpret several social situations that are depicted in a series of black-and-white drawings (e.g., one sequence of drawings shows a sequence of events at a Halloween party). Some of the situations involve antisocial behavior (e.g., aggression), and other situations involve prosocial behavior (e.g., sharing). The researchers compare the interpretations that boys with ADHD make with the interpretations that boys without ADHD make with respect to both kinds of situations (Milch-Reich, Campbell, Pelham, Connelly, & Geva, 1999).

A Sample Dissertation

To illustrate how an experimental study might appear in its written form, we present excerpts from Virginia Kinnick's doctoral dissertation conducted at the University of Colorado (Kinnick, 1989). The researcher, a faculty member in the School of Nursing at another university, had considerable experience teaching nursing students the knowledge and skills they would need when working with women who were in the process of delivering a baby, and her interest lay in learning more about teaching such knowledge and skills effectively.

During a woman's labor prior to the delivery of her baby, a fetal monitor is often used to assess the baby's heart rate, and the maternity nurse must frequently check the monitor for signs that the baby might be experiencing exceptional and potentially harmful stress. The researcher wanted to determine whether a particular method of teaching concepts (one described by Tennyson and Cocchiarella) might be more effective for teaching fetal monitoring skills than the method traditionally used in nursing education programs. In her own words, the researcher's problem statement was as follows:

This study is designed to determine if use of an instructional design model for concept attainment in teaching the critical concepts related to fetal monitoring will make a significant difference in preparation of nursing students in this skill, compared to the traditional teaching method which exists in most schools. (Kinnick, 1989, p. 8)

The research design is not one of the designs we have specifically described in this chapter. Instead, it involves administering three different instructional methods to three treatment groups (with participants assigned randomly to groups) and then observing the effects of the treatments at two different times: once immediately after instruction and then later after

students had completed the clinical rotation portion of their nursing program. Thus, the design of the study was the following:

	Group		Time →	
	Group 1	Tx1	Obs	Obs
	Group 2	Tx2	Obs	Obs
	Group 3	Tx3	Obs	Obs

In the following pages, we present excerpts from the methodology chapter of the researcher's dissertation. Our comments and observations appear on the right-hand side.

Dissertation ANALYSIS 6

METHODOLOGY

[After an introductory paragraph outlining the chapter's contents, the author describes the sample—students enrolled in maternity nursing courses at two universities—used in the study. Then, as she begins a discussion of her procedure, she explains that the experimental treatments were based on the Tennyson-Cocchiarella concept-teaching model (1986) and presents the key elements of the model. We pick up the methodology chapter at the point where the author describes the specific treatments used for each of the three treatment groups.]

Description of the Treatment Groups

[The author first explains that, for each of the three groups, treatment consisted of instruction in the basic concepts of fetal monitoring, plus additional instructional strategies, or "teaching variables," that differed for the groups.] . . . Starting with a basic class and adding new teaching variables to each treatment group, however, did require additional time. The length of time required for teaching the three treatment groups varied between 1 and 2 hours. These timeframes were established based on the results of the survey of baccalaureate nursing schools, in which 36% of the schools responding had less than 1 hour to teach fetal monitoring theory, and 52% had 1 to 2 hours (Kinnick, 1989).

The teaching variables for the first treatment group included labels and definitions, and presentation of best examples. According to Merrill and Tennyson, these variables usually include additional information needed to aid in the clarification and understanding of the concepts (Merrill & Tennyson, 1977, p. 100). Therefore, the design of this didactic presentation began with a very basic overview of physiology at the uterofetoplacental unit. Electronic fetal monitoring patterns are a reflection of uterofetoplacental physiology. Understanding the normal physiology and changes in the physiology that cause inadequate fetal oxygenation help the learner to identify the various patterns, and whether patterns are normal or abnormal. Understanding the physiology is also the basis to identifying appropriate nursing intervention which promotes normal physiology (reduction or even elimination of fetal distress) when abnormal patterns occur.

Comments

The author points out a possible confounding variable in her study: The three forms of instruction took varying amounts of time.

The survey to which the author refers was administered during a pilot study that she conducted prior to conducting the dissertation itself. She published the pilot study as a research article, which she cites here.

In this and subsequent paragraphs the author describes the treatment used for each treatment group; in a later "Procedure" section, she describes the general procedure she used to conduct the study. More often, a researcher will include a description of how each group was treated within the procedure section itself. Either approach is acceptable, however, as long as the writer makes the organization of the methodology section clear (e.g., through headings and subheadings).

When the classes were taught, the majority of students did not have any theory about the process of labor and delivery. In addition, they had not seen a fetal monitor. Methods of monitoring the fetus and a brief description and discussion of external versus internal monitoring, therefore, needed to be discussed. In addition, it was necessary to show the students a print-out of a fetal monitor as well as explain what the graphs meant. Before the basic concepts related to interpretation of the fetal heart could be taught, the student also needed to recognize critical characteristics of a contraction pattern as seen on a monitor strip. Contraction patterns can be a cause of physiological changes at the uterofetoplacental site. After these areas had been covered, the concept label, definitions, and best examples were presented

This 1 hour presentation included labels, definitions, best examples, and clarifying information. In the experience of this researcher, this presentation reflects closely the method for teaching fetal monitoring used in most schools of nursing, especially when the allocated time for teaching this content is limited. This treatment group is referred to as Group 1 throughout the study.

The second treatment group began with the same presentation used with the first treatment group, plus the addition of expository presentations for each major concept. An expository presentation was added after the labels, definition, and best examples of each set of coordinate concepts had been completed. For example, following the definition and display of the best examples of baseline fetal heart rate and its coordinate concepts, an expository presentation was done of the coordinate concepts. When that was completed, the concept of baseline variability was introduced and the same order of teaching variables was used. The addition of the expository presentations added approximately half an hour, so that this treatment group was scheduled for one and one-half hours. This group (labels, definitions, best examples and expository presentation) is referred to as Group 2.

The design in Group 2 was chosen based on the results of Dunn's research (1984) on concept learning with college age students. . . . [The author briefly describes Dunn's findings and their relevance for the instruction presented to Group 2.]

The treatment design for the third group used the same teaching variables as in Group 2, plus the addition of an interrogatory presentation to follow each expository presentation. This involved the addition of . . . transparencies specifically developed for the interrogatory presentation. When a fetal monitor pattern was shown on the screen, students were requested to compare it with their handout of definitions (list of critical characteristics) and best examples, and to identify the concept shown on the fetal monitoring pattern. This treatment design incorporated all of the teaching variables of the Tennyson-Cocchiarella concept-teaching model.

Development of the Instruments

[In this section, the author describes the tests she used to assess what participants knew about fetal monitoring following instruction, as well as a short questionnaire she used to determine the extent to which each participant knew something about fetal monitoring before instruction.]

This description of what most students knew (and did not know) before instruction gives the reader greater confidence that the results observed after instruction (i.e., students' test performance) were probably due to the instructional treatments, rather than to any earlier learning experiences that the students may have had.

Notice that the author's notion of what is "traditional" instruction is based on her own experiences, and she says so here.

After describing Group 1, the author proceeds to descriptions of Group 2 and then Group 3 in a logical and systematic fashion. The use of three subheadings (something along the lines of Treatment for Group 1 or Group 1 Instruction) might have been helpful, however.

By "expository presentation," the author means giving a short explanation or lecture about important ideas and concepts.

A rationale for a particular experimental treatment strengthens any research report. A brief rationale can easily be incorporated into the description of procedures; a longer one should probably be presented earlier in the research report.

By "interrogatory presentation," the author means asking questions to assess students' understanding of, and ability to apply, what they have learned.

Procedure

Prior to implementing this research, approval for the project was obtained from the Human Research Committee at the University of Colorado, and the Internal Review Board for Research at the University of Northern Colorado (Appendix E). The researcher then met with all students in each maternity nursing course during their first class to explain the research and ask their consent to participate. Consent forms were provided for each student (Appendix E). Once this process was completed, the research design was implemented.

Each maternity nursing course had three groups participating in the research. Students in each of the courses were randomly assigned to one of these three groups. One group received the instructional method described in the Tennyson-Cocchiarella model of concept attainment. A second group received the same instructional method with the exception of the interrogatory presentation. The third group had a didactic presentation using only labels, definitions, best examples and clarifying information. In other words, both the expository and interrogatory presentations were eliminated from the presentation for the third group. In both schools, the researcher taught all three methods. A script (or lecture) was developed for the researcher to use in all the treatment groups so that the content was the same in each group (Appendix F). The students were tested in a class session within 2 to 3 days following the class (treatment).

After the completion of the clinical experience of all groups in each university, a parallel form of the classification test was again administered. The sequence can be summarized as follows:

Class instruction → Posttest → Clinical
Rotation → Delayed Test upon Completion of
Clinical Rotation

In addition, each student was requested to keep a record of the number of contacts each of them had with fetal monitoring tracings, the context, and type of pattern (Appendix G). For example, the student may have been assigned to a labor patient who had a normal pattern. The contact, however, could have been in clinical conference where actual monitor strips of patients were discussed, or also in a prenatal clinic where a nonstress test was done on a patient. The purpose of keeping these records [was] to identify the number of interrogatory examples the students encountered clinically and the range of examples. This information [could] be compared with the post test results.

Ideally, none of the students were to have had any contact in the clinical setting before the instruction and first test were done. However, it was impossible to schedule all three treatments before students in each maternity nursing course were assigned to the clinical setting since they began their clinical experiences the second week of classes. A few students in this situation were assigned to patients with fetal monitors attached. Since they did not have any theory on fetal monitoring, they were not responsible for interpretation of fetal monitor patterns. However, staff nurses and/or clinical instructors may have demonstrated how to attach and detach the equipment and talked about tracings seen by each student on their individual patients.

Because the author conducted the study at two universities, she followed the necessary human research review procedures at both institutions.

As noted earlier in Chapter 9, random assignment is one effective way of ruling out the possible effects of confounding variables.

The first group mentioned here ("one group") is actually Group 3, and the last ("the third group") is actually Group 1; this reversal might cause confusion for the reader.

The use of a "script" here would help the researcher teach the content similarly for all three treatment groups (except, of course, for the things she intentionally wanted to do differently for the three groups). Thus, it should help minimize any influences the researcher's hypotheses might have on her delivery of different instructional methods.

This graphic display of the procedure used is a helpful summary for the reader.

The author presumably asked students to keep such records as a way of helping her interpret any unexpected results related to the delayed (postclinical rotation) test. Keep in mind, however, that such self-reporting techniques, dependent as they are on participants' diligence and memories, will not always yield totally accurate information.

Here the author points out a potential weakness in her study: Some students had additional exposure to fetal monitoring outside of the instruction she had given them in their respective treatment groups. The exposure was apparently minimal, however, and so probably did not jeopardize the quality of her study. Such honesty is essential in any research report.

Statistical Analysis

[The author continues with a discussion of the statistical analyses she used to compare the performance of the three groups.]

NOTE: Excerpt is from *Learning Fetal Monitoring Under Three Conditions of Concept Teaching* (pp. 58–69) by V. Kinnick, 1989, unpublished doctoral dissertation, University of Colorado, Boulder. Reprinted with permission.

For Further Reading

- Antony, J. (2003). *Design of experiments for engineers and scientists*. Oxford, England: Butterworth-Heinemann/Elsevier.
- Barlow, D. H., Nock, M. K., & Hersen, M. (2008). *Single case experimental designs: Strategies for studying behavior change* (3rd ed.). Boston: Allyn & Bacon.
- Bausell, R. B. (1994). *Conducting meaningful experiments: Forty steps to becoming a scientist*. Thousand Oaks, CA: Sage.
- Canavos, G. C., & Koutrouvelis, I. A. (2009). *An introduction to the design and analysis of experiments*. Upper Saddle River, NJ: Prentice Hall/Pearson Education.
- Deming, S. N., & Morgan, S. L. (1993). *Experimental design* (2nd ed.). New York: Elsevier Science.
- Friedman, D., & Sunder, S. (1994). *Experimental methods: A primer for economists*. New York: Cambridge University Press.
- Glass, D. J. (2007). *Experimental design for biologists*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Glass, G. V. (1988). Quasi experiments: The case of interrupted time series. In R. M. Jaeger (Ed.), *Complementary methods for research in education* (pp. 445–464). Washington, DC: American Educational Research Association.
- Heath, D. (1995). *An introduction to experimental design and statistics for biology*. Bristol, PA: Taylor & Francis.
- Kirk, R. E. (1994). *Experimental design: Procedures for the behavioral sciences* (3rd ed.). Pacific Grove, CA: Brooks/Cole.
- Montgomery, D. C. (2009). *Design and analysis of experiments* (7th ed.). New York: Wiley.
- Morgan, D. L., & Morgan, R. K. (2001). Single participant research design: Bringing science to managed care. *American Psychologist*, 56, 119–127.
- Phillips, D. C. (1981). Toward an evaluation of the experiment in educational contexts. *Educational Researcher*, 10(6), 13–20.
- Pukelsheim, E. (2006). *Optimal design of experiments (Classics in applied mathematics)*. New York: Wiley.
- Quinn, G. P., & Keough, M. J. (2002). *Experimental design and data analysis for biologists*. Cambridge, England: Cambridge University Press.
- Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). *Estimating causal effects using experimental and observational designs*. Washington, DC: American Educational Research Association.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Solso, R. S., & MacLin, K. (2008). *Experimental psychology: A case approach* (8th ed.). Boston: Pearson/Allyn & Bacon.

ANSWERS TO THE CONCEPTUAL ANALYSIS EXERCISE

"Identifying Research Designs" on page 251:

1. This is a *quasi-experimental* study. In particular, it involves a *multiple baseline* design: Each of the horses begins training on a different day. In the section of the chapter "Using Designs 11, 12, and 13 in Single-Subject Studies," a multiple baseline study involving one 2-year-old girl is described. In this example, however, we see the approach being used with five horses, each of which is treated identically except for the date on which training begins.
2. This is an *experimental* study in which the researchers randomly assign participants to one of five groups, each of which is asked a different question.
3. Don't let the random selection of treatment and control groups fool you. This is a *quasi-experimental* study because the participants are not randomly assigned as *individuals* to the treatment and control groups. More specifically, the study is a *nonrandomized control group pretest–posttest* design (Design 8).
4. This, too, is a *quasi-experimental* study. It is a *time-series* design in which the effects of no intervention (Phase 1) are compared to the effects of two different interventions (the two new signs) imposed either singly or in combination. Of the designs described in this

chapter, it is probably most similar to Design 12. Note, however, that no phase of the study is repeated; this omission is a decided weakness in the design.

5. This is a *combined experimental and ex post facto factorial* design with two independent variables, one of which is a *within-subjects* variable. One independent variable is the presence or absence of attention-deficit hyperactivity disorder, which the researchers do not (and *cannot*) manipulate; this is the ex post facto component of the design. The other independent variable is the content of the drawings (aggression vs. prosocial behavior); this is the experimental, within-subjects component of the design.

Now go to MyEducationLab at www.myeducationlab.com to take assessments to evaluate your mastery of chapter concepts. Assignments and Activities exercises and Building Skills in reading Research exercises are also available to help you master concepts and skills. Feedback for many of these exercises is provided so that you can see why your answers are correct or incorrect.

10

Mixed-Methods Research

Many research problems have both qualitative and quantitative dimensions. To fully address them, then, the researcher must use both qualitative and quantitative techniques. Qualitative and quantitative methodologies are not necessarily a case of either-or, but rather a case of more-or-less.

Some research problems practically scream for both qualitative and quantitative data. These problems call for **mixed-methods research**. Such research involves not only collecting, analyzing, and interpreting both qualitative and quantitative data but also *integrating* conclusions from those data into a cohesive whole.

As you may have noticed, this chapter is a short one. It is short not because mixed-methods studies are quick and easy—they definitely are *not* quick and easy, not by a long shot—but because mixed-methods research draws largely on the qualitative and quantitative research strategies we have previously addressed in Chapters 6 through 9. The trickiest part of a mixed-methods study is in *combining* the two methodological traditions into a research endeavor in which all aspects substantially contribute to a single, greater whole. In other words, a good mixed-methods study is one that effectively and convincingly “hangs together.”

Other things being equal, a mixed-methods study is the most challenging study a researcher can undertake, because it requires both qualitative and quantitative research skills. In a typical mixed-methods investigation, the researcher must be well versed in most or all of the following skill sets (Creswell & Plano Clark, 2011):

- Identifying focused and useful research questions
- Formulating and strategically testing hypotheses
- Choosing one or more samples that enable appropriate inferences about a larger population
- Controlling for confounding variables
- Creating and using measurement instruments that have validity and reliability for their purposes
- Conducting structured, semistructured, and open-ended interviews
- Analyzing qualitative data (identifying units suitable for coding, applying the codes, discerning general themes, etc.)
- Drawing and persuasively arguing for reasonable conclusions from qualitative data (through triangulation, negative case analysis, thick description, etc.).
- Calculating and drawing inferences from descriptive and inferential statistics

Thus, if you are thinking about conducting a mixed-methods study, we urge you to do so only after you have had some experience with strictly qualitative studies *and* with strictly quantitative studies.

When Mixed-Methods Designs Are Most Useful and Appropriate

In some disciplines, mixed-methods research is becoming increasingly fashionable, trendy, “hip”—so much so that many universities now offer specific courses in mixed-methods techniques. But fashion and trendiness are hardly legitimate reasons to conduct a mixed-methods

study. Ultimately, decisions about research design must be driven by the research problem and its subproblems. Some problems and subproblems call for quantitative data, some call for qualitative data, and some—but only some—call for both.

On average, a mixed-methods research study requires more of the researcher's time and energy than a strictly qualitative or quantitative study. Why, then, would a researcher want to go to the trouble of collecting, analyzing, interpreting, and integrating both qualitative and quantitative data? Following are several good reasons (Bryman, 2006; Greene, Caracelli, & Graham, 1989):

- *Completeness*: A researcher can fully address a research problem and its subproblems only by collecting, analyzing, and interpreting both qualitative and quantitative data.
- *Complementarity*: Quantitative aspects of the study can compensate for weaknesses in qualitative research, and vice versa—for example, the results of unstructured interviews with only a small number of individuals (which might raise concerns about generalizability) can be replicated by administering a questionnaire to a larger, more representative sample.
- *Hypothesis generation and testing*: Qualitative data often provide insights that help a researcher form hypotheses about cause-and-effect relationships—hypotheses that the researcher can subsequently test through controlled, quantitative research.
- *Development of appropriate research tools and strategies*: One type of data can inform and guide subsequent collection of another type of data—for example, unstructured interviews (yielding qualitative data) can guide the construction of appropriate questions for a survey (which will yield quantitative data).
- *Triangulation*: A researcher can make a more convincing case for particular conclusions if both qualitative and quantitative data lead to those conclusions.
- *Resolution of puzzling findings*: In a quantitative study, various results can sometimes seem inconsistent or contradictory; qualitative data may reveal underlying nuances and meanings that can help the researcher make sense of the numbers.

In deciding whether to conduct mixed-methods research, the shrewd researcher also takes into account the philosophical orientations and research traditions of his or her discipline (Hesse-Biber, 2010). For example, if prominent scholars in a researcher's field agree that people's subjective "realities" and personal, idiosyncratic "truths" are legitimate topics of investigation, the researcher might reasonably incorporate phenomenological and/or grounded-theory (i.e., qualitative) elements into a research project. If, instead, the researcher's professional colleagues largely adhere to the notion that scientific inquiry should be an objective endeavor—one based solely on the "hard facts" that only numbers can reveal—the researcher might be better advised to stick with a solely quantitative methodology.

Common Mixed-Methods Designs

In previous chapters we have already touched on a few ways in which qualitative and quantitative methods might be combined into a single study. For example, historical researchers often make use of both qualitative and quantitative data in a single research inquiry (see Chapter 7). Interviews in a large-scale survey might include a few open-ended questions that yield complex responses requiring qualitative analysis (see Chapter 8). And the effectiveness of people's study strategies might be measured both quantitatively (e.g., by the number of facts they recall) and qualitatively (e.g., by the content of their think-aloud study sessions; see Chapter 9).

The ways in which a researcher might combine qualitative and quantitative methods are almost limitless, restricted only by the researcher's imagination and creativity and, of course, by the nature of the research problem. The following four general designs can help you envision some of the possibilities.¹

¹In describing these designs, we use labels suggested by Creswell and Plano Clark (2011). Other labels are possible, of course; for instance, Gay, Mills, and Airasian (2009) use the terms QUAN–QUAL, QUAL–Quan, and QUAN–Qual for convergent designs, exploratory designs, and explanatory designs, respectively, with the upper- and lowercase letters reflecting the relative emphasis the researcher might place on quantitative versus qualitative methods.

Convergent Designs. In a **convergent design**, a researcher collects both qualitative and quantitative data in parallel, usually at the same time and with respect to the same research question(s). The researcher gives equal weight to the two types of data and strives for triangulation, with the hope that analyses of both data sets lead to similar conclusions about the phenomenon under investigation.

Embedded Designs. An **embedded design** is similar to a convergent design, in that both qualitative and quantitative data are collected within the same general time frame. However, one general approach dominates—perhaps a qualitative approach, but more often a quantitative one—with the other approach serving in a secondary, supplementary role. In many cases the two data sets are collected to address somewhat different (although certainly related) research questions.

Exploratory Designs. An **exploratory design** typically encompasses two phases. In Phase 1, a researcher uses one or more qualitative methods to get a general sense of characteristics, phenomena, and/or issues related to the topic of study. The qualitative data—perhaps from observations, interviews, or both—provides a basis for a more systematic, quantitative study in Phase 2. For example, as mentioned earlier, qualitative observations of a phenomenon in the real world might help the researcher develop hypotheses to be systematically tested in an experimental study, or the results of a few unstructured interviews might help the researcher develop appropriate questions for a questionnaire administered to a much larger sample.

Explanatory Designs. Like an exploratory design, an **explanatory design** is usually a two-phase process, but in this case the quantitative phase comes first. More specifically, Phase 1 involves collecting considerable quantitative data, perhaps in an experiment, ex post facto study, or survey. However, this first phase yields only numbers, in the forms of percentages, averages, and so on. Collecting qualitative data in a Phase 2 follow-up—for instance, asking participants to describe what they were thinking or feeling during an experimental intervention or to elaborate on their answers to survey questions—can help the researcher give greater substance and meaning to the numbers.

When mixed-methods designs are lopsided in favor of one form of data over the other, they usually lean in the quantitative direction (Hesse-Biber, 2010). Unless there's a compelling reason to do otherwise, we urge novice mixed-methods researchers to lean in the quantitative direction as well. A more quantitatively oriented approach can provide a reasonable structure to guide the overall research project—a structure that can keep a researcher on task and consistently focused on addressing the research problem. As one mixed-methods expert has put it, “Undertaking a qualitative approach to mixed methods is like taking a journey without always being in control of your destination” (Hesse-Biber, 2010, p. 211).

Planning a Mixed-Methods Study

Some mixed-methods designs are, in advance, determined in precise detail from start to finish. Others have a more *emergent* quality to them, with data collected early in the process revealing other kinds of data that might also be useful (recall the concept of *emerging design* in Chapter 6). For example, early data might yield difficult-to-interpret inconsistencies or suggest potentially fruitful new avenues of investigation; in either case, a subsequent phase of data collection might be in order.

But regardless of whether a mixed-methods study has a fixed or emerging design, it requires considerable advance planning. We have discussed planning in great depth in earlier chapters (especially in Chapter 4), so here we alert you only to a few additional things to keep in mind as you plan a mixed-methods research project.

Identifying Research Questions and Hypotheses

Curious people tend to ask questions—*lots* of questions—and different questions can't necessarily all be addressed in the same way. Such multiple question-asking can give rise to mixed-methods research. Often, then, a mixed-methods study revolves around two or more research problems rather than just one.

Occasionally both quantitative and qualitative data might, on the surface, be relevant to a single research question. As an example, we return to a study described in Chapter 9, one in which three groups of college professors—geographers, sociologists, and educational psychologists—studied maps in an effort to remember the maps' contents (J. E. Ormrod, Ormrod, Wagner, & McCallin, 1988). The overarching research question was this:

Do geographers study and remember maps differently than nongeographers do?

Both quantitative data and qualitative data were potentially relevant to the question: Quantitative data could reveal whether geographers tend to remember *more* map content than nongeographers, and qualitative data could reveal whether geographers study and remember maps in distinctly different ways than nongeographers do. In a sense, then, the research question could be broken down into two subproblems:

Do geographers remember more details of a new map than nongeographers do? (This question calls for quantitative data.)

Do geographers think about a new map in different ways than nongeographers do? (This question calls for qualitative data.)

The answer to the first (quantitative) question was, yes, geographers remember more details of a new map *if* the map is consistent with well-established geographical principles of how various land forms and human constructions are arranged on the earth's surface. The answer to the second (qualitative) question was, yes, geographers think about a new map differently than nongeographers do—in particular, geographers try to make sense of the arrangement of features in the map, whereas nongeographers are more likely to engage in rote memorization. The answers to both subproblems, then, converged to answer the overall research question.

As you plan a mixed-methods study, we urge you to identify two or more separate problems—or at least two or more subproblems—to guide you in your investigation. Following are general frameworks you might consider in the *quantitative* questions you ask:

- To what extent do _____ [certain kinds of people, animals, plants, inanimate objects] exhibit _____ [certain kinds of behaviors or characteristics] _____?
- Do _____ have more/less of _____ than _____ do?
- Is there a predictable correlation between _____ and _____?
- Does _____ have an effect on _____?

Meanwhile, your *qualitative* questions might take forms such as these:

- What is the general nature of _____ [a certain group or phenomenon] _____?
- How do people think or feel about _____?
- How do participants in the sample explain _____?

We also recommend one or more questions that require an *integration* of the study's quantitative and qualitative elements—perhaps questions along these lines:

- Do the quantitative data and qualitative data converge to support the conclusion that _____?
- Can the qualitative data help to explain and elaborate on the quantitative findings?
- Do the qualitative data suggest hypotheses you might test with quantitative data?

Our last question brings up the issue of *hypotheses*. Mixed-methods studies may or may not involve the testing of certain hypotheses. In some cases, one or more hypotheses may be posed in advance, presumably as a result of a review of the related literature. In other cases—for instance, in the two-phase exploratory design previously described—qualitative data collected early in the study might yield hypotheses that the researcher subsequently tests more systematically by collecting quantitative data.

Conducting the Literature Review

As is true for strictly quantitative and strictly qualitative designs, a mixed-methods researcher should conduct much of his or her literature review at the very beginning of the project. (Grounded theory studies are occasionally exceptions to this rule; see Chapter 6.) A review of the related literature can help the researcher pin down appropriate questions and hypotheses, suggest possible research designs, and reveal potentially helpful measurement instruments.

Especially in studies with two or more phases, additional visits to the library or its online databases may be useful midway through the research project (Hesse-Biber, 2010). For example, in the first, qualitative phase of a two-phase exploratory study, you may unearth intriguing ideas about which you need to learn more so that you can better plan the second, quantitative phase. Or, in a two-phase explanatory study, your initial quantitative findings may necessitate a search for new literature that will assist you as you plan and conduct the subsequent, qualitative phase.

Addressing Validity Concerns

Whether conducting a qualitative, quantitative, or mixed-methods study, any researcher must ensure that its measurement techniques—in the most elementary sense of what *measurement* involves—are valid indicators of the variables under investigation (see Chapter 4). Here our focus is not on the validity of measurement instruments but rather on the validity of the overall research effort, including:

- Its *internal validity*—the extent to which the study enables defensible conclusions about cause-and-effect and other cross-variable relationships
- Its *external validity*—the extent to which the study's results can be generalized to a larger population or broader context
- Its general *credibility* and *trustworthiness*—the extent to which others perceive the study's findings to be convincing and worth taking seriously

All three of these ideas should look familiar, as we previously discussed each of them in Chapter 4.

Additional validity issues come up in mixed-methods research, especially in relation to how a study's quantitative and qualitative components come together—or in some cases *don't* come together. Drawing on Creswell and Plano Clark's list of potential threats to validity in mixed-methods research (2011, pp. 240–241), we urge you to consider and address questions such as the following whenever you conduct a mixed-methods study:

- Are the samples for the quantitative and qualitative aspects of the study the same or sufficiently similar to justify comparisons between the quantitative and qualitative data?
- Are the quantitative and qualitative data equally relevant to the same or related topics and research question(s)?
- Are the quantitative and qualitative data weighted equally in drawing conclusions? If not, what is your rationale for giving priority to one type of data over the other?
- Can specific statements or artifacts from the qualitative element of the study be used to support or illustrate some of the quantitative results?
- Can apparent discrepancies between the quantitative and qualitative data be reasonably resolved?

Special Ethical Considerations in Mixed-Methods Research

In conducting any mixed-methods research study, the usual ethical guidelines apply, including protection from harm, voluntary and informed consent, and participants' right to privacy regarding anything they might reveal about themselves. Furthermore, the researcher must obtain permission from the appropriate committee at his or her institution for any research involving human beings or nonhuman animals—from the internal review board (IRB) in the case of human participants, or otherwise from the institutional animal care and use committee (IACUC).

We alert you to two issues that can arise for mixed-methods studies in particular, and especially for two-phase studies. One issue comes into play in either an exploratory design or an explanatory design, in which the results of Phase 1 in some way guide the implementation of Phase 2. In this situation, the researcher may have to submit two separate proposals to the IRB or IACUC, one for each phase of the study. Alternatively, the researcher's first proposal might describe Phase 1 in detail and give a general overview of the intended procedure in Phase 2, followed by an updated, more specific (but perhaps briefer) proposal for Phase 2 (Creswell & Plano Clark, 2011). Either way, securing permission for the entire study will require more effort than would be the case for a one-shot research study. Don't let this fact discourage you: Your two-phase project may very well make a more significant contribution to your field than a shorter, more limited investigation could possibly do.

A second ethical issue arises in an explanatory design, especially if the researcher wants to use the results of the Phase 1 quantitative data to choose a subsample to interview in Phase 2 (Hesse-Biber, 2010). When participants respond in Phase 1—say, to a questionnaire about a potentially sensitive topic—they can reasonably expect that you won't be able to attach particular responses to particular people *unless you have explicitly told them otherwise during the informed consent process*. Yet in fact, you may very well want to use participants' responses in Phase 1 to identify an appropriate and informative subsample for your follow-up in Phase 2. You must plan for this eventuality before you implement Phase 1 and must, of course, describe your intentions in your IRB proposal and informed consent materials. One commonly used strategy is to give participants in Phase 1 data collection the option of participating—*voluntarily*—in a second, follow-up part of the study and providing a place to include their name and a telephone number or email address so that you can contact them.

PRACTICAL APPLICATION Deciding Whether to Use a Mixed-Methods Design

By its very nature, a mixed-methods study is more complex than either a solely qualitative or solely quantitative one. Other things being equal, it is likely to take more of the researcher's time and energy and may also require more resources. Is it, then, worth the trouble? The following checklist can help you answer this question.

CHECKLIST

Pinning Down the Logistics and Feasibility of a Mixed-Methods Study

1. Can your research problem(s) be better addressed with both quantitative and qualitative data than with only one form of data or the other? If so, explain how each kind of data will contribute to your inquiry.

_____ 2. Does your proposed project reflect a convergent, embedded, exploratory, or explanatory design? If so, which one? If not, describe the general nature and structure of your design.

_____ 3. What specific steps should you take to ensure that your proposed study has:

- Internal validity: _____
- External validity: _____
- Credibility and trustworthiness: _____

_____ 4. Given your design, how much time is the study likely to take?

_____ 5. What specific research skills do you need to collect the data? Do you currently have these skills? If not explain how you might reasonably acquire them.

_____ 6. What specific research skills do you need to analyze and interpret the data? Do you currently have these skills? If not explain how you might reasonably acquire them.

_____ 7. What special resources do you need to complete the study?

_____ 8. Do you have the time, energy, skills, and resources to carry out the study as you have designed it? If not, how might you scale down your study so that you can still address your research problem? Alternatively, how might you revise your research problem so that you can reasonably address it, given the time, skills, and resources you have?

Analyzing and Interpreting Mixed-Methods Data

The final activity in any mixed-methods study is, of course, to analyze and then *interpret* the collected data. At this point in the book, this should hardly be news, but we do want to raise two points about data analysis and interpretation that are especially relevant to mixed-methods research.

First, you must make a decision—ideally, before you collect your data—about whether you will (a) give the qualitative data and quantitative data equal weight in drawing your conclusions, or (b) give higher importance to one form of data over the other. You certainly can't disregard one set of data simply because they displease you! But with an appropriate rationale, you might make one form of data secondary and subservient to the other form of data. For example, this might be the case if you were using people's qualitative responses in follow-up interviews to shed light on their earlier (quantitative) responses to questionnaire items.

Second, you must decide how you might most effectively integrate your quantitative and qualitative findings as you interpret and report them. Creswell (2012) has suggested a couple of

possibilities that can be helpful in this regard. When conducting both a survey (quantitative) and open-ended interviews (qualitative), you might count the number of times each theme appears in the interviews and then compare the frequency counts with the quantitative analysis of survey responses. Another option is to create a two-column table—one column for qualitative findings and another for quantitative findings—that can reveal parallels and any discrepancies between the two data sets.

We have only skimmed the surface of mixed-methods research, but we have, we hope, given you food for thought about the appropriateness of mixed-methods approaches for your own research problems and questions. If you wish to delve into these approaches in greater depth, we urge you to consult one or more of the resources in the “For Further Reading” list at the end of the chapter.

A Sample Dissertation

We conclude the chapter with an example of a mixed-methods study that used a two-phase, explanatory design. Laura Lara (who is now Laura Lara-Brady) conducted the study for her doctoral dissertation in educational psychology at the University of Northern Colorado (Lara, 2009). Her focus was on factors that might influence the college success of Latina/o students, especially those with Mexican American backgrounds. Phase 1 of her study involved the administration of three questionnaires; hence, it had a descriptive, quantitative nature. Phase 2 involved in-depth interviews with a small subsample of Phase 1 participants; it made use of methods common in qualitative phenomenological studies and content analysis.

In her dissertation, Lara expresses concern about “the relatively low numbers of Latina/os attending and graduating from higher education institutions” (p. 3). She then draws on related research literature to identify four potentially important factors in Latina/o students’ college success—family, religion, other people’s support, and motivation—and she ties these factors to theories of child development and ethnic identity. We pick up the dissertation at the beginning of Chapter III, in which she repeats the four research questions she first posed in Chapter I and then describes her methodology. Although most of the excerpts we present are from Chapter III, we also include a paragraph from Chapter IV to give you an idea of how Lara effectively integrates her quantitative and qualitative results in interpreting her data.

Dissertation ANALYSIS

7

Chapter III

METHODOLOGY

The purpose of this mixed method study is to identify the perceptions of predominantly Mexican American college students regarding their academic successes and challenges, with special attention to the role of parents, views of education, meaning of success, and the importance of religion. This chapter describes the *quantitative* procedure used to answer the following questions:

1. How are parenting, education, meaning of success, and religion associated with the academic achievement of predominantly Mexican American college students with low and high GPAs?
2. How is ethnic identity associated with the academic achievement of predominantly Mexican American college students?

Comments

The word “associated” in Questions 1 and 2 implies that the author will compute correlation coefficients, which she indeed does in the quantitative phase of her study.

This chapter also describes the *qualitative* procedure used to answer the following questions:

3. How do predominantly Mexican American college students describe aspects of family, religion, meaning of success, and motivation in terms of being protective factors and risk factors in their academic achievement?
4. Are there any additional protective or risk factors related to the academic achievement of predominantly Mexican American college students?

Both quantitative and qualitative research methods were employed to study the academic achievement of predominantly Mexican American college students. Although there are numerous types of mixed-method designs, I selected the Sequential Explanatory Design (Creswell, 2003; Taskakori & Teddlie, 2003) due to its use of qualitative research to explain and interpret quantitative findings. The Sequential Explanatory Design is characterized by the collection and analysis of quantitative data followed by the collection and analysis of qualitative data. In this design, qualitative data are used to expand and provide depth of meaning to the experiences of diverse populations (Merchant & Dupuy, 1996).

[In the pages that follow, the author gives an advance organizer for the chapter, describes the participants in her study, and provides details about the three questionnaires administered in the study's quantitative part and the general protocol used in conducting follow-up qualitative interviews. Then, in a "Design/Procedure" section, she presents a rationale for using a Sequential Explanatory Design. We pick up the chapter where she begins to describe data collection.]

Quantitative Data Collection

The first phase of the study involved understanding participants' backgrounds, influences on academic achievement, and their level of Latina/o ethnic identity (exploration, resolution, affirmation). Potential participants received a link to complete the questionnaires via e-mail. After giving consent to participate in this study, participants were directed to the demographic questionnaire (Appendix A), Academic Factors Questionnaire (Appendix B), and the Ethnic Identity Scale (EIS) (Umaña-Taylor, Yazedjian, & Bámaca-Gómez, 2004) (Appendix C) online. The completion of all three questionnaires took approximately 30–40 minutes. All participants were invited to share their academic achievement stories.

To protect participants' confidentiality, questionnaire responses were separated from qualitative interview responses, and a separate online form asked participants' age, place of birth, languages spoken, as well as five demographic questions (ethnicity, generation, gender, GPA, and socioeconomic status). Consent forms and questionnaires were encrypted online and accessed only through a secure password by me, the lead researcher.

[In two subsequent paragraphs, the author describes how she statistically analyzed responses to the questionnaires.]

Notice that the research questions are slightly different—but clearly related—for the quantitative and qualitative aspects of the study.

Protective factors and risk factors are sometimes seen as moderating variables in cause-and-effect relationships (see Chapter 2). In this case, however, the author recognizes that her design does not allow for hard-and-fast conclusions about cause and effect, and so she chooses her words carefully.

Notice the use of e-mail for the quantitative phase of the study. This is quite appropriate for the population being studied here, as students at virtually all four-year colleges and universities are quite accustomed to and comfortable with using e-mail for many school-related tasks. In earlier decades, however, such a procedure might have led to considerable bias in a researcher's sample.

Notice the precautions that the author takes to ensure confidentiality and participants' right to privacy.

Qualitative Data Collection

Participants who completed the questionnaires were asked to provide their contact information for a follow-up interview. Students who gave consent to participate in a follow-up interview were selected based on gender, ethnicity, and GPA. Interviews took approximately 40–60 minutes and explored issues of family life, views of success, motivation, and support given at home that impacted the academic achievement of predominantly Mexican American college students, as well as other factors related to participants' academic success. All participants were allowed to select a pseudonym or their own name during the interview. However, none chose a pseudonym.

Epoché. In order to understand the phenomenon exactly as participants experience it, the concept of epoché became central. Epoché evolves from the Greek word "check." Originated by Husserl, the epoché is the separation or "bracketing" of the researcher's biases, prejudices, and any preconceived ideas about the phenomenon being studied (Field & Morse, 1985; Stanghellini, 2005). The epoché allowed for each participant's experience to be considered as a single entity in and of itself. This perception of the phenomenon thus calls for looking, watching, and becoming aware without importing the researcher's judgment (Moustakas, 1994). As suggested by Moustakas (1994), researchers should engage in the epoché process before conducting each interview to minimize any biases. In this study, every attempt was made to bracket any prejudices and biases of the researcher, by noting them in a journal along with any expectations prior to and subsequent to each interview. For example, before each interview I would briefly describe my expectations and other ideas I had on my mind, such as having an expectation for students in the high GPA group to be more strongly supported by their families than were students in the low GPA group. By confronting my expectations, I tried to minimize their influences as I listened to and interpreted what the participants said.

Selecting interview participants. The goal of purposeful sampling is to understand a specific phenomenon, not to represent a population, by selecting information-rich cases for research (Creswell, 2003). Studying information-rich cases yields in-depth understanding of the phenomenon that gives insight into questions under study (Patton, 2002). One strategy of purposeful sampling that captures variations between cases studied is stratified purposeful sampling. Stratified purposeful sampling illustrates characteristics of specific subgroups to facilitate comparisons by selecting participants based on key dimensions (Patton, 1990). Potential cases are then divided into "strata" containing variations of the phenomenon. In this study, "strata" to be researched were participants' GPA. Ten participants (five students in each group) were chosen for follow-up interviews based on purposeful sampling [using participants'] GPAs.

[In the next three paragraphs, under the subhead "Phenomenology," the author describes using a phenomenological approach to focus on the "lived experiences of the participants" and to "find meaning in participants' actual experiences." [See Chapter 6 of this book for a more in-depth description of the phenomenological approach to qualitative research.]

Data analysis. Interview questions were aimed at expanding the responses gathered in the quantitative portion of the study, specifically responses gathered from

Here the author briefly alludes to the sampling procedure she uses in her study's qualitative phase. Details follow a bit later, as you will see.

Epoché is an important strategy for maximizing a researcher's objectivity in a phenomenological study (see Chapter 6).

The author uses what we have, in this book, called purposive sampling; she also engages in stratified sampling (see Chapter 8 for discussions of both purposive sampling and stratified sampling). In particular, the author intentionally selects students of Mexican American heritage (both genders), including some with high GPAs and some with low GPAs.

the Academic Factor Questionnaire. Interview questions were refined after the quantitative results were gathered (for a copy of the final interview questions, please refer to Appendix D.) Interviews took approximately 40–60 minutes at a coffee shop across the street from the university. All interviews were recorded using a voice recorder and an external microphone and were later uploaded to the researcher's computer. Views of success, family life, motivation, and support given at home were explored as a basis for the interview. The focus of the interviews was to understand the contextual factors surrounding participants' academic success, as well as the differences between academically successful and non-academically successful students. All of the questions were open-ended, and I probed participants for clarification and detail. For consistency, all participants were given the same interview questions. Interviews were administered in English (although a Spanish version was available).

Qualitative data were transcribed in a personal computer using the software HyperTRANSCRIBE. This software allowed for easy transcription of MP3 audio files into a Microsoft Word document. After all interviews were collected, they were stored in the researcher's computer and transferred into HyperTRANSCRIBE for transcription into a Word document. Another researcher with expertise in educational psychology and I later conducted a content analysis of the interviews for triangulation (Campbell & Fiske, 1959; Creswell, 1998. . .). The content analysis included both coding and theme analysis (Moustakas, 1994). Coding of the data consisted of looking at the content of the responses elicited by participants and arranging them with a color scheme in terms of frequency/repetition and theme. A thematic analysis followed, and it analyzed all components in participants' interviews to form a comprehensive picture of a collective experience. As a continuation of the content analysis, the second educational psychology researcher and I separately conducted a thematic exploration to ensure uniformity and validity of the results. Once patterns were established, we compared results and developed all patterns into themes to finally compare them with the quantitative analysis.

[In the remaining two paragraphs of the chapter, the author continues the discussion of triangulating the results and then describes strategies used to maintain participants' confidentiality and right to privacy.]

[We now jump way ahead to a section called "Interpretation of Final Analysis" in Chapter IV ("Results") in the author's dissertation. We include this paragraph to give you a flavor of how she pulled the quantitative and qualitative results into a single, integrated discussion.]

[S]tudents in the low and high GPA groups showed distinct views of success, family, religion, and motivation. Three of the strongest differences were view of education, view of religion, and academic preparation. Students in the high GPA group viewed education as a privilege, and as something that had to be earned. One student in the high GPA group mentioned, "I don't get grades, I earn grades. I mean you just don't . . . you know what I mean? Teachers don't just hand things out. You get what you deserve." Students in the high GPA group also saw college as the next step in their lives: "I think

Notice the informal setting for the qualitative interviews. Such a setting might be more comfortable for participants, setting them at ease and enabling them to be more candid and expansive in their responses.

Notice the Spanish option here—presumably another strategy for making participants feel comfortable and at ease.

The author makes good use of technology to facilitate interview transcription and content analysis.

it's just one of the next steps as far as being able to do something that I'm passionate about." One student in the low GPA group indicated, "Well I'm . . . I know it's cocky, and whatever, but I'm smarter than the good chunk of the students in my class, but because I don't put any effort into it, no one ever sees it." Students in the high GPA group viewed grades as the result of personal effort, while students in the low GPA group described seeing themselves as being already smart and not needing to work hard to prove it.

NOTE: Excerpt is from *A Mixed Method Study of Factors Associated with the Academic Achievement of Latina/o College Students from Predominantly Mexican American Backgrounds: A Strengths-Based Approach* (pp. 73–74, 86–89, 91–93, 145) by L. G. Lara, 2009, doctoral dissertation, University of Northern Colorado, Greeley. (available through the online database ProQuest Dissertations & Theses: Full text; publication number 3397099). Reprinted with permission.

Notice how participants' qualitative responses give depth and meaning to the more abstract idea that perception of personal effort might be an important factor in high-GPA students' academic success.

For Further Reading

Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research*, 6(1), 97–113.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.

Greene, J. C. (2007). *Mixed methods in social inquiry*. San Francisco: Jossey-Bass.

Hesse-Biber, S. N. (2010). *Mixed methods research: Merging theory with practice*. New York: Guilford.

Onwuegbuzie, A. J., & Johnson, R. B. (2006). The validity issue in mixed research. *Research in the Schools*, 13(1), 48–63.

Tashakkori, A., & Teddlie, C. (Eds.). (2010). *Handbook of mixed methods in social and behavioral research* (2nd ed.). Thousand Oaks, CA: Sage.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks, CA: Sage.

Now go to MyEducationLab at www.myeducationlab.com to take assessments to evaluate your mastery of chapter concepts. Assignments and Activities exercises and Building Skills in Reading Research exercises are also available to help you master concepts and skills. Feedback for many of these exercises is provided so that you can see why your answers are correct or incorrect.

11

Strategies for Analyzing Quantitative Data

Numbers are meaningless unless we analyze and interpret them in order to reveal the truth that lies beneath them. With statistics, we can summarize large numerical data sets, make predictions about future trends, and determine when different experimental treatments have led to significantly different outcomes. Thus, statistical procedures are among the most powerful tools in the researcher's toolbox.

In quantitative research, we try to make better sense of the world through the numbers we obtain. Sometimes these numbers represent aspects of the observable, physical world, such as the pull of gravity on a concrete object, the temperature of a gas, or the number of people engaging in a particular activity. We may also use numbers to represent nonphysical phenomena, such as how much students learn in the classroom, what beliefs people have about controversial topics, or how much influence various news media are perceived to have. We can then summarize and interpret the numbers by using statistics. In general, we can think of statistics as a group of computational procedures that enable us to find patterns and meaning in numerical data.

To some beginning researchers, the field of statistics can appear to be a never-never land in which advanced mathematicians conjure up elusive, hard-to-grasp numerical entities. But in reality, statistics are invaluable and often indispensable tools in research. They provide a means through which numerical data can be made more meaningful, so that the researcher can see their nature and better understand their interrelationships. The first and last question of statistics is precisely the same question that every researcher needs to ask: What do the data mean? In other words, What message do they communicate?

Exploring and Organizing a Data Set

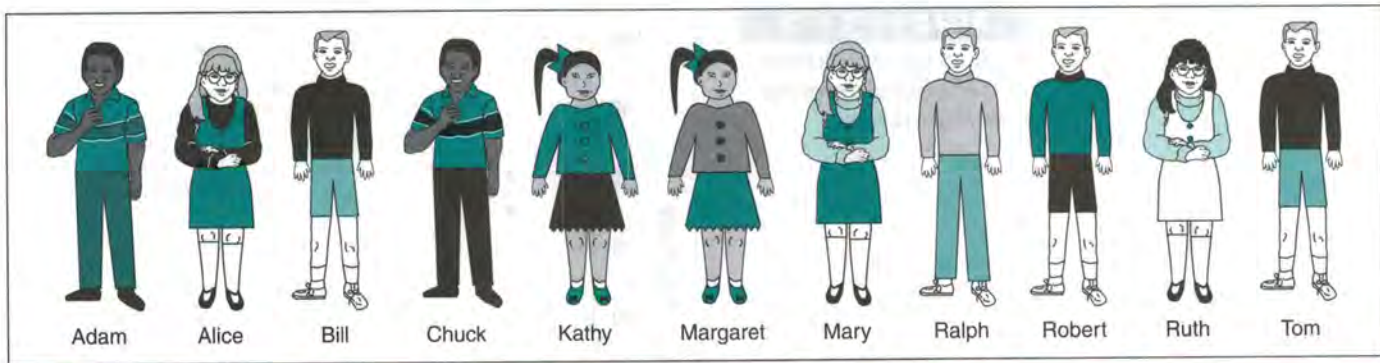
Before employing any statistical procedure—before making a single computation—look closely at your data and explore various ways of organizing them. Using an open mind and your imagination, look for patterns in the numbers. Nothing takes the place of looking carefully, inquiringly, critically—perhaps even naively—at the data.

We take a simple example to illustrate the point. Following are the scores on a reading achievement test for 11 children: Ruth, 96; Robert, 60; Chuck, 68; Margaret, 88; Tom, 56; Mary, 92; Ralph, 64; Bill, 72; Alice, 80; Adam, 76; Kathy, 84. What do you see? Jot down a few observations before you read further.

Now let's try various arrangements of the scores to see what patterns they might reveal. Some of the information may be irrelevant to our research problem. No matter. Careful researchers discover everything possible about their data, whether the information is immediately useful or not.

We begin by forming an alphabetical list of the children's names and their test scores:

Adam	76	Mary	92
Alice	80	Ralph	64
Bill	72	Robert	60
Chuck	68	Ruth	96
Kathy	84	Tom	56
Margaret	88		

**FIGURE 11.1**

The 11 children in our reading achievement test sample

When we display the children's test scores in this manner, the scores are no more meaningful, but we have at least isolated individuals and scores so that we can inspect them more easily. Does this arrangement show us anything? Yes. It shows that the highest score was earned by a girl and that the lowest score was earned by a boy. Silly, you say, and meaningless. Perhaps. But it's an observable fact, and it might come in handy later on.

Let's keep the arrangement but view it in another way. In Figure 11.1, we see these 11 boys and girls lined up in a row, still arranged in alphabetical order according to first names. Look! Now we can discern a symmetrical pattern that wasn't previously apparent. No matter whether we start from the left or from the right, we have *one* boy, then *one* girl, then *two* boys, *three* girls, *two* boys, *one* girl, and *one* boy. Putting adjacent children of the same sex together, the arrangement is this:

♂ ♀ ♂♂ ♀♀♀ ♂♂ ♀ ♂

or:

1 1 2 3 2 1 1

Now let's arrange the data differently, separating girls from boys:

Girls		Boys	
Alice	80	Adam	76
Kathy	84	Bill	72
Margaret	88	Chuck	68
Mary	92	Ralph	64
Ruth	96	Robert	60
		Tom	56

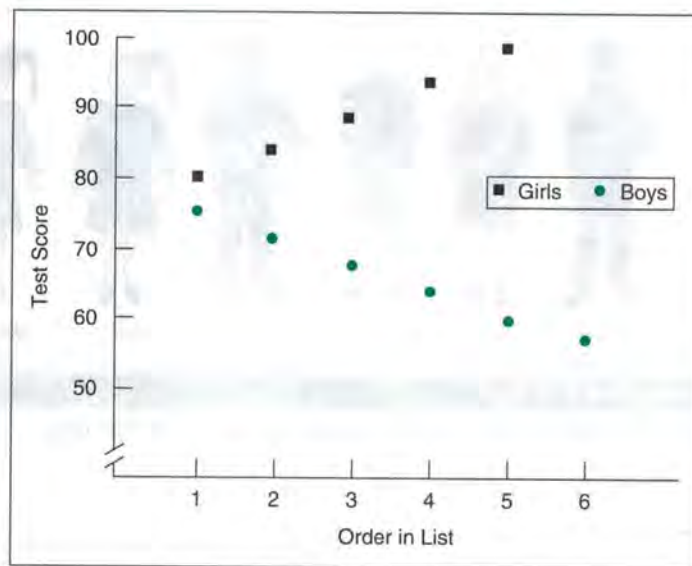
Represented graphically in Figure 11.2, the trends are quite dramatic: The girls' scores increase as we proceed through the alphabet, and the boys' scores decrease.

Not only are there opposite trends, but now we are aware of a very obvious fact that may, up to this point, have escaped our attention: The scores are equidistant from one another. Each score is 4 points either above or below the preceding one.

Whatever we have observed may have no relevance whatsoever for our project, but because it represents *dynamics within the data*, it is important that we see it. That is the point: The researcher should be aware of the dynamics—the phenomena—that are active within the data, whether those phenomena are important to the purpose of the research or not. The astute researcher overlooks nothing.

FIGURE 11.2

A visual representation of the reading achievement test scores



The preceding exercise was, of course, an artificial one. We would be hard pressed to find much meaning in diverging trends for girls versus boys that appear simply through an alphabetical arrangement of first names. Yet for the researcher working in an area of science, observations of a similar kind may reveal important new knowledge. Take the case of a paleontologist and an astronomer who examined growth marks on the spiral-shaped shells of a particular marine mollusk, the chambered nautilus (Kahn & Pompea, 1978). They noticed that each chamber in a shell had an average of 30 growth lines and deduced that (a) the growth lines had appeared at the rate of 1 per day and (b) one chamber had been laid down every lunar month, specifically every 29.53 days. They also concluded that, if their interpretation of the data was correct, it might be possible to determine from fossil shells the length of the ancient lunar months. Because the distance of the moon from Earth can be calculated from the length of the lunar month, the scientists examined nautilus fossils—some of them 420 million years old—and noticed a gradual decrease in the number of growth lines in each chamber as the fossils came from further and further back in prehistoric time. This finding indicated that the moon was once closer to Earth and revolved around it more rapidly than it does now—an observation consistent with generally accepted scientific theory.

In the two examples just presented, we find a fundamental principle about data exploration: *How the researcher prepares the data for inspection or interpretation will affect the meaning that those data reveal. Therefore, every researcher should be able to provide a clear, logical rationale for the procedure used to arrange and organize the data.* We had no rationale whatsoever for arranging the data according to the children's first names. Had we used their last names—which would have been equally illogical—we would still have seen that the girls had higher scores than the boys, but we would not necessarily have seen the diverging trends depicted in Figure 11.2.

In research questions regarding the physical world, the method for organizing data is apt to be fairly straightforward. Data often come to the scientist prepackaged and prearranged. The sequence of growth rings on a nautilus shell is already there, obvious and nondebatable. But in other disciplines—for instance, in the social sciences, humanities, and education—a researcher may need to give considerable thought to the issue of how best to organize the data.

Organizing Data to Make Them Easier to Think about and Interpret

As we mentioned in Chapter 1, the human mind can think about only so much information at one time. A data set of, say, 5,000 tidbits of information is well beyond a human being's mental capacity to consider simultaneously. In fact, unless the researcher has obtained *very* few pieces of

data (perhaps only seven or eight numbers), he or she will want to organize them in one or more ways to make them easier to inspect and think about.

In the preceding example of 11 children and their reading achievement test scores, we experimented with several simple organizational schemes in an effort to find patterns in the data. Let's take another everyday example. Joe is in high school. In February he gets the following quiz grades: 92, 69, 91, 70, 90, 89, 72, 87, 73, 86, 85, 75, 84, 76, 83, 83, 77, 81, 78, 79. Here Joe's grades are listed in a *simple linear sequence*—the order in which Joe earned them. These are the raw numerical facts—the data—derived directly from the situation. As they appear in the preceding list, they don't say very much, except that Joe's performance seems to be inconsistent.

Let's put Joe's grades in a *two-dimensional table* organized by weeks and days:

Grade Record for February					
	Monday	Tuesday	Wednesday	Thursday	Friday
First week	92	69	91	70	90
Second week	89	72	87	73	86
Third week	85	75	84	76	83
Fourth week	83	77	81	78	79

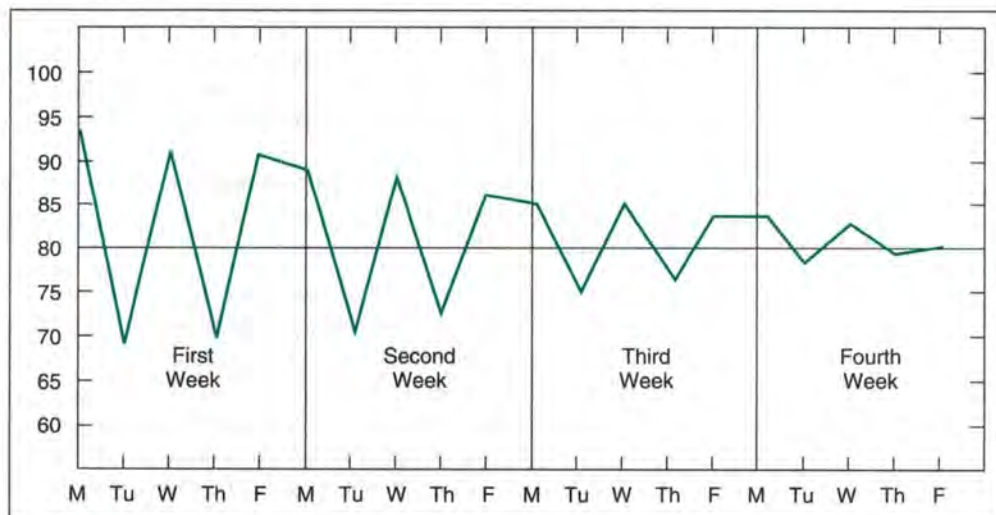
The table reveals some patterns in Joe's grades. If we compare the five columns, we quickly notice that the grades on Mondays, Wednesdays, and Fridays are considerably higher than those on Tuesdays and Thursdays. And if we look at successive numbers in each column, we see that the grades get progressively worse on Mondays, Wednesdays, and Fridays, but progressively better on Tuesdays and Thursdays.

Now let's represent Joe's grades in the form of a *simple line graph*, shown in Figure 11.3. In this graph, we see phenomena that were not readily apparent in the two-dimensional table. It's hard to miss the considerable variability in grades during the first and second weeks, followed by a gradual leveling-out process in the latter part of the month. A profile of this sort should prompt the alert researcher to explore the data further in an attempt to explain the pattern the graph reveals.

Graphing data is often quite useful for revealing patterns in a data set. For example, let's return to a study first described in a Practical Application exercise near the end of Chapter 9:

FIGURE 11.3

Line graph of Joe's daily grades



Two researchers want to see if a particular training program is effective in teaching horses to enter a horse trailer without misbehaving in the process—that is, without rearing, trying to turn around, or in some other way resisting entry into the trailer. Five horses (Red, Penny, Shadow, Sammy, and Fancy) go through the training, with each horse beginning training on a different day. For each horse, an observer counts the number of misbehaviors every day prior to and during training, with data being collected for a time span of at least 45 days. (Ferguson & Rosales-Ruiz, 2001)

In Chapter 9 we were concerned only with the design of this study, concluding that it was a quasi-experimental (and more specifically, a multiple baseline) study. But now let's look at the results of the study. When the researchers plotted the numbers of five different misbehaviors for each horse before and during training, they constructed the graph presented in Figure 11.4. Was the training effective? Absolutely yes! Once training began, Penny had one really bad day plus another day in which she turned a couple of times, and Shadow and Fancy each tossed their heads during one of their loading sessions. Aside from these four occasions, the horses behaved perfectly throughout the lengthy training period, despite the fact that all five had been quite ornery prior to training. These data have what we might call a *hit-you-between-the-eyes* quality: We don't need a fancy statistical analysis to tell us that the training was effective.

Time-series studies often yield data that show clear hit-you-between-the-eyes patterns; for another example, return to Figure 9.3 on page 241. But generally speaking, simply organizing the data in various ways will not, in and of itself, reveal everything the data have to offer. Instead, a quantitative researcher will need to perform statistical analyses to fully discover the patterns and meanings the data hold. Before we turn to the nature of statistics, however, let's briefly look at how a researcher can use computer software to assist with the data organization process.

Using Computer Spreadsheets to Organize and Analyze Data



The process of organizing large amounts of data was once a cumbersome, time-consuming, and tedious task. Fortunately, the advent of computers has made the process much simpler and more efficient. One important tool is an **electronic spreadsheet**, a software program that allows a researcher to enter and then manipulate data in a two-dimensional table. Undoubtedly the best known spreadsheet software is Microsoft's Excel, but other software packages are available as well, including "freeware" you can download without charge from the Internet (e.g., Sphymic Software Spreadsheet, Simple Spreadsheet, Spread32).

The beauty of electronic spreadsheets is that once you enter data into them, the software can quickly and easily help you organize the data and make simple calculations. For example, you can add several test scores together to create a new column that you might call "Total of Test Scores," or you might divide the numbers in one column by the numbers in another column to get proportions that are potentially meaningful in the context of your study. If you change a data point—for example, perhaps you discover that you miskeyed a test score and so must correct it—all of the relevant calculations are automatically updated. The software typically also lets you copy (or *import*) data from databases, word processing documents, or other spreadsheets into a new spreadsheet.

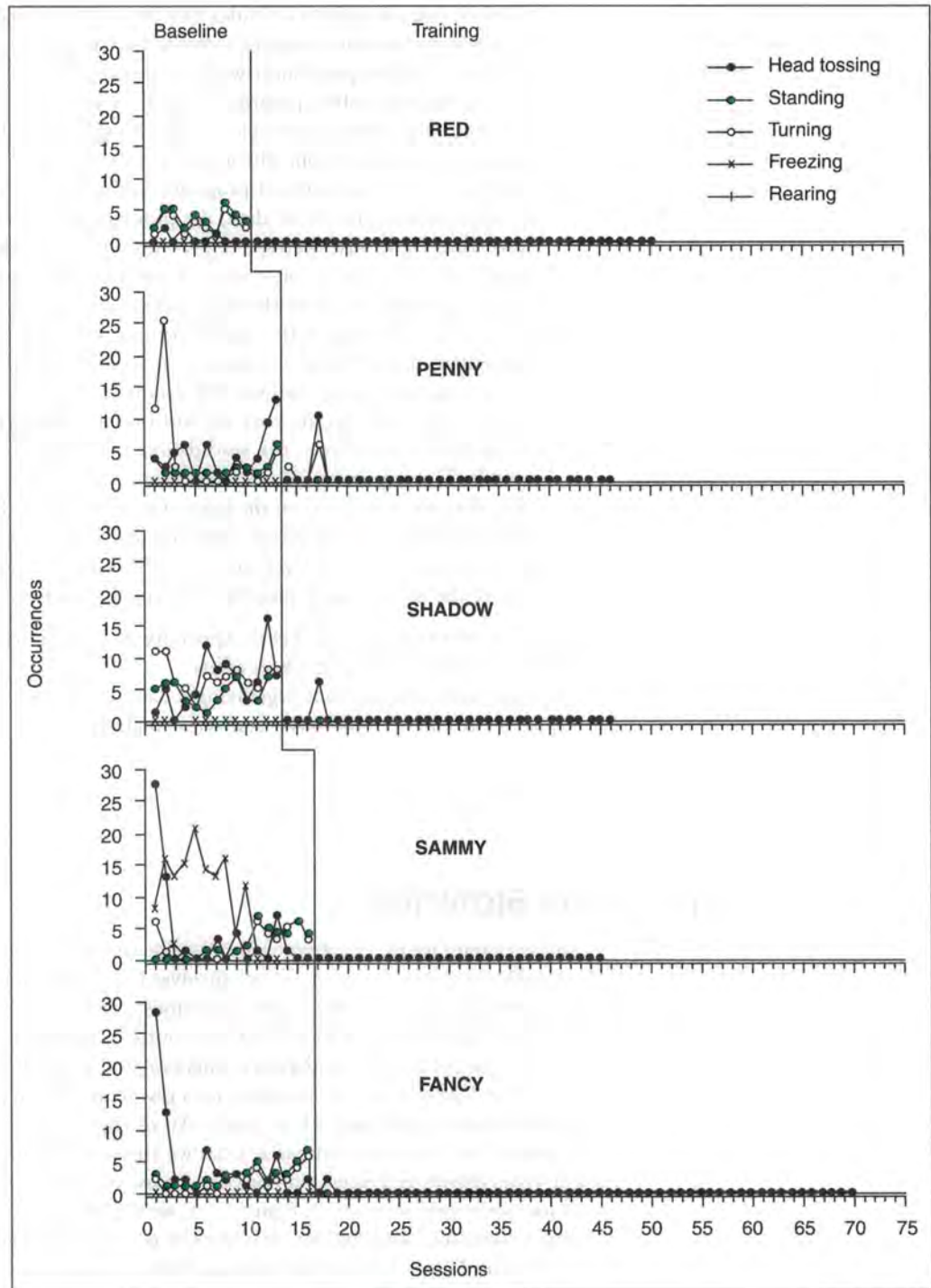
Spreadsheets would be useful to researchers even if they were capable only of listing data and adding up different columns and rows. But in fact, they allow the researcher to do many other things as well:

- **Sorting.** Once the data have been organized into rows and columns, it is possible to reorganize them in any way you wish. For example, suppose you have math test scores for a large number of children of various ages. You originally entered the scores in the order in which you obtained them. But now you decide that you want to consider them on the basis of the children's ages. In a matter of seconds, an electronic spreadsheet can sort the scores by age and list them from youngest to oldest child, or vice versa.
- **Recoding.** A spreadsheet typically allows you to make a new column that reflects a transformation of data in an existing column. For instance, imagine that you have

FIGURE 11.4

Undesirable trailer-loading behaviors for five horses before and after training

Reprinted from "Loading the Problem Loader: The Effects of Target Training and Shaping on Trailer-Loading Behavior of Horses" by D. L. Ferguson & J. Rosalez-Ruiz, 2001, *Journal of Applied Behavior Analysis*, 34, p. 419. Reprinted with permission of the Society for the Experimental Analysis of Behavior, Inc.



reading scores for children from ages 7 to 15. Perhaps you want to compare the scores for children in three different age groups: Group 1 will consist of children who are 7 to 9 years old, Group 2 will include 10- to 12-year-olds, and Group 3 will include 13- to 15-year-olds. You can tell the computer to form a new column called "Group" and to give each child a group number (1, 2, or 3) depending on the child's age.

■ **Formulas.** Current spreadsheet programs have the capability to calculate many complex mathematical and statistical formulas. Once the data are organized into rows and

columns, you can specify formulas that describe and analyze one or more groups of data. For example, you can enter the formula for computing the average, or mean, of a set of numbers, and the spreadsheet will perform the necessary calculations. Many commonly used formulas are often preprogrammed, so you merely select the statistic or function you need (e.g., you might select “AVERAGE”) and highlight the data you wish to include in the calculation. The software does the rest.

- **Graphing.** Most spreadsheet programs have graphing capabilities. After you highlight the appropriate parts of the data, the program will automatically produce a graph from those data. Generally, the type of graph produced is selected from several options (e.g., line graphs, bar graphs, pie charts). Users can select how the axes are labeled, how the legend is created, and how the data points are depicted.
- **“What Ifs.”** Thanks to the speed and ease with which an electronic spreadsheet can manipulate and perform calculations on large bodies of data, you can engage in numerous trial-and-error explorations. For example, if you are examining data for a sample of 5,000 people and decide that an additional comparison between certain subgroups might prove interesting, the spreadsheet can complete the comparison in a matter of seconds. This capability allows you to continually ask *what if . . . ?*—for instance, What if the data were analyzed on the basis of gender, rather than on the basis of age? or What if results from administering only one level of a specific medication were analyzed instead of grouping all levels together? This *what-if* capability allows the researcher to explore the data in many possible ways quickly and easily.

In the discussion of Microsoft Excel in Appendix A, you can learn how to use some of the many features that an electronic spreadsheet offers.

We have said enough about organizing a data set. We now turn to one of the most important tools in a researcher’s toolbox—statistical analysis.

Choosing Appropriate Statistics

In a single chapter we cannot thoroughly describe the wide variety of statistical procedures available to researchers. Here we must limit ourselves to a description of basic statistical concepts and principles and a brief overview of some of the most commonly used procedures. We are assuming that you have taken or will take at least one course in statistics—better still, take two, three, or even more!—to get a firm foundation in this essential research tool.

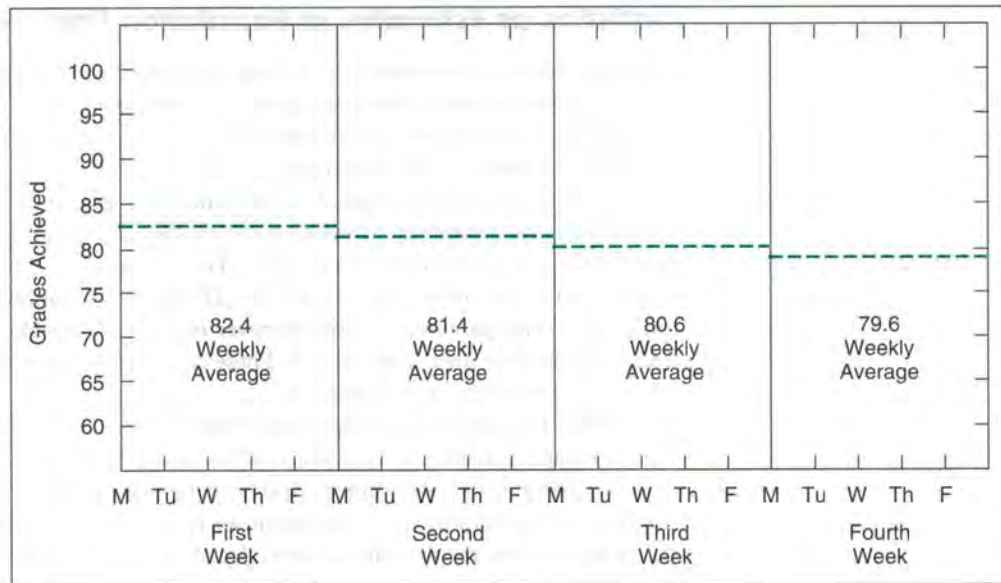
In a preceding section, we looked at Joe’s test scores in three ways: a simple linear sequence, a two-dimensional table, and a line graph. All of these depicted Joe’s day-to-day performance. Now, instead, let’s begin to summarize what we are seeing in the test scores. We can, for example, use a statistic known as a *mean*—in everyday terms, an *average*—to take out the jagged irregularities of Joe’s daily performance. In Figure 11.5, we represent Joe’s average scores for the four weeks of February with four broken lines. When we do this, we get an entirely new view of Joe’s achievement. Whereas Figure 11.3 showed only an erratic zigzagging between daily extremes, with the zigzags becoming less extreme as the weeks went by, the dotted lines in Figure 11.5 show that, week by week, very little change actually occurred in Joe’s average test performance.

Yet it may be that we also want to summarize how much Joe’s grades *vary* each week. The means presented in Figure 11.5 tell us nothing about how consistent or inconsistent Joe’s grades are in any given week. We would need a different statistic—perhaps a *range* or a *standard deviation*—to summarize the variability we see each week. (We’ll describe the nature of such measures of variability shortly.)

Thus far, we have discovered an important point: *Looking at data in only one way yields an incomplete view of those data and thus provides only a portion of the meaning those data hold.* For that reason, we have many statistical techniques, each of which is suitable for a different purpose.

FIGURE 11.5

Line graph of Joe's weekly average grades



Each technique extracts a somewhat different meaning from a particular set of data. Every time you apply a new statistical treatment to your data, you derive new insights and see more clearly and completely the data's underlying dynamics.

In the next few pages, we consider two general functions that statistics can serve. We also discuss the various ways in which the nature of the data may dictate the particular statistical procedures that can be used.

Functions of Statistics

Statistics have two major functions. Some statistics describe what the data look like—where their center or midpoint is, how broadly they are spread, how closely two or more variables within the data are intercorrelated, and so on. Such statistics are, appropriately, called **descriptive statistics**.

Other statistics, known as **inferential statistics**, serve a different purpose: They allow us to draw *inferences* about large populations by collecting data on relatively small samples. For example, imagine that you are an immigration officer. Although you have never been to Egypt, you have met numerous Egyptians as they debark from incoming planes and ships. Perhaps you have even become well acquainted with a small number of Egyptians. From this small sample of the Egyptian population, you might infer what Egyptian people in general are like. (Your inferences may or may not be accurate because your sample, which consists entirely of visitors and immigrants to your own country, is not necessarily representative of the entire population of Egypt. However, that is a sampling problem, not a statistical one.)

More generally, inferential statistics involve using a small sample of a population and then *estimating* the characteristics of the larger population from which the sample has been drawn. For instance, we might estimate a population mean from the mean we obtain for a sample. Or we might determine whether two or more groups of people are actually different, given the differences we observe between samples taken from each of those groups. Inferential statistics provide a way of helping us make reasonable guesses about a large, unknown population by examining a small sample that *is* known. In the process, they also allow us to test hypotheses regarding what is true for that large population.

Statistics as Estimates of Population Parameters

Especially when we use statistics to draw inferences about a population from which a research sample has been drawn, we are using them as *estimates of population parameters*. A **parameter** is a characteristic or quality of a population that, in *concept*, is a constant; however, its *value* is variable.

As an illustration, let's consider a circle. One of the parameters of a circle is its radius. In concept, the radius is a constant: It is the same for every circle—the distance from the center of the circle to the perimeter. In value, it varies, depending on the size of the circle. Large circles have long radii; small circles, short radii. The value—that is, the length of the radius in linear units (centimeters, feet, etc.)—is variable. Thinking of a parameter in this way, we see that each circle has several parameters: The diameter is always twice the radius (r), the circumference is always $2\pi r$, and the area is always πr^2 . These concepts are constants, even though their particular values vary from one circle to the next.

Within the context of quantitative data analysis, a parameter is a particular characteristic (e.g., a mean or standard deviation) of the entire population—which is sometimes called a *universe*—about which we want to draw conclusions. In most cases, we can study only a small sample of a population. Any calculation we perform for the sample rather than the population (the sample mean, the sample standard deviation, etc.) is called a **statistic**. Statisticians distinguish between population parameters and sample statistics by using different symbols for each. Table 11.1 presents a few commonly used symbols in statistical notation.

Considering the Nature of the Data

As you begin to think about the statistical procedures that might be most appropriate for your research problem, keep in mind that different statistics are suitable for different kinds of data. In particular, you should consider whether your data

- Have been collected for a single group or, instead, for two or more groups
- Involve continuous or discrete variables
- Represent nominal, ordinal, interval, or ratio scales
- Reflect a normal or non-normal distribution

After we look at each of these distinctions, we will relate them to another distinction—that between parametric and nonparametric statistics.

Single-Group versus Multi-Group Data

In some cases, a research project yields data about a single group of people, objects, or events. In other cases, it may yield parallel sets of data about two or more groups. Analyzing characteristics of a single group will often require different statistical techniques than those for making comparisons among two or more groups.

TABLE 11.1

Conventional statistical notation for population parameters and sample statistics

The Factor in Question	The Symbol Used to Designate the Factor	
	Population Parameter	Sample Statistic
The mean	μ	$\bar{M}$ or X
The standard deviation	σ	s or SD
Proportion or probability	P	p
Number or total	N	n

Note: The symbol μ is the lowercase form of the Greek letter *mu*. The symbol σ is the lowercase form of the Greek letter *sigma*.

Continuous versus Discrete Variables

In Chapter 2 we define a *variable* as a quality or characteristic in a research investigation that has two or more possible values. Simply put, a variable *varies*. However, it may vary in different ways. A **continuous variable** reflects an infinite number of possible values falling along a particular continuum. A simple example is chronological age. The participants in a research study can be an infinite number of possible ages. Some might be 2 years old, others might be 92, and we might have virtually any age (including fractions of years) in between. Even if the study is limited to a small age range—say, 2- to 4-year-old children—we might have children who are exactly 2 years old, children who are 2 years and 1 month old, children who are 2 years and 2 months old, and so on. We could, in theory, be even more precise, perhaps specifying participants' ages in days, hours, minutes, seconds, or even fractions of a second.

In contrast, a **discrete variable** has a finite and small number of possible values. A simple example is a student's high school grade level. At a four-year high school, a student can be in only one of four grades: 9th, 10th, 11th, or 12th. At most high schools, it isn't possible to be in anything else. One cannot be somewhere between two grade levels, such as in the "9.25th grade."

Nominal, Ordinal, Interval, and Ratio Data

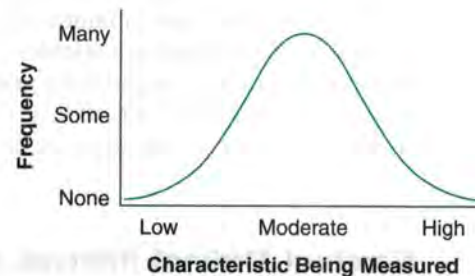
In Chapter 4 we describe four different scales of measurement; these scales, in turn, dictate how we can statistically analyze the numbers we obtain relative to one another. To refresh your memory, we briefly describe each of the scales again.

- **Nominal data** are those for which numbers are used only to identify different categories of people, objects, or other entities; they do not reflect a particular quantity or degree of something. For instance, a researcher might code all males in a data set as 1 and all females as 2. The researcher might also code political affiliation with numbers, perhaps using 1 for Republicans, 2 for Democrats, 3 for Independents, and so on. In neither case do the numbers indicate that participants have more or less of something; girls don't have more "sex" than boys, and Independents don't have more "political affiliation" than Republicans or Democrats.
- **Ordinal data** are those for which the assigned numbers reflect an order or sequence. They tell us the degree to which people, objects, or other entities have a certain quality or characteristic (a variable) of interest. They do not, however, tell us anything about how great the differences are between the people, objects, or other entities. For example, in a group of graduating high school seniors, each student might have a class rank that reflects his or her relative academic standing in the group: A class rank of 1 indicates the highest grade point average (GPA), a rank of 2 indicates the second highest GPA, and so on. These numbers tell us which students surpassed others in terms of GPA, but it does not tell us precisely how similar or different the GPAs of any two students in the sequence are.
- **Interval data** reflect equal units of measurement. As is true for ordinal data, the numbers reflect differences in degree or amount. But in addition, differences between the numbers tell us *how much difference* exists in the characteristic being measured. As an example, scores on intelligence tests (IQ scores) are, because of the way in which they are derived, assumed to reflect an interval scale. Thus, if we take four IQ scores at equal intervals—for instance, 85, 95, 105, and 115—we can assume that the 10-point difference between each pair reflects equivalent differences in intelligence between the people who have obtained those scores. The one limitation of interval data is that a value of zero (0) does *not* necessarily reflect a complete lack of the characteristic being measured. For example, it is sometimes possible to get an IQ score of 0, but such a score does not mean that a person has no intelligence whatsoever.
- **Ratio data** are similar to interval data but have an additional feature: a true zero point. Not only do the numbers reflect equal intervals between values for the characteristic being measured, but in addition a value of 0 tells us that there is a complete absence of that characteristic. An example would be income level: People with an annual income of

\$30,000 make \$10,000 more than people with an annual income of \$20,000, and people with an annual income of \$40,000 make \$10,000 more than people with an annual income of \$30,000. Furthermore, people who make \$0 a year have *no* income.

Normal and Non-Normal Distributions

Numerous theorists have proposed that many characteristics of living populations (e.g., populations of maple trees, platypuses, human beings, or a certain subgroup of human beings) reflect a particular pattern, one that looks like this:

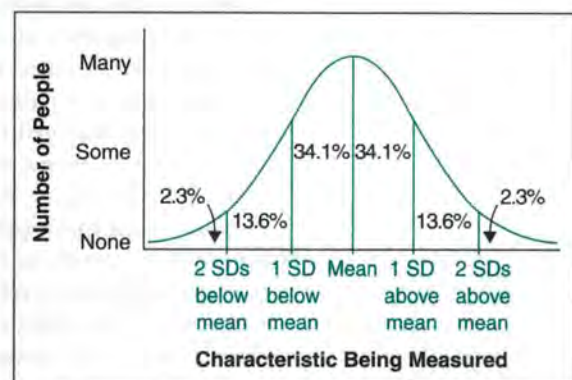


This pattern, commonly called the **normal distribution** or **normal curve**—you may also see the term *bell curve*—has several distinguishing characteristics:

- *It is horizontally symmetrical.* One side is the mirror image of the other.
- *Its highest point is at its midpoint.* More people (or whatever other entities are the focus of investigation) are located at the midpoint than at any other point along the curve. In statistical terms, three widely used measures of central tendency—the mode, the median, and the mean (all to be described shortly)—are equivalent.
- *Predictable percentages of the population lie within any given portion of the curve.* If we divide the curve according to its standard deviation (also to be described shortly), we know that certain percentages of the population lie within each portion. In particular, approximately 34.1% of the population lies between the mean and one standard deviation below the mean, and another 34.1% lies between the mean and one standard deviation above the mean. Approximately 13.6% of the population lies between one and two standard deviations below the mean, with another 13.6% lying between one and two standard deviations above the mean. The remaining 4.6% lies two or more standard deviations away from the mean, with 2.3% at each end of the distribution. This pattern is shown in Figure 11.6. The proportions of the population lying within any particular section of the normal distribution can be found in most introductory statistics books. You can also find them online by using the key words “normal distribution table” in a search engine such as Google or Yahoo!

FIGURE 11.6

Percentages within each portion of the normal distribution



To better understand the normal distribution, take any fortuitous happening and analyze its distribution pattern. For example, let's take the corn production of Iowa farms during a single year. If we could survey the per-acre yield of every farmer in Iowa—the total population, or universe, of the cornfields and corn farmers in Iowa—we would probably find that a few farmers had an unusually poor yield of corn per acre for no discernible reason except that “that’s the way it happened.” A few other farmers, for an equally unknown reason, likely had unusually heavy yields from their fields. However, most farmers would have had a middle-of-the-road yield, sloping gradually in either direction toward the greater-yield or the lesser-yield direction. The normal curve would describe the Iowa corn production. No one planned it that way; it is simply how nature behaves.

Watch an approaching thunderstorm. An occasional flash of lightning heralds the coming of the storm. Soon the flashes occur more frequently. At the height of the storm, the number of flashes per minute reach a peak. Gradually, with the passing of the storm, the number of flashes subsides. The normal curve is at work once again.

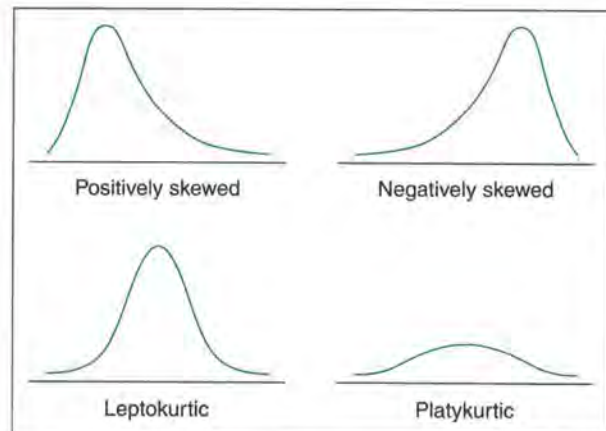
We could think of thousands of situations, only to find that nature often behaves in a way consistent with the normal distribution. The curve is a constant; it is always bell-shaped. In any one situation, the *values* within it vary. The mean is not always the same number, and the overall shape may be more broadly spread or more compressed, depending on the situation.

Sometimes, however, a variable doesn’t fall in a normal distribution. For instance, its distribution might be lopsided, or skewed. If the peak lies to the left of midpoint, the distribution is **positively skewed**; if the peak lies to the right of midpoint, the distribution is **negatively skewed**. Or perhaps a distribution is unusually pointy or flat, such that the percentages within each portion of the distribution are notably different from those depicted in Figure 11.6. Here we are talking about **kurtosis**, with an unusually peaked, or pointy, distribution reflecting a **leptokurtic curve** and an unusually flat one being a **platykurtic curve** (see Figure 11.7).

Of course, some data sets don’t resemble a normal distribution, not even a lopsided, pointy, or overly flattened variation of one. In general, ordinal data, by virtue of how they are created, *never* fall in a normal distribution. For instance, a data set might look more like a stairway that progresses upward in regular intervals. Or, take a graduating high school class. If each student is given a class rank according to academic grade point average, Luis might rank first, Janene might rank second, Marietta third, and so on. We don’t see a normal distribution in this situation because we have only one student at each academic rank. If we construct a graph that depicts the frequencies of the class ranks, we see a low, flat distribution rather than one that rises upward and peaks in the middle.

Percentile ranks, too, form a flat distribution rather than a bell-shaped curve. Percentile ranks are often used to report performance on scholastic aptitude and achievement tests. To calculate them, a researcher first determines the **raw score**—the number of

FIGURE 11.7
Common departures from
the normal distribution



test items correctly answered or number of points accumulated—that each person in the sample earns on a test or other research instrument. Each person's percentile rank is then calculated as follows:

$$\text{Percentile rank} = \frac{\text{Number of other people scoring lower than the person}}{\text{Total number of people in the sample}}$$

By the very nature of how they are calculated, percentile ranks spread people evenly over the number of possible ranks one might get; for instance, there will be roughly the same number of people earning percentile ranks of 5, 35, 65, and 95. Furthermore, although percentile ranks tell us how people have performed relative to one another, they do not tell us *how much* they differ from one another in the characteristic being assessed. In essence, percentile ranks are *ordinal data* and must be treated as such.

Choosing between Parametric and Nonparametric Statistics

Your choice of statistical procedures must depend to some degree on the nature of your data and the extent to which they reflect a normal distribution. Some statistics, known as **parametric statistics**, are based on certain assumptions about the nature of the population in question. Two of the most common assumptions are these:

- The data reflect an interval or ratio scale.
- The data fall in a normal distribution (e.g., the distribution has a central high point, and it is not seriously skewed, leptokurtic, or platykurtic).

When either of these assumptions is violated, the results one obtains from parametric statistics may be suspect.

Other statistics, called **nonparametric statistics**, are not based on such assumptions. For instance, some nonparametric statistics are appropriate for data that are ordinal rather than interval in nature. Others may be useful when a population is highly skewed in one direction or the other.

You may be thinking, Why not use nonparametric statistics all the time to avoid having to make (and possibly violate) any assumptions about the data? The reason is simple: Our most complex and powerful inferential statistics are based on parametric statistics. Nonparametric statistics are, by and large, appropriate only for relatively simple analyses.

On an optimistic note, we should point out that some statistical procedures are **robust** with respect to certain assumptions. That is, they yield generally valid results even when an assumption isn't met. For instance, a particular procedure might be as valid with a leptokurtic or platykurtic distribution as it is with a normal distribution; it might even be valid with ordinal rather than interval data. When using any statistical technique, you should consult with a statistics textbook to determine what assumptions are essential for that technique and what assumptions might reasonably be disregarded. Some statistical software packages routinely provide information about whether a particular data set meets or violates certain assumptions and make appropriate adjustments for non-normal distributions.

Descriptive Statistics

As their name implies, descriptive statistics *describe* a body of data. Here we discuss how to determine three things a researcher might want to know about a data set: points of central tendency, amount of variability, and the extent to which two or more variables are associated with one another.

Measures of Central Tendency

A *point of central tendency* is a point around which the data revolve, a middle number around which the data regarding a particular variable seem to hover. In statistical language, we use the term *measures of central tendency* to refer to techniques for finding such a point. Three commonly used measures of central tendency are the mode, the median, and the mean, each of which has its own characteristics and applications.

The **mode** is the single number or score that occurs most frequently. For instance, in this data set

3 4 6 7 7 9 9 9 9 10 11 11 13 13 13 15 15 21 26

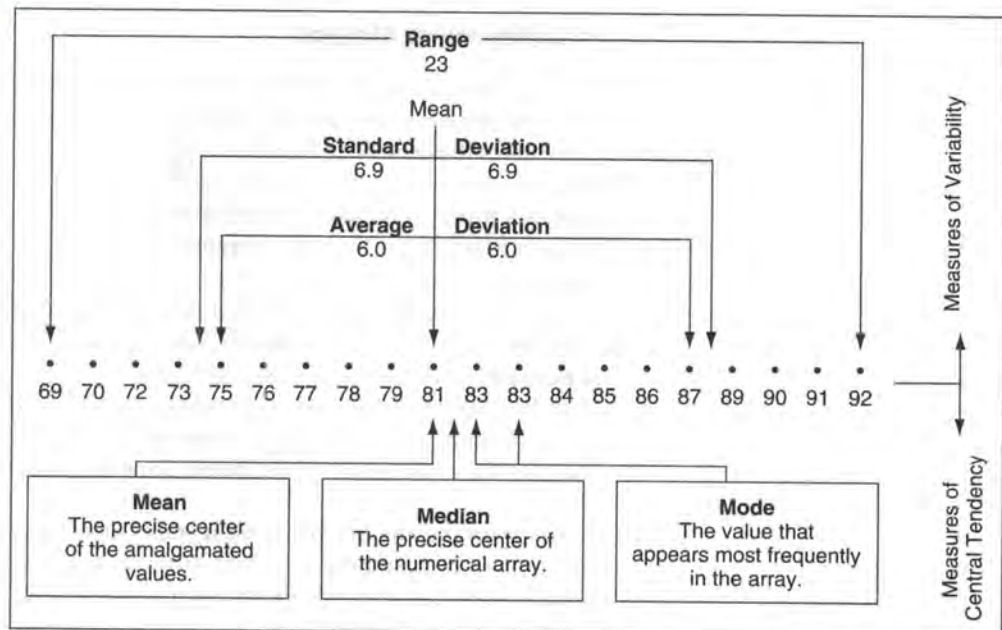
the mode is 9, because 9 occurs more frequently (four times) than any other number. Similarly, if we look at the list of Joe's grades (see p. 273), we see that only one score (83) appears more than once; thus, 83 is the mode. As a measure of central tendency, the mode is of limited value, in part because it doesn't always appear near the middle of the distribution and in part because it isn't very stable from sample to sample. However, the mode is the *only* appropriate measure of central tendency for nominal data.

The **median** is the numerical center of a set of data, with exactly as many scores above it as below it. The word *median* is a derivation of the Latin word for "middle," and so the median score is the one precisely in the middle of the series. Recall that Joe's record has 20 grades. Thus, 10 grades are above the median, and 10 are below it. The median is midway in the series between the 10th and 11th scores, or in this case, midway between the scores of 81 and 83—that is, 82 (see Figure 11.8).

You might think of the **mean** as the fulcrum point for a set of data: It represents the single point at which the two sides of a distribution "balance." Mathematically, the mean is the *arithmetic average*¹ of the scores within the data set. To find it, we calculate the sum of all the scores (adding each score every time it occurs) and then divide by the total number of scores. If we use

FIGURE 11.8

Measures of central tendency and variability for Joe's grades



¹When the word *arithmetic* is used as an adjective, as it is here, it is typically pronounced with emphasis on the third syllable ("ar-ith-MET-ic").

the symbol X to refer to each score in the data set and the symbol N to refer to the total number of scores, we calculate the mean as follows:

$$M = \frac{X_1 + X_2 + X_3 + \dots + X_N}{N}$$

Statisticians frequently use the symbol Σ (uppercase form of the Greek letter *sigma*) to designate adding all of the numbers related to a particular variable; thus, we can rewrite the formula for a mean as follows:

$$M = \frac{\Sigma X}{N}$$

Using the formula, we find that the mean for Joe's grades is 81, as illustrated in Figure 11.8. (The variation in Joe's grades, depicted in the figure as *measures of variability*, is discussed shortly.)

The mean is the measure of central tendency most commonly used in statistical analyses and research reports. However, it is appropriate only for interval or ratio data, because it makes mathematical sense to compute an average only when the numbers reflect equal intervals along a particular scale.

The median is more appropriate for dealing with ordinal data. The median is also used frequently when a researcher is dealing with a data set that is highly skewed in one direction or the other. As an example, consider this set of scores:

3 4 5 5 6 9 15 17 125

The mean for these scores is 21, a number that doesn't give us a very good idea of the point near which most of the scores are located. The median, which in this case is 6, is a better reflection of central tendency because it isn't affected by the single extreme score of 125. Similarly, medians are often used to reflect central tendency in family income levels, home values, and other such financial variables; most family incomes and home values are clustered at the lower end of the scale, with only a very few extending into the million-dollar range.

Curves Determine Means

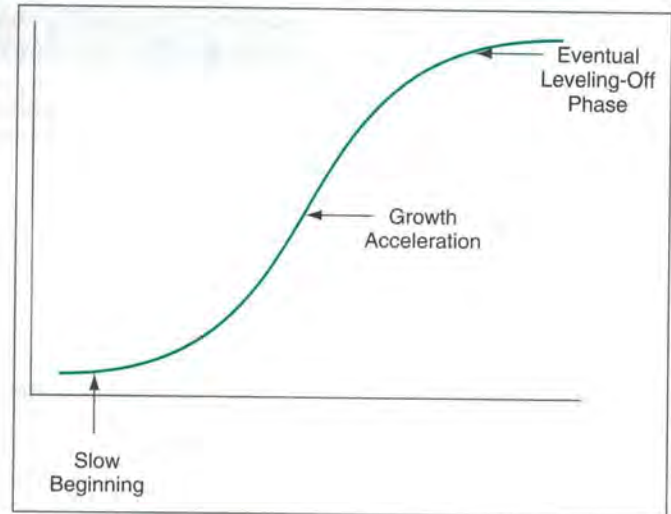
The mean as we have just described it—sometimes known as the *arithmetic mean*²—is most appropriate when we have a normal distribution, or at least a distribution that is somewhat symmetrical. But not all phenomena fit a bell-shaped pattern. Growth is one: It often follows an ogive curve that eventually flattens into a plateau (see Figure 11.9).

Growth is a function of geometric progression. As an example, let's consider the work of Thomas Robert Malthus, an English clergyman and economist who theorized about the potential for a population explosion and resulting worldwide famine. In *An Essay on the Principle of Population* (1826/1963), Malthus contended that, when unchecked, a population increases at an exponential rate, in which each successive value depends, multiplicatively, on the preceding value; for example, in the series 2, 4, 8, 16, 32, 64, 128 . . . , each number is twice the preceding number. But Malthus also predicted that the size of the human population would eventually flatten out because there is an upper limit to what Mother Earth can produce in the way of food to sustain the population. Thus, many growth curves resemble an S, as illustrated in Figure 11.9.

If we are recording the growth of bean stalks in an agronomy laboratory, we do not find the average growth by assuming a normal distribution and calculating the arithmetic mean. The statistical technique does not fit the natural fact. Instead, we use the **geometric mean**, which is computed by *multiplying* all of the scores together and then finding the N th root of the product. In other words, the geometric mean, which we can symbolize as M_g , is calculated as follows:

²Again, the emphasis in *arithmetic* is on the third syllable ("ar-ith-MET-ic").

FIGURE 11.9
Typical growth curve



$$M_g = \sqrt[N]{(X_1)(X_2)(X_3) \dots (X_N)}$$

For growth phenomena, we use the geometric mean because that is the way things grow. That is the way cells divide—geometrically.

Biologists, physicists, ecologists, and economists all encounter growth phenomena in one form or another. They all witness the same typical aspects of change: a slow beginning (a few settlers in an uninhabited region, a few bacteria on a culture); then, after a period of time, rapid expansion (the boom period of city growth, the rapid multiplication of microorganisms); and finally—sometimes but not always—a leveling-off period (the land becomes scarce and the city sprawl is contained by geographical and economic factors, the bacteria have populated the entire culture). Following are examples of situations in which the application of the geometric mean is appropriate:

- Biological growth
- Population growth
- Increments of money at compound interest
- Decay or simple decelerative situations

In every situation, one basic principle applies: The configuration of the data dictates the measure of central tendency most appropriate for that particular situation. If the data fall in a distribution that approximates a normal curve (as most data do), they call for one measure of central tendency. If they assume an ogive curve configuration (characteristic of a growth situation), they demand another measure. A *polymodal distribution*—one with several distinct peaks—might call for still a third approach; for instance, the researcher might describe it in terms of its two or more modes. Only after careful and informed consideration of the characteristics of the data can the researcher select the most appropriate statistical measure.

Thus, we must emphasize an essential rule for researchers who use statistics in their data analysis: *The nature of the data—the facts of life—governs the statistical technique, not the other way around.* Just as the physician must know what drugs are available for specific diseases and disorders, so the researcher must know what statistical techniques are suited to specific research demands. Table 11.2 presents a summary of the measures of central tendency and their uses, together with the various types of data for which each measure is appropriate.

TABLE 11.2

Using measures of central tendency for different types of data

<i>Measure of Central Tendency</i>	<i>How It Is Determined (N = number of scores)</i>	<i>Data for Which It Is Appropriate</i>
Mode	The most frequently occurring score is identified.	• Data on nominal, ordinal, interval, and ratio scales • Multimodal distributions (two or more modes may be identified when a distribution has multiple peaks)
Median	The scores are arranged in order from smallest to largest, and the middle score (when <i>N</i> is an odd number) or the midpoint between the two middle scores (when <i>N</i> is an even number) is identified.	• Data on ordinal, interval, and ratio scales • Data that are highly skewed
Arithmetic mean	All the scores are added together, and their sum is divided by the total number (<i>N</i>) of scores.	• Data on interval and ratio scales • Data that fall in a normal distribution
Geometric Mean	All the scores are multiplied together, and the <i>N</i> th root of their product is computed.	• Data on ratio scales • Data that fall in an ogive curve (e.g., growth data)

Measures of Central Tendency as Predictors

Some researchers regard the matter of central tendency from a somewhat different standpoint. They consider it from the perspective of optimal chance: What is the best prediction?

As an example, consider this situation. Suppose you are walking down the street. Suddenly you come to a crowd of people forming in a normal-curve-like manner. Where, based on your best prediction, will you find the cause for the crowd forming? The answer is simple. Where the crowd is deepest, where the greatest number of people are, you will probably find the cause for the gathering. It may be an accident, a street fight, or a person giving away free candy bars. But whatever the occasion, your best guess about the cause of the gathering lies at the point where the human mass is at its peak.

Similarly, we can often make reasonable predictions about a population based on our knowledge of central tendency. When we speak of "the average citizen," "the average student," and "the average wage earner," we are referring to those citizens, students, and wage earners who are huddled around the point of central tendency. In the broad spectrum of possibilities, we are betting on the average as being the best single guess about the nature of the total population.

Measures of Variability: Dispersion and Deviation

Up to this point, we have been discussing the question, What is the best guess? Now we turn to the opposite question: What are the worst odds? This, too, is important to know. The more that the data cluster around the point of central tendency, the greater is the probability of making a correct guess about where any particular data point lies. The farther the data are dispersed from the central axis, the greater the margin of predictive error becomes. Consider, for example, the two curves shown in Figure 11.10. The data are more similar if they cluster about the mean. Scatter them, and they lose some of their uniformity; they become more diverse, more heterogeneous. As specific data points recede farther from the mean, they lose more and more of the quality that makes them "average."

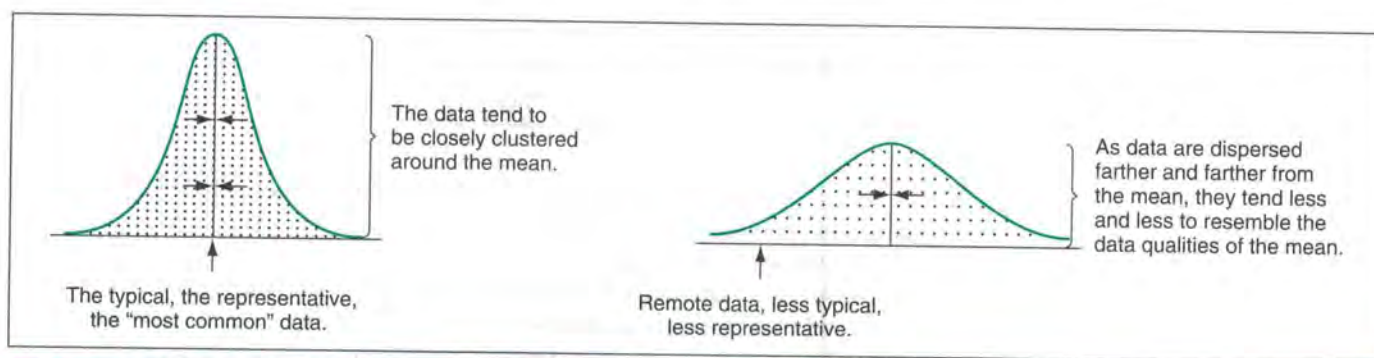


FIGURE 11.10
Distributions that differ in variability

To derive meaning from data, then, it's important to determine not only their central tendency but also their spread. And it often helps to pin down their spread in terms of one or more quantitative values.

How Great Is the Spread?

The simplest measure of variability is the **range**, which indicates the spread of the data from lowest to highest value:

$$\text{Range} = \text{Highest score} - \text{Lowest score}$$

For instance, the range for Joe's test scores is $92 - 69$, or 23 (see Figure 11.8).

Although the range is easy to compute, it has limited usefulness as a measure of variability and may even be misleading if the extreme upper or lower limits are atypical of the other values in the series. Let's take an example. Following are the numbers of children in each of ten families: 1, 3, 3, 3, 4, 4, 5, 5, 6, 15. We might say that the families range from one with a single child to a family of 15 children (a range of $15 - 1$, or 14). But this figure is misleading: It suggests that the sample shows a great deal of variability in family size. We give a more realistic estimate of variability in this sample if we say that 80% of the families have from 3 to 6 children.

Other measures of variability use the median or mean as a starting point. One such measure is the **interquartile range**. If we divide the distribution into four equal parts, Quartile 1 lies at a point where 25% of the members of the group are below it. Quartile 2 divides the group into two equal parts and is identical to the median. Quartile 3 lies at a point where 75% of the values are below it.³ The interquartile range is equal to Quartile 3 (the 75th percentile point) minus Quartile 1 (the 25th percentile point), as follows:

$$\text{Interquartile range} = \text{Quartile 3} - \text{Quartile 1}$$

Thus, the interquartile range gives us the range for the middle 50% of the cases in the distribution. Because quartiles are associated with the median, any researcher employing the median as a measure of central tendency should also consider the quartile deviation as a possible statistical measure for variability.

But now let's instead use the mean as a starting point. Imagine that we determine how far away from the mean each piece of data is in the distribution. That is, we calculate the *difference* between each score and the mean score (we call this difference the *deviation*). If we were to add all of these differences (ignoring the plus and minus signs) and then divide the sum by the *number* of scores (which reflects the number of score–mean differences as well), we get the *average* of the

³If, instead of dividing the data into 4 equal parts, we divide them into 10 equal parts, each part is called a *decile*; if into 100 equal parts, each part is called a *percentile*.

differences between any score and the mean. This number is sometimes called the *average deviation* (AD). The equation for the average deviation is:

$$AD = \frac{\sum |X - M|}{N}$$

Here $|X - M| = X - M$ without regard for plus and minus signs; in other words, it is the *absolute value* of the difference between each score and the mean.

The average deviation is easily understood and, for that reason, has some merit. It is acceptable when no further statistical procedure is contemplated. It is a little-used value, however, and the measures of standard deviation and variance have largely replaced it in most research projects.

The **standard deviation** (σ or s , for population parameter and sample statistic, respectively) is the measure of variability most commonly used in statistical procedures. To understand the reason for using the standard deviation, we must think about what happens mathematically when we find the average deviation just described. Imagine that we were to compute the average deviation *without* using the absolute values of the differences. For any number lying to the left of the mean (and thus having a value smaller than the mean), the difference between the number and the mean ($X - M$) would be a negative number. In contrast, any number lying to the right of the mean would yield a positive $X - M$ value. When we added all of the positive and negative deviations together, they would entirely counterbalance one another, essentially “canceling out” one another and yielding an overall sum of zero. To circumvent this problem, of course, the average deviation uses absolute values, thereby allowing us to ignore the plus and minus signs. Yet this is a rather dubious procedure. It is neither sound mathematics nor sound statistics to ignore what we do not like. We can, however, change negatives to positives in a perfectly acceptable mathematical manner. In arithmetic, if we multiply a negative value by itself, it becomes positive. Thus, when we square all negative differences, they become positive.

To calculate a standard deviation, we follow a procedure similar to calculating an average deviation. However, rather than taking the absolute value of the score–mean differences, we *square* the differences. After we have added the squared differences together and then divided by the number of scores, we find the square root of the quotient. Thus, the formula for a standard deviation is as follows:

$$s = \sqrt{\frac{\sum (X - M)^2}{N}}$$

It is important to note that we square the differences first and *then* add them together. If we were to add them together first and then square their sum, we would get an entirely different—and incorrect!—result.

The average deviation and standard deviation are usually similar but not identical values. As an example, if we calculate these two statistics for Joe’s grades, we get an average deviation of 6.0 and a standard deviation of 6.9 (see Figure 11.8).

Many statistical procedures use another measure of variability in addition to or instead of the standard deviation. This statistic is known as the **variance**, which is simply the standard deviation squared:

$$s^2 = \frac{\sum (X - M)^2}{N}$$

The various measures of variability we have just described (omitting the average deviation, which, as we’ve said, is rarely used) are summarized in Table 11.3.

Using the Mean and Standard Deviation to Calculate Standard Scores

Earlier we mentioned the term *raw score*, the number of correct answers or points that a person gets on a test or other measurement instrument. Such scores typically need some context to

TABLE 11.3

Using measures of variability for different types of data

Measure of Variability	How It's Determined (N = number of scores)	Data for Which It's Appropriate
Range	The difference between the highest and lowest scores in the distribution	Data on ordinal, interval, and ratio scales*
Interquartile range	The difference between the 25th and 75th percentiles	- Data on ordinal, interval, and ratio scales - Especially useful for highly skewed data
Standard deviation	$s = \sqrt{\frac{\sum(X - M)^2}{N}}$	- Data on interval and ratio scales - Most appropriate for normally distributed data
Variance	$s^2 = \frac{\sum(X - M)^2}{N}$	- Data on interval and ratio scales - Most appropriate for normally distributed data - Especially useful in inferential statistical procedures (e.g., analysis of variance)

* Measures of variability are usually inappropriate for nominal data. Instead, frequencies or percentages of each number are reported.

make them meaningful. For instance, if we say that Mary has gotten a score of 35 on a test of extroversion (i.e., on a test assessing her tendency to be socially outgoing), you might ask: What does that score *mean*? Is it high? Low? Somewhere in the middle? Without a context, a score of 35 has no meaning. We have no idea how extroverted or introverted Mary is.

Sometimes researchers provide context by converting raw scores to **norm-referenced scores**, scores that reflect where each person is positioned relative to other members of the person's group. We have already seen one example of a norm-referenced score: A *percentile rank* is the percentage of people in the group that a particular individual has scored *better than*. For example, if Mary scores at the 95th percentile on a test of extroversion, then we know that she is quite outgoing—more so than 95% of the people who have taken the test. But as noted earlier in the chapter, percentile ranks have a definite limitation: They are ordinal data rather than interval data, and so we cannot perform even such basic arithmetic operations as addition and subtraction on them. Accordingly, we will be very limited in the statistical procedures we can use with percentile ranks.

More useful in statistical analyses are standard scores. Simply put, a **standard score** tells us how far an individual's performance is from the mean with respect to standard deviation units. The simplest standard score is a *z-score*, which is calculated by using an individual's raw score (which we will symbolize as X), along with the mean and standard deviation for the entire group, as follows:

$$z = \frac{X - M}{s}$$

As an illustration, let's return to Mary's score of 35 on the extroversion test. If the mean of the scores on this test is 25, and if the standard deviation is 5, we would calculate Mary's *z-score* as follows:

$$z = \frac{35 - 25}{5} = \frac{10}{5} = 2$$

When we calculate *z-scores* for an entire group, we get a distribution that has a mean of 0 and a standard deviation of 1.

Because about half of the *z-scores* for any group of people will be a negative number—as just noted, the *mean* for the group is 0—researchers sometimes change *z-scores* into other standard-score scales that yield only positive numbers. To convert a *z-score* to another scale, we

would simply multiply the z by the new scale's standard deviation (s_{new}) and then add the new scale's mean (M_{new}) to the product obtained, as follows:

$$\text{New standard score} = (z \times s_{\text{new}}) + M_{\text{new}}$$

Let's take an example. One common standard-score scale is the IQ scale, which uses a mean of 100 and a standard deviation of 15. (As you might guess, this scale is the one on which intelligence test scores are typically based.) If we were to convert Mary's extroversion score to the IQ scale, we would plug her z -score of 2, plus a standard deviation of 15 and a mean of 100, into the preceding formula, as follows:

$$\text{IQ score} = (2 \times 15) + 100 = 130$$

Thus, using the IQ scale, Mary's score on the extroversion test would be 130.

Another commonly used standard-score scale is the **stanine**. Stanines have a mean of 5 and a standard deviation of 2. Mary's stanine would be 9, as we can see from the following calculation:

$$\text{Stanine} = (2 \times 2) + 5 = 9$$

Stanines are *always* a whole number from 1 to 9. If our calculations gave us a number with a fraction or decimal, we would round it off to the nearest whole number. If some of our calculations resulted in numbers of 0 or less, or 10 or more, we would change those scores to 1 and 9, respectively.

Standard scores take a variety of forms, each with a prespecified mean and standard deviation; z -scores, IQs, and stanines are just three examples.⁴ But in general, standard scores give us a context that helps us interpret the scores: If we know the mean and standard deviation on which the scores are based, then we also know where in the distribution any particular score lies. For instance, an IQ score of 70 is two standard deviations (30 points) below the mean of 100, and a stanine score of 6 is one half of a standard deviation (1 is half of 2) above the mean of 5.

Converting data to standard scores does not change the shape of the distribution; it merely changes the mean and standard deviation of that distribution. But imagine that, instead, we *do* want to change the nature of the distribution. Perhaps we want to change a skewed distribution into a more balanced, normally distributed one. Perhaps, in the process, we also want to change ordinal data into interval data. Several procedures exist for doing such things, but describing them would divert us from the basic nature and functions of statistics that we need to focus on here. You can find discussions of *normalizing* a data set in many basic statistics textbooks; another good resource is Harwell and Gatti (2001).

Keeping Measures of Central Tendency and Variability in Perspective

Statistics related to central tendency and variability help us summarize our data. But—so as not to lose sight of our ultimate goal in conducting research—we should remind ourselves that statistical manipulation of the data is *not*, in and of itself, research. Research goes one step further and demands *interpretation* of the data. In finding medians, means, interquartile ranges, or standard deviations, we have not interpreted the data, nor have we extracted any *meaning* from them. We have merely described the center and spread of the data. We have attempted only to see what the data look like. After learning their basic nature, we should then look for conditions that are forcing the data to behave as they do. For example, if we toss a pair of dice 100 times and one particular die yields a 5 in 80 of those tosses, we will have a distribution for that die much different from what we would expect. This may suggest to us that a reason lurks behind the particular results we have obtained. For example, perhaps we are playing with a loaded die!

⁴A standard score gaining increasing popularity for reporting academic achievement test results is the *NCE score*, which has a mean of 50 and a standard deviation of 21.06. With this particular (and seemingly very odd) standard deviation, an NCE score of 1 is equivalent to a percentile score of 1 and, likewise, an NCE score of 99 is equivalent to a percentile score of 99.

Measures of Association: Correlation

The statistics described so far—measures of central tendency and variability—involve only a single variable. Oftentimes, however, we also want to know whether two or more variables are in some way associated with one another. For example, relationships exist between age and reading ability (as illustrated in Figure 8.1 in Chapter 8), between emotional state and physical health, between the amount of rainfall and the price of vegetables in the marketplace. Consider, too, the relationships between temperature and pressure, between the intensity of light and the growth of plants, between the administration of a certain medication and the resulting platelet agglutination in the blood. Relationships among variables are everywhere. One function of statistics is to capture the nature and strength of such relationships.

The statistical process by which we discover whether two or more variables are in some way associated with one another is called *correlation*. The resulting statistic, called a **correlation coefficient**, is a number between -1 and $+1$; most correlation coefficients are decimals (either positive or negative) somewhere between these two extremes. A correlation coefficient for two variables simultaneously tells us two different things about the relationship between those variables:

- **Direction.** The direction of the relationship is indicated by the *sign* of the correlation coefficient—in other words, by whether the number is a positive or negative one. A positive number indicates a **positive correlation**: As one variable increases, the other variable also increases. For example, there is a positive correlation between self-esteem and school achievement: Students with higher self-esteem achieve at higher levels (e.g., Marsh, Gerlach, Trautwein, Lüdtke, & Brettschneider, 2007). In contrast, a negative number indicates an inverse relationship, or **negative correlation**: As one variable increases, the other variable *decreases*. For example, there is a negative correlation between the number of friends children have and the likelihood that they'll be victims of bullying: Children who have many friends are *less* likely to be bullied by their peers than are children who have few or no friends (e.g., Espelage & Swearer, 2004).
- **Strength.** The strength of the relationship is indicated by the *size* of the correlation coefficient. A correlation of $+1$ or -1 indicates a *perfect* correlation: If we know the degree to which one characteristic is present, we know exactly how much of the other characteristic exists. For example, if we know the length of a horseshoe crab in inches, we also know—or at least we can quickly calculate—exactly what its length is in centimeters. A number close to either $+1$ or -1 (e.g., $+.89$ or $-.76$) indicates a *strong* correlation: The two variables are closely related, such that knowing the level of one variable allows us to predict the level of the other variable with considerable accuracy. For example, we often find a strong relationship between two intelligence tests taken at the same time: People tend to get similar scores on both tests, especially if both tests cover similar kinds of content (e.g., McGrew, Flanagan, Zeith, & Vanderwood, 1997). In contrast, a number close to 0 (e.g., $+.15$ or $-.22$) indicates a *weak* correlation: Knowing the level of one variable allows us to predict the level of the other variable, but we cannot predict with much accuracy. For example, there is a weak relationship between intellectual giftedness and emotional adjustment: Generally speaking, people with higher IQ scores show greater emotional maturity than people with lower scores (e.g., Janos & Robinson, 1985), but many people are exceptions to this rule. Correlations in the middle range (for example, those in the .40s and .50s, positive or negative) indicate a *moderate* correlation.

The most widely used statistic for determining correlation is the Pearson product moment correlation, sometimes called the Pearson r . But there are numerous other correlation statistics as well. As in the case of the central tendency, the nature of the data determines the technique that is most appropriate for calculating correlation. In Table 11.4, we present several parametric and nonparametric correlational techniques and the kinds of data for which they are appropriate.

One especially noteworthy statistic in Table 11.4 is the *coefficient of determination*, or R^2 . This statistic, which is the square of the Pearson r , tells us *how much of the variance is accounted for* by the correlation. Although you will see this expression used frequently in research reports, researchers usually don't stop to explain what it means. By *variance*, we are specifically referring to a particular

TABLE 11.4

Examples of correlational statistics

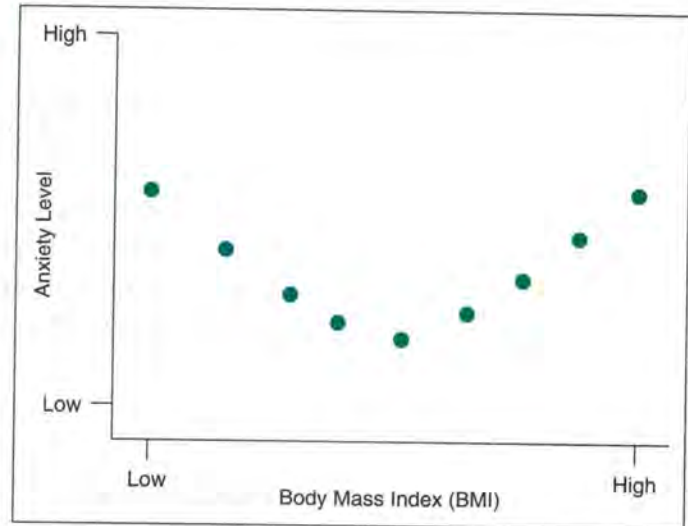
Statistic	Symbol	Data for Which It's Appropriate
Parametric Statistics		
Pearson product moment correlation	r	Both variables involve continuous data.
Coefficient of determination	R^2	This is the square of the Pearson product moment correlation; thus, both variables involve continuous data.
Point biserial correlation	r_{pb}	One variable is continuous; the other involves discrete, dichotomous, and perhaps nominal data (e.g., Democrats vs. Republicans, males vs. females).
Biserial correlation	r_b	Both variables are continuous, but one has been artificially divided into an either-or dichotomy (e.g., "above freezing" vs. "below freezing," "pass" vs. "fail").
Phi coefficient	ϕ	Both variables are true dichotomies.
Triserial correlation	r_{tri}	One variable is continuous; the other is a trichotomy (e.g., "low," "medium," "high").
Partial correlation	$r_{12.3}$	The relationship between two variables exists, in part, because of their relationships with a third variable, and the researcher wants to "factor out" the effects of this third variable (e.g., what is the relationship between motivation and student achievement when IQ is held constant statistically?).
Multiple correlation	$R_{1.23}$	One variable is related to two or more variables; here the researcher wants to compute the first variable's <i>combined</i> relationship with the others.
Nonparametric Statistics		
Spearman rank order correlation (Spearman's rho)	ρ	Both variables involve rank-ordered data and so are ordinal in nature.
Kendall coefficient of concordance	W	Both variables involve rankings (e.g., rankings made by independent judges regarding a particular characteristic) and hence are ordinal data, and the researcher wants to determine the degree to which the rankings are similar.
Contingency coefficient	C	Both variables involve nominal data.
Kendall's tau correlation	τ	Both variables involve ordinal data; the statistic is especially useful for small sample sizes (e.g., $N < 10$).

measure of variability mentioned earlier: the square of the standard deviation, or s^2 . For example, if we find that, in our data set, the R^2 between Variable 1 and Variable 2 is .30, we know that 30% of the variability in Variable 1 is reflected in its relationship with Variable 2. This knowledge will allow us to control for—and essentially *reduce*—some of the variability in our data set through such statistical procedures as partial correlation and analysis of covariance (described in Table 11.4 and later in Table 11.5, respectively). It is important to note, too, that the correlation statistics presented in Table 11.4 are all based on an important assumption: that the relationship between the two variables is a *linear* one—that is, as one variable continues to increase, the other continues to increase (for a positive correlation) or decrease (for a negative correlation). Not all relationships take a linear form, however. For example, consider the relationship between body mass index (a general measure of a person's body fat; often abbreviated as BMI) and anxiety. In one recent study (Scott, McGee, Wells, & Browne, 2008), researchers found that anxiety was highest in people who were either very underweight or very overweight; anxiety was lowest for people of relatively *average* weight. Such a relationship is known as a *U-shaped relationship* (see Figure 11.11). U-shaped and other nonlinear relationships can be detected through scatter plots and other graphic techniques, as well as through certain kinds of statistical analyses (e.g., see B. Thompson, 2008).

FIGURE 11.11

U-shaped relationship
between body mass
index (BMI) and anxiety

Based on Scott et al., 2008.



Always keep in mind that the nature of the data governs the correlational procedure that is appropriate for those data. Don't forget the cardinal rule: *Look at the data!* Determine their nature, scrutinize their characteristics, and then select the correlational technique suitable for the type of data with which you are working.

How Validity and Reliability Affect Correlation Coefficients

Beginning researchers should be aware that the extent to which one finds a statistical correlation between two characteristics depends, in part, on how well those characteristics have been measured. Even if there really *is* a correlation between two variables, a researcher won't necessarily find one if the measurement instruments he or she uses have poor validity and reliability. For instance, we are less likely to find a correlation between age and reading level if the reading test we use is neither a valid (accurate) nor reliable (consistent) measure of reading achievement.

Over the years, we authors have had many students find disappointingly low correlation coefficients between two variables that they hypothesized would be highly correlated. By looking at the correlation coefficient alone, a researcher cannot determine the reason for a low correlation any more than he or she can determine the reason for a high one. Yet one thing is certain: *You will find substantial correlations between two characteristics only if you can measure both characteristics with a reasonable degree of validity and reliability.* We refer you back to the section "Validity and Reliability in Measurement" in Chapter 4, where you can find strategies for determining and enhancing both of these essential qualities of sound measurement.

A Reminder about Correlation

Whenever you find evidence of a correlation within your data, you must remember one important point: *Correlation does not necessarily indicate causation.* For example, if you find a correlation between self-esteem and classroom achievement, you cannot necessarily conclude that students' self-esteem *influences* their achievement. Only experimental studies, such as those described in Chapter 9, allow you to draw definitive conclusions about the extent to which one thing causes or influences another.

Finding a correlation in a data set is equivalent to discovering a signpost. That signpost points to the fact that two variables are associated, and it reveals the nature of the association (positive or negative, strong or weak). It should then lead you to wonder, What is the underlying reason for the association? But the statistic alone will not be able to answer that question.

Inferential Statistics

As mentioned earlier, inferential statistics allow us to draw inferences about large populations from relatively small samples. More specifically, inferential statistics have two main functions:

1. To estimate a population parameter from a random sample
2. To test statistically based hypotheses

In this text, we do not have the space to venture too far into these areas; statistics textbooks can give you more detailed information. However, we comment briefly about several general concepts and principles.

Estimating Population Parameters

When we conduct research, more often than not we use a sample to learn about the larger population from which the sample has been drawn. Typically we compute various statistics for the sample we have studied. Inferential statistics can tell us how closely these sample statistics approximate parameters of the overall population. For instance, we often want to estimate population parameters related to central tendency (the mean, or μ), variability (the standard deviation, or σ), and proportion (P). These values in the population compare with the M or $\bar{X}$, the s , and the p of the sample (see Table 11.1, page 278).

To show you what we mean by estimation, we use a simple illustration. Jan is a production manager for a large corporation. The corporation manufactures a piece of equipment that requires a connecting-rod pin, which the corporation also manufactures. The pin fits snugly into a particular joint in the equipment, permitting a metal arm to swivel within a given arc. The pin's diameter is critical: If the diameter is too small, the arm will wobble while turning; if it is too large, the arm will stick and refuse to budge. Jan has received complaints from customers that some of the pins are faulty. She decides to estimate, on the basis of a random sample of the connecting-rod pins, how many units of the equipment may have to be recalled in order to replace their faulty pins. From this sample, Jan wants to know three facts about the thousands of equipment units that have been manufactured and sold:

1. What is the average diameter of the pins?
2. How widely do the pins vary in diameter?
3. What proportion of the pins are acceptable in the equipment units already sold?

The problem is to determine population parameters on the basis of the sample statistics. From the sample, Jan can estimate the mean and variability of the pin diameters and the proportion of acceptable pins within the population universe. These are the values represented by μ , the σ , and the P .

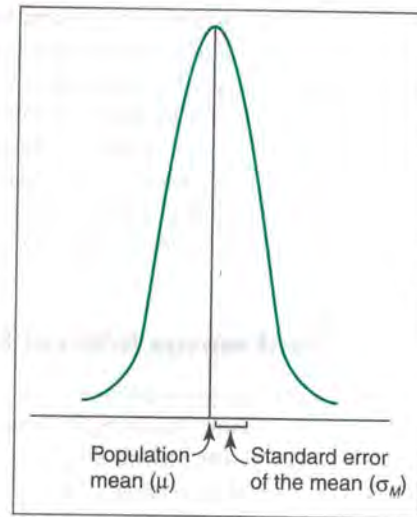
Statistical estimates of population parameters are based on the assumption that *the sample is randomly chosen and representative of the total population*. Only when we have a random, representative sample can we make reasonable guesses about how closely our statistics estimate population parameters. To the extent that a sample is nonrandom and therefore nonrepresentative—to the extent that the sample's selection has been *biased* in some way—our statistics may be poor reflections of the population from which it has been drawn.

An Example: Estimating a Population Mean

Imagine that we want to estimate the average (mean) height of 10-year-old boys in the state of New Hampshire. Measuring the heights of the entire population would be incredibly time-consuming, so we decide to measure the heights of a random and presumably representative sample of, let's say, 200 boys.

FIGURE 11.12

Distribution of sample means



Random samples from populations—please note the word *random* here—display roughly the same characteristics as the populations from which they were selected. Thus, we should expect the mean height for our sample to be approximately the same as the mean for the overall population. It will not be *exactly* the same, however. In fact, if we were to collect the heights for a second random sample of 200 boys, we would be likely to compute a slightly different mean than we had obtained for the first sample.

Different samples—even when each has been randomly selected from the same population—will almost certainly yield slightly different estimates of the overall population. The difference between the population mean and a sample mean constitutes an *error* in our estimation. Because we don't know what the exact population mean is, we also don't know how much error is in our estimate. We *do* know three things, however:

1. The means we might obtain from an infinite number of random samples form a normal distribution.
2. The *mean of this distribution of sample means* is equal to the mean of the population from which the samples have been drawn (μ). In other words, the population mean equals the average, or mean, of all the sample means.
3. The standard deviation of this distribution of sample means is directly related to the standard deviation of the characteristic in question for the overall population.

This situation is depicted in Figure 11.12.

The third characteristic just listed—the standard deviation for the distribution of sample means—is called the **standard error of the mean**. This index tells us how much the particular mean we calculate is likely to vary from one sample to another *when all samples are the same size and are drawn randomly from the same population*. Statistically, when all of the samples are of a particular size (n), the standard error of the mean is represented as

$$\sigma_M = \frac{\sigma}{\sqrt{n}}$$

Here we are faced at once with a problem. The formula we just presented involves using the population standard deviation (σ), but the purpose of using the sample was to *avoid* having to measure the entire population. Fortunately, statisticians have devised a way to estimate the standard error of the mean from the standard deviation of a *sample* drawn from the population. This formula is

$$\text{Estimated } \sigma_M = \frac{s}{\sqrt{n-1}}$$

Notice how, in both formulas, the standard error of the mean is directly related to the standard deviation of the characteristic being measured: More variability in the population leads to a larger standard error of the mean—that is, to greater variability in the sample means that we might obtain. In addition, the standard error is *inversely* related to n , the size of the sample. As the sample size increases, the standard error of the mean decreases. Thus, a larger sample size will give us a sample mean that more closely approximates the population mean. This principle holds true for estimates of other population parameters as well. In general, *larger samples yield more accurate estimates of population parameters.*

Point versus Interval Estimates

When using sample statistics to estimate population parameters, we can make two types of estimates: point estimates and interval estimates.

A **point estimate** is a single statistic that is used as a reasonable estimate of the corresponding population parameter; for instance, we might use a sample mean as a close approximation to the population mean. Although point estimates have the seeming benefit of being precise, in fact this precision is illusory. A point estimate typically does *not* correspond exactly with its equivalent in the population. Let's return to our previous example of the connecting-rod pins. Perhaps the company has produced 500,000 pins, and Jan has selected a sample of 100 of them. When she measures the diameters of these pins, she finds that the mean diameter is 0.712 centimeter, and the standard deviation is 0.020 centimeter. She guesses that the mean and standard deviation of the diameters of *all* of the pins are also 0.712 and 0.020, respectively. Her estimates will probably be close—and they are certainly better than nothing—but they won't necessarily be dead-on.

A more accurate approach—although still not 100% dependable—is to identify **interval estimates** of parameters. In particular, we specify a range within which a population parameter probably lies, and we state the probability that it actually lies there. Such an interval is often called a **confidence interval** because it attaches a certain level of probability to the estimate—a certain level of *confidence* that the estimated range includes the population parameter.

As an example, Jan might say that she is 95% certain that the mean of the 500,000 connecting-rod pin diameters her company has produced is somewhere between 0.708 and 0.716. What Jan has done is to determine that the standard error of the mean is 0.002 (see the previously presented formula for estimated σ_M). Jan knows that sample means fall in a normal distribution (look once again at Figure 11.12). She also knows that normal distributions have predictable proportions within each section of the curve (look once again at Figure 11.6). In particular, Jan knows that about 68% ($34.1\% + 34.1\%$) of the sample means lie within one standard error of the population mean, and that about 95% ($13.6\% + 34.1\% + 34.1\% + 13.6\%$) lie within two standard errors of the population mean. What she has done, then, is to go two standard errors (2×0.002 , or 0.004) to either side of her sample mean (0.712) to arrive at her 95% confidence interval of 0.708 to 0.716.

We have said enough about estimation for you to appreciate its importance. To venture further would get us involved in specific statistical procedures that are not the province of this text. For additional guidance, we urge you to consult one or more statistics textbooks, such as those listed in the "For Further Reading" list at the end of the chapter.

Testing Hypotheses

The second major function of inferential statistics is to test hypotheses. At the outset, we should clarify our terminology. The term *hypothesis* can confuse you unless you understand that it has two different meanings in research literature. The first meaning relates to a *research hypothesis*; the second relates to a *statistical hypothesis*.

Most of our discussions of hypotheses in earlier chapters have involved the first meaning of the word *hypothesis* (e.g., recall Chapter 1's discussion of the homeowner who speculates about why a table lamp may have failed). In forming a research hypothesis, a researcher speculates about how the research problem or one of its subproblems might be resolved. A research hypothesis is a reasonable conjecture, an educated guess, a theoretically or empirically based prediction. Its purpose is a practical one: It provides a temporary objective, an operational target, a logical framework that guides a researcher as he or she collects and analyzes data.

When we encounter the phrase "testing a hypothesis," however, the matter is entirely different. Here the word *hypothesis* refers to a statistical hypothesis, usually a *null hypothesis*. A **null hypothesis** (often symbolized as H_0) postulates that any result observed is the result of chance alone. For instance, if we were to compare the means of two groups, our null hypothesis would be that both groups are parts of the same population and that any differences between them—including any difference we see between their means—are strictly the result of the fact that *any* two samples from the population will yield slightly different estimates of a population parameter.

Now let's say that we look at the *probability* that our result is due to chance alone. If, for example, we find that a difference between two group means would, if due entirely to chance, occur *only one time in a thousand*, we could reasonably conclude that the difference is *not* due to chance—that, instead, something in the situation we are studying (perhaps an experimental treatment we have imposed) is systematically leading to a difference in the groups' means. This process of comparing observed data with the results that we would expect from chance alone is called *testing the null hypothesis*.

At what point do researchers decide that a result has *not* occurred by chance alone? One common cutoff is a 1-in-20 probability: Any result that would occur by chance only 5% of the time—that is, a result that would occur, on average, only one time in every 20 times—probably is *not* due to chance but instead to another, systematic factor that is influencing the data. Other researchers use a more rigorous 1-in-100 criterion: The observed result would occur by chance only one time in 100. The probability that researchers use as their cutoff point, whether .05, .01, or some other figure, is the **significance level**, or **alpha** (α). A result that, based on this criterion, we deem *not* to be due to chance is called a **statistically significant** result. When we decide that a result is due to something other than chance, we *reject the null hypothesis*.

In the "Results" section of a research report, you will often see the researcher's alpha level implied in parentheses. For example, imagine that a researcher reports that "a *t*-test revealed significantly different means for the two treatment groups ($p < .01$).¹" The " $p < .01$ " here means that the difference in means for the two groups would occur by chance less than one time in 100 *if* the two groups had been drawn from the same population. Sometimes, instead, a researcher will state the actual probability with which a result might occur by chance alone. For example, a researcher might report that "a *t*-test revealed significantly different means for the two treatment groups ($p = .003$).²" The " $p = .003$ " here means that a difference this large would occur only three times in 1,000 for two groups that come from the same population. In this situation, then, chances are good that the two groups come from *different* populations—a round-about way of saying that the two treatments differentially affected the outcome.

When we reject the null hypothesis, we must look to an alternative hypothesis—which may, in fact, be the *research hypothesis*—as being more probable. For example, if our null hypothesis is that two groups are the same and we then obtain data that lead us to reject this hypothesis, we indirectly support the opposite hypothesis: that the two groups are *different*.

In brief, we permit a certain narrow margin of variation within our data, which we deem to be natural and the result of pure chance. Any variation within this statistically permissible range is not considered to be important enough to claim our attention. Whatever exceeds these limits, however, is considered to be the result of some determinative factor other than chance, and so the influence is considered to be an important one. The term *significant*, in the statistical sense in which we have been using it, is close to its etymological meaning—namely, "giving a signal" that certain dynamics are operating within the data and merit attention.

Making Errors in Hypothesis Testing

It is possible, of course, that we may make a mistake when we decide that a particular result is not the result of chance alone. In fact, *any* result could conceivably be due to chance; our sample, although selected randomly, may be a fluke that displays atypical characteristics simply through the luck of the draw. If we erroneously conclude that a result was not due to chance when in fact it *was* due to chance—if we incorrectly reject the null hypothesis—we are making a **Type I error** (also called an *alpha error*).

Yet in another situation, we might conclude that a result is due to chance when in fact it is *not*. In such a circumstance, we have failed to reject a null hypothesis that is actually false—something known as a **Type II error** (also called a *beta error*). For example, imagine that we are testing the relative effects of a new medication versus the effects of a placebo in lowering blood cholesterol. Perhaps we find that people who have been taking the new medication have, on average, a lower cholesterol level than people taking the placebo, but the difference is a small one. We might find that such a difference could occur 25 times out of 100 due to chance alone, and so we *retain the null hypothesis*. If, in actuality, the medication does reduce cholesterol more than a placebo does, we have made a Type II error.

Statistical hypothesis testing is all a matter of probabilities, and there is always the chance that we could make either a Type I or Type II error. We can decrease the odds of making a Type I error by lowering our level of significance, say, from .05 to .01, or perhaps to an even lower level. In the process of doing so, however, we increase the likelihood that we will make a Type II error—that we will fail to reject a null hypothesis that is, in fact, incorrect. To decrease the probability of a Type II error, we would have to increase our significance level (α), which, because it increases the odds of rejecting the null hypothesis, also increases the probability of a Type I error. Obviously, then, there is a trade-off between Type I and Type II errors: Whenever you decrease the risk of making one, you increase the risk of making the other.

To illustrate this trade-off, we return to our study of the potentially cholesterol-reducing medication. There are four possibilities:

1. We correctly conclude that the medication reduces cholesterol.
2. We correctly conclude that it does not reduce cholesterol.
3. We mistakenly conclude that it is effective when it *isn't*.
4. We mistakenly conclude that it isn't effective when it *is*.

These four possibilities are illustrated in Figure 11.13. The three vertical lines illustrate three hypothetical significance levels we might choose. Imagine that the dashed middle line, Line A, represents a significance level of, say, .05. In this particular situation (such will not always be the case), we have a slightly greater chance of making a Type I error (represented by the upper shaded area) than of making a Type II error (represented by the lower shaded area). But the significance level we choose is an arbitrary one. We could reduce our chance of a Type I error by decreasing our significance level to, say, .03. Line B to the right of Line A in the figure represents such a change; notice how it would create a smaller box (lower probability) for a Type I error but create a larger box (greater probability) for a Type II error. Alternatively, if we raise the significance level to, say, .06 (as might be represented by Line C, to the left of Line A in the figure), we decrease the probability of a Type II error but increase the probability of a Type I error.

There is perhaps nothing more frustrating for the novice researcher than obtaining insignificant results—those that, from a statistical standpoint, could have been due to chance alone. Following are three suggestions for decreasing the likelihood of making a Type II error and thereby increasing the likelihood of correctly rejecting an incorrect null hypothesis. In other words, these are suggestions for increasing the **power** of a statistical test:

- *Use as large a sample size as is reasonably possible.* The larger the sample, the less the statistics you compute will diverge from actual population parameters.⁵

⁵Formulas exist for computing the power of statistical procedures for varying sample sizes. For example, see Lipsey (1990) or Murphy, Myers, and Wolach (2009).

FIGURE 11.13

The trade-off between
Type I and Type II errors

		Our conclusion about the medication	
		No beneficial effect	Beneficial effect
Reality (the Truth) about the medication	No beneficial effect	Correct conclusion	Type I error
	Beneficial effect	Type II error	Correct conclusion
		C	A B

- *Maximize the validity and reliability of your measures.* Measures of variables in a research study rarely have perfect (100%) validity and reliability, but some measures are more valid and reliable than others. Research projects that use measures with high validity and reliability are more likely to yield statistically significant results. (Again we refer you to the section “Validity and Reliability in Measurement” in Chapter 4.)
- *Use parametric rather than nonparametric statistics whenever possible.* As a general rule, nonparametric statistical procedures are less powerful than parametric techniques. By “less powerful,” we mean that nonparametric statistics typically require larger samples to yield results that enable the researcher to reject the null hypothesis. When characteristics of the data meet the assumptions for parametric statistics, then, we urge you to use these statistics. (Look once again at the section “Choosing between Parametric and Nonparametric Statistics” earlier in this chapter.)

It is important—in fact, critical—to keep in mind that *whenever we test more than one statistical hypothesis, we increase the probability of making at least one Type I error*. Let’s say that, for a particular research project, we have set the significance level at .05, such that we will reject the null hypothesis whenever we obtain results that would be due to chance alone only 1 time in 20. And now let’s say that as we analyze our data, we perform 20 different statistical tests, always setting α at .05. In this situation, although we won’t necessarily make a Type I error, the odds are fairly high that we will.⁶

Another Look at Statistical Hypotheses versus Research Hypotheses

Novice researchers sometimes become so wrapped up in their statistical analyses that they lose track of their overall research problem or hypothesis. In fact, testing a null hypothesis involves

⁶When testing 20 hypotheses at a .05 significance level, the probability of making at least one Type I error is .642—in other words, chances are better than 50-50 that at least one Type I error is being made. In general, the probability of making a Type I error when conducting multiple statistical tests is $1 - (1 - \alpha)^n$, where α (alpha) is the significance level and n is the number of tests conducted.

nothing more than a statistical comparison of two distributions of data—one hypothetical (a theoretical ideal) and one real (the distribution of data collected from a research sample). A researcher simply uses one or more statistical procedures to determine whether calculated values sufficiently diverge from the statistical ideal to reject the null hypothesis.

Testing a statistical hypothesis does not, in and of itself, contribute much to the fulfillment of the basic aim of research: a systematic quest for undiscovered knowledge. Certainly statistical analyses are invaluable tools that enable us to find patterns in the data and thus help us detect possible dynamics working within the data. But we must never stop with statistical procedures that yield one or more numerical values. We must also *interpret* those values and give them meaning. The latter process includes the former, but the two should never be confused.

It is often the case that the statistical hypothesis is the opposite of the research hypothesis. For example, we might, as our research hypothesis, propose that two groups are different from one another. As we begin our statistical analysis, we set out to test the statistical hypothesis that the two groups are the same. *By disconfirming the null hypothesis, we indirectly find support for our research hypothesis.* This is, to be sure, a backdoor approach to finding evidence for a research hypothesis, yet it is the approach that is typically taken. The reasons for this approach are too complex to be dealt with in a text such as this one. Suffice it to say that it is mathematically much easier to test a hypothesis that an equivalence exists than to test a hypothesis that a difference exists.

Examples of Statistical Techniques for Testing Hypotheses

Table 11.5 lists many commonly used parametric and nonparametric statistical techniques for testing hypotheses. We hope it will help you make decisions about the techniques that are most appropriate for your own research situation. As you can see in the table, however, nonparametric techniques exist only for relatively simple statistical analyses, such as comparing measures of central tendency or testing the statistical significance of correlations. When your research problem calls for a relatively sophisticated analysis (e.g., multiple regression or structural equation modeling), parametric statistical procedures—and the underlying assumptions about the data these procedures require—are your only viable option.

We urge you to consult one or more statistics texts to learn as much as you can about whatever statistical procedures you use. Better still, enroll in one or more statistics courses! You can successfully solve your research problem only if you apply statistical procedures appropriately and thereby conduct accurate analyses of your data.

Meta-Analysis

Occasionally researchers use inferential statistics not to analyze and draw conclusions from data they have collected but instead to analyze and draw conclusions about *other researchers' statistical analyses*. Such analysis of analyses is known as **meta-analysis**. A meta-analysis is most useful when many studies have already been conducted on a particular topic or research problem and another researcher wants to pull all of the results together into a neat and mathematically concise package.

The traditional approach to synthesizing previous studies related to a particular research question is simply to describe them all, pointing out which studies yield which conclusions, which studies contradict others, and so on. In a meta-analysis, however, the researcher integrates the studies statistically rather than verbally. After pinning down the research problem, the researcher:

1. *Conducts a fairly extensive search for relevant studies.* The researcher does not choose arbitrarily among studies that have been reported about the research problem. Instead, he or she uses some systematic and far-reaching approach (e.g., searching in several pre-specified professional journals, using certain keywords in a search of online databases) to identify studies that have addressed the topic of interest.

TABLE 11.5

Examples of inferential statistical procedures and their purposes

Statistical Procedure	Purpose
Parametric Statistics	
Student's <i>t</i> -test	To determine whether a statistically significant difference exists between two means. A <i>t</i> -test takes slightly different forms depending on whether the two means come from separate, independent groups (an <i>independent-samples t</i> -test) or, instead, from a single group or two interrelated groups (a <i>dependent-samples t</i> -test).
Analysis of variance (ANOVA)	To examine differences among three or more means by comparing the variances (s^2) both within and across groups. As is true for <i>t</i> -tests, ANOVAs take slightly different forms for separate, independent groups and for a single group; in the latter case, a <i>repeated-measures</i> ANOVA is called for. If an ANOVA yields a significant result (i.e., a significant value for <i>F</i>), you should follow up by comparing various pairs of means using a <i>post hoc comparison of means</i> .
Analysis of covariance (ANCOVA)	To look for differences among means while controlling for the effects of a variable that is correlated with the dependent variable (the former variable is called a <i>covariate</i>). This technique can be statistically more powerful than ANOVA (i.e., it decreases the probability of a Type II error).
<i>t</i> -test for a correlation coefficient	To determine whether a Pearson product moment correlation coefficient (<i>r</i>) is larger than would be expected from chance alone.
Regression	To examine how effectively one or more variables allow(s) you to predict the value of another (dependent) variable. A <i>simple linear regression</i> generates an equation in which a single independent variable yields a prediction for the dependent variable. A <i>multiple linear regression</i> yields an equation in which two or more independent variables are used to predict the dependent variable.
Factor analysis	To examine the correlations among a number of variables and identify clusters of highly interrelated variables that reflect underlying themes, or <i>factors</i> , within the data.
Structural equation modeling (SEM)	To examine the correlations among a number of variables—often with different variables measured for a single group of people at different points in time—in order to identify possible causal relationships (<i>paths</i>) among the variables. SEM encompasses such techniques as <i>path analysis</i> and <i>confirmatory analysis</i> and is typically used to test a previously hypothesized model of how variables are causally interrelated. SEM enables a researcher to identify a <i>mediator</i> in a relationship: a third variable that may help explain why Variable A seemingly leads to Variable B (i.e., Variable A affects the mediating variable, which in turn affects Variable B). SEM also enables a researcher to identify a <i>moderator</i> of a relationship: a third variable that alters the nature of the relationship between Variables A and B (e.g., Variables A and B might be correlated when the moderating variable is high but not when it is low, or vice versa). (Mediating and moderating variables are discussed in more detail in Chapter 2.) When using SEM, the researcher must keep in mind that the data are <i>correlational</i> in nature; thus, any conclusions about cause-and-effect relationships are speculative at best.
Nonparametric Statistics	
Mann-Whitney <i>U</i>	To compare the medians of two groups when the data are ordinal rather than interval in nature. This procedure is the nonparametric counterpart of the independent-samples <i>t</i> -test in parametric statistics.
Kruskal-Wallis test	To compare three or more group medians when the data are ordinal rather than interval in nature. This procedure is the nonparametric counterpart of ANOVA.
Wilcoxon signed-rank test	To compare the medians of two correlated variables when the data are ordinal rather than interval in nature. This procedure is a nonparametric equivalent of a dependent-samples <i>t</i> -test in parametric statistics.
Chi-square (χ^2) goodness-of-fit test	To determine how closely observed frequencies or probabilities match expected frequencies or probabilities. A chi-square can be computed for nominal, ordinal, interval, or ratio data.
Odds ratio	To determine whether two dichotomous nominal variables (e.g., smokers vs. non-smokers and presence vs. absence of heart disease) are significantly correlated. This is one nonparametric alternative to a <i>t</i> -test for Pearson's <i>r</i> .
Fisher's exact test	To determine whether two dichotomous variables (nominal or ordinal) are significantly correlated when the sample sizes are quite small (e.g., $n < 30$). This is another nonparametric alternative to a <i>t</i> -test for Pearson's <i>r</i> .

2. *Identifies appropriate studies to include in the meta-analysis.* The researcher limits the chosen studies to those that involve a particular experimental treatment (in experimental studies), pre-existing condition (in ex post facto studies), or other variable that is the focus of the meta-analysis. He or she may further restrict the chosen studies to those that involve particular populations, settings, assessment instruments, or other factors that may impact a study's outcome.
3. *Converts each study's results to a common statistical index.* Previous researchers may possibly have used different statistical procedures to analyze their data. For example, if each researcher has compared two or more groups that received two or more different experimental interventions, one investigator may have used a *t*-test, another may have conducted an analysis of variance, and a third may have conducted a multiple regression. The meta-analytic researcher's job is to find a common denominator here. Typically, when an experimental intervention has been studied, an **effect size** is calculated for each study; that is, the researcher determines how much of a difference the intervention makes (in terms of standard deviation units) in each study. The effect sizes of all of the studies are then used to compute an average effect size for that intervention.

The statistical procedures used in meta-analyses vary widely, depending, in part, on the research designs of the included studies; for instance, correlational studies require different meta-analytic procedures than experimental studies. We must point out, too, that meta-analyses, although they can make an important contribution to the knowledge bases of many disciplines, are not for the mathematically fainthearted. If you are interested in conducting a meta-analysis, several of the resources listed in the "For Further Reading" section at the end of this chapter should prove helpful.

Using Statistical Software Packages



Earlier in the chapter, we mentioned that general-purpose spreadsheet programs can be used to describe and analyze sets of quantitative data. However, many spreadsheets are limited in their statistical analysis capabilities. As an alternative, you may want to consider using one of the several statistical software packages now widely available for use on personal computers (e.g., SPSS, SAS, SYSTAT, Minitab, Statistica). Such packages have several advantages:

- *Increased user-friendliness.* As statistical software programs become increasingly powerful, they also become more user-friendly. In most cases, the programs are logical and easy to follow, and results are presented in easy-to-read table format. Selection of the proper statistics and interpretation of the results, however, are still left to the researcher.
- *Range of available statistics.* Many of these programs include a wide variety of statistical procedures, and they can easily handle large data sets, multiple variables, and missing data points.
- *Assumption testing.* A common feature of statistical software packages is to test for characteristics (e.g., skewness, kurtosis) that might violate the assumptions on which a parametric statistical procedure is based.
- *Speed of completion.* As always, a major benefit of using the computer is the speed with which it accomplishes tasks. Even relatively simple statistical procedures might take several hours if executed by hand; more complex analyses are, for all practical purposes, impossible for a researcher to conduct using only paper, pencil, and a hand-held calculator.
- *Graphics.* Many statistical programs allow the researcher to summarize and display data in tables, pie charts, bar graphs, or other graphics.

In Appendix B, we show you some of the basics of one statistical software program, SPSS, and use a small data set to illustrate some of the ways you might use it.⁷

For frugal researchers—especially those whose research problems require small data sets and relatively simple statistical procedures (e.g., computing standard deviations, correlation coefficients, or chi-squares)—online statistics calculators provide another option. Two examples are www.easycalculation.com (www.easycalculation.com/statistics/statistics.php) and GraphPad Software's QuickCalcs (www.graphpad.com/quickcalcs). A Google or Yahoo! search for “online statistics calculator” can identify other helpful websites as well.

Yet we must caution you: *A computer cannot and should not do it all for you.* You may be able to perform sophisticated calculations related to dozens of statistical tests and present the results in a variety of ways, but if you do not understand how the results relate to your research problem, or if you cannot otherwise make logical, theoretical, or pragmatic sense of what your analyses have revealed, then all your efforts have been for naught. Powerful statistical software programs make it all too easy to conduct studies so large and complex that the researcher loses sight of the initial research question. In the words of Krathwohl (1993), the researcher eventually behaves “like a worker in a laboratory handling radioactive material, . . . manipulating mechanical hands by remote control from a room outside a sealed data container. With no sense of the data, there is little basis for suspecting an absurd result, and we are at the mercy of the computer printout” (p. 608).

Ultimately *you* must be in control of your analyses; you must know what calculations are being performed and why. Only by having an intimate knowledge of the data can you derive true meaning from the statistics computed and use them to address your research problem.

Interpreting the Data

To the novice researcher, statistics can be like the voice of a bevy of sirens. For those who have never studied or have forgotten the works of Homer, the *Odyssey* describes the perilous straits between Scylla and Charybdis. On these treacherous rocks resided a group of Sirens—svelte maidens who, with enticing songs, lured sailors in their direction and, by so doing, caused ships to drift and founder on the jagged shores.

For many beginning researchers, statistics hold a similar appeal. Subjecting data to elegant statistical routines may lure novice researchers into thinking they have made a substantial discovery, when in fact they have only calculated a few numbers. Behind every statistic lies a sizable body of data; the statistic may summarize these data in a particular way, but it cannot capture all the nuances of the data. The entire body of data collected—not any single statistic calculated—is what ultimately must be used to resolve the research problem. There is no substitute for the task the researcher ultimately faces: to discover the meaning of the data and its relevance to the research problem. Any statistical process you may employ is only ancillary to this central quest.

At the beginning of the chapter, we presented a hypothetical data set for 11 school children and discovered that the 5 girls in the sample had higher reading achievement test scores than the 6 boys. Shortly thereafter, we presented actual data about growth marks on the shells of the chambered nautilus. Perhaps these examples piqued your curiosity. For instance, perhaps you wondered about questions such as these:

- Why were all of the girls' scores higher than those of the boys?
- Why were the intervals between each of the scores equidistant for both boys and girls?
- What caused the nautilus to record a growth mark each day of the lunar month?
- Is the relationship between the forming of the partitions and the lunar cycle singular to the nautilus, or are there other similar occurrences in nature?

⁷At the instructor's request, this book can be packaged with the Student Version of SPSS at a discount; the CD for the software provides versions for both Windows and Macintosh users. Please contact your local Pearson representative if you are an instructor who is interested in setting up such a package for your students.

Knowledge springs from questions like these. But we must be careful not to make snap judgments about the data we have collected. It is all too easy to draw hasty and unwarranted conclusions. Even the most thorough research effort can go astray at the point of drawing conclusions from the data.

For example, from our study of 11 children and their reading achievement scores, we might conclude that girls read better than boys. But if we do so, we are not thinking carefully about the data. Reading is a complex and multifaceted skill. The data *do not* say that girls read better than boys. The data *do* say that, on a particular test given on a particular day to a particular group of 11 children, all girls' scores were higher than all boys' scores and that, for both boys and girls, the individual scores differed by intervals of 4. The apparent excellence of the girls over the boys was limited to test performance in those reading skills that were specifically measured by the test. Honesty and precision dictate that all conditions in the situation be considered and that we make generalizations only in strict accordance with the data. On the following day, the same test given to another 11 children might yield different results.

In general, interpreting the data means several things:

1. *Relating the findings to the original research problem and to the specific research questions and hypotheses.* Researchers must eventually come full circle to their starting point—why they conducted a research study in the first place and what they hoped to discover—and relate their results to their initial concerns and questions.
2. *Relating the findings to pre-existing literature, concepts, theories, and research studies.* To be useful, research findings must in some way be connected to the larger picture—to what people already know or believe about the topic in question. Perhaps the new findings confirm a current theoretical perspective, perhaps they cast doubt on common “knowledge,” or perhaps they simply raise new questions that must be addressed before humankind can truly understand the phenomenon in question.
3. *Determining whether the findings have practical significance as well as statistical significance.* Statistical significance is one thing; **practical significance**—whether findings are actually useful—is something else altogether. For example, let's return to that new medication for lowering blood cholesterol level mentioned earlier in the chapter. Perhaps we randomly assign a large sample of individuals to one of two groups; one is given the medication, and the other is given a placebo. At the end of the study, we measure cholesterol levels for the two groups and then conduct a *t*-test to compare the group means. If our sample size is quite large, the standard error of the mean will be very small, and we may therefore find that even a minor difference in the cholesterol levels of the two groups is statistically significant. Is the difference *practically* significant as well? That is, do the benefits of the medication outweigh its costs and any unpleasant side effects? A calculation of *effect size*—how different the cholesterol levels are for the treatment and control groups relative to the standard deviation for one or both groups—can certainly help us as we struggle with this issue. But ultimately a statistical test cannot, in and of itself, answer the question. Only the human mind—the researcher, practitioners in the field of medicine, and so on—can answer it.
4. *Identifying limitations of the study.* Finally, interpreting the data involves outlining the weaknesses of the study that yielded them. No research study can be perfect, and its imperfections inevitably cast at least a hint of doubt on its findings. Good researchers know—and also report—the weaknesses along with the strengths of their research.

A Sample Dissertation

To illustrate this final step in the research process—interpretation of the data—we present excerpts from Kimberly Mitchell's doctoral dissertation in psychology conducted at the University of Rhode Island (Mitchell, 1998). The researcher was interested in identifying possible causal factors leading to eating disorders and substance abuse, and she hypothesized that

family dynamics and child abuse might be among those factors. She drew on three theoretical perspectives that potentially had relevance to her research question: problem behavior theory, social cognitive theory, and the theory of cognitive adaptation. She administered several surveys to a large sample of undergraduate students and obtained a large body of correlational data about the students' childhoods, eating habits, drug use, and so on. She then used *structural equation modeling* (described briefly in Table 11.5) as a means of revealing possible—we must emphasize the word *possible*—cause-and-effect relationships in her data set.

The dissertation refers to several psychological theories and concepts with which many of our readers may not be familiar. Nevertheless, as you read the excerpts, you should be able to see how the author frequently moves back and forth between her results and the broader theoretical framework. We pick up the dissertation at the point where Mitchell begins to summarize and interpret her results.

Dissertation ANALYSIS 8

DISCUSSION

Summary of Results and Integration

The purpose of this study was to integrate several theories that are beneficial for understanding health-risk behaviors. Problem Behavior Theory (Jessor, 1987), Social Cognitive Theory... (Bandura, 1977a), and the Theory of Cognitive Adaptation (Taylor, 1983) are similar in that they all pose a cognitive component within the individual that is crucial to overcome the potential negative consequences of life stressors... This study supports these three theories, as well as previous research in the field. It extends the research by linking these theories into a single comprehensible framework for understanding the link between the childhood stressors of sexual abuse and negative family functioning and adult substance misuse of alcohol, illicit drugs, and eating.

A series of structural equation models revealed the powerful impact individuals' perceptions of their confidence and their interactions with their environment play on health-risk behavior. The first three models examined various ways childhood stressors (sexual abuse and family functioning) could predict current health-risk behaviors (alcohol use, illicit drug use, and binge eating). Examination of the first three models (Full, Direct, and Mediation) and chi-square difference tests revealed that the mediators (self-efficacy, life satisfaction, and coping) are extremely important in predicting health-risk behaviors. This [finding] supports Jessor's (1987) theory that problem behavior is the result of the interaction of the personality system, perceived environment, and the behavioral system. The personality system is measured by the cognitive mediator constructs; the perceived environment by the family functioning construct; and the behavioral system by the outcome constructs... [T]he socialization an individual encounters throughout childhood through interactions with family members appears to influence both how the individual perceives the self and the environment around him/her. These factors seem to propel individuals to behave in ways that may or may not be risky for their health.

Comments

The author capitalizes the names of the three theories. More often, researchers use lowercase letters when referring to particular theoretical perspectives. Either approach is acceptable as long as the author is consistent.

Notice how the author begins with a "grand conclusion" of sorts, which she supports in subsequent paragraphs. She also explains how she has expanded on existing theories by integrating them to explain the phenomenon she has studied.

The "models" she refers to here are multivariable flowcharts that reflect how some variables may influence other variables, perhaps directly or perhaps indirectly through additional, mediator variables.

Self-efficacy refers to people's confidence in their ability to perform a task (e.g., resist the temptation to abuse alcohol) successfully. It is a central concept in Bandura's social cognitive theory, one the three theoretical frameworks on which the author bases her study.

Furthermore, Jessor (1987) suggests that problem behaviors in which adolescents engage are interrelated and co-vary. Donovan and Jessor (1985) suggest that diverse problem behavior, such as alcohol abuse, risky sexual behavior, and drug use constitute a single behavioral syndrome. The current study supports this notion. All of the structural models revealed a positive relationship between alcohol and drug use, as well as a positive relationship between drug use and binge eating. Although the relationship between alcohol use and binge eating was not found to be significant, they are indirectly related through drug use. Such relationships support the idea that these health-risk behaviors constitute a single behavioral syndrome. Future research with a longitudinal design is needed to see if there is a linear trend among these variables. . . .

[The author continues with a discussion of more specific aspects of her findings and their relevance to the three theoretical frameworks. We pick up her discussion again when she summarizes her conclusions.]

Summary of Conclusions

There are several conclusions that can be drawn from this study. First, in support of Problem Behavior Theory (Jessor, 1987), health-risk behaviors may be part of a single behavioral syndrome. The consistent relationships found throughout the models between alcohol use and drug use, as well as [between] drug use and binge eating, reveal the presence of a higher order behavioral syndrome.=

Second, there is a complex relationship between child sexual abuse and family functioning in terms of their ability to predict life satisfaction, coping, and self-efficacy. While child sexual abuse was found to significantly predict coping and life satisfaction, the inclusion of family functioning into the model made these paths disappear. The initial finding indicates a confounding of child sexual abuse and family functioning rather than sexual abuse itself. Furthermore, the constant relationship between child sexual abuse and family functioning shows that, although child sexual abuse does not directly predict the mediator constructs, it plays a role in the prediction indirectly.

Third, family functioning and cognitive mediators interact in specific and consistent ways to determine health-risk behaviors. Those students with high levels of family functioning are likely to have high life satisfaction, more effective coping strategies, and higher self-efficacy for alcohol use, drug use, and eating. In turn, these cognitive factors interact to predict health-risk behavior.

[The author continues with additional conclusions, and then turns to the limitations of her study.]

Study Limitations

The present study offers several important findings to the literature. Yet, there are some limitations to the study as well. First, the design was cross-sectional rather than longitudinal. Structural equation modeling is a multivariate technique that is well utilized with longitudinal data (Maruyama, 1998). By incorporating longitudinal data into the overall design, one can begin to establish causality in the results. The use of cross-sectional data with this sample does not allow the researcher to make causal statements about the findings. For example, the data cannot tell us whether self-efficacy for

Notice how the author continually connects her findings with the theoretical frameworks she is using.

Here the author points out both what she has found and what she has not found.

Although the author has previously presented each of her conclusions, she summarizes them all here. Such a summary is typical of lengthy research reports. It is quite helpful to readers, who might easily lose track of some important conclusions as they read earlier portions of a report.

The author makes the point that two of her independent (predictor) variables, child sexual abuse and family functioning, are highly interrelated. Their strong correlation is reflected in the models identified through her structural equation modeling procedures.

The author's use of the term cross-sectional is somewhat different from our use of it in Chapter 8. She simply means that she collected all data from her sample at one time, rather than following the sample over a lengthy period and collecting data at two or more times. As the author states, a longitudinal design would have better enabled her to identify important factors that preceded—and so may have had a causal effect on—other factors.

alcohol use comes before actual alcohol use or vice versa. Furthermore, the study asks the participant to answer a portion of the survey retrospectively, such as [is true for] the child sexual abuse and family functioning items. This brings up problems with how reliable the responses are due to the length of time that has passed between the incident(s) in question and the time of the study. . . .

A second limitation to this study is the nature of the sample itself. Although the sample size is excellent ($n=469$), there were disproportionate numbers of men and women (125 and 344, respectively). Furthermore, the sample was extremely homogeneous (87% White; 91% freshman or sophomore; 74% with family income over \$35,000; and 73% Catholic or Protestant). This degree of similarity among participants limits the generalizability of the study results to other populations. Yet the results are still important because this is a population at high risk for alcohol use, drug use, and bulimia-related binge eating.

Another limitation to this study is the lack of response to the probing sexual abuse questions. Approximately one half of the 91 students who reported sexual abuse did not respond to the in-depth questions regarding the abuse experience(s) (e.g., degree of trust with perpetrator, frequency of abuse). This could be due to the nature of the survey itself or [to] the environment in which students filled out the survey. In terms of the nature of the survey, once students responded to the overall sexual abuse questions geared to determine whether they were abuse survivors or not, they were instructed to skip the next five questions if their responses to the previous seven questions were all "Never." It is possible that students who did not respond "Never" to the seven questions skipped the follow-up questions anyway in a desire to finish the survey quickly. The second possibility to the lack of response is the environment in which students took the survey. Students were asked to sign up for a designated one-hour time slot to participate in the study. It is highly likely that students signed up for the same time slots as their friends in class and subsequently sat next to each other while filling out the survey. Due to the close proximity and the sensitive nature of the questions, some sexual abuse survivors may not have wanted to fill out additional questions in fear that their friends might see. Better procedures in the future would be to have all students fill out all questions, whether they are abuse survivors or not, and/or to allow them to have more privacy while taking the survey. . . .

A final limitation of the study is the use of self-report data only. Self-report data may be fraught with problems derived from memory restrictions and perception differences. A more comprehensive design would include actual physical ways to measure the outcome variables. For example, the researcher could have strengthened the design by taking blood or urine samples to examine drug use. The problem here is that [the latter] method requires a great deal of time and money to undertake.

[The researcher concludes the discussion by talking about potential implications of her findings for clinical practice and social policy.]

The author points out a problem with using surveys to learn about people's prior life experiences: Human memory is not always accurate. Her use of the word reliable here refers to accuracy and dependability (i.e., validity) of the results, rather than to reliability as we have previously defined the term.

The author explains ways in which her sample was not completely representative of the overall population of older adolescents and young adults but also makes a good case for the value of studying this sample.

The author identifies gaps (missing data) in her survey data and suggests plausible explanations for them. At the end of the paragraph, she offers suggestions for how future research might minimize such gaps.

By perception differences, the author is presumably referring to how different participants may have interpreted their prior experiences and/or items on the survey. An additional weakness of self-report data is that some participants may have intentionally misrepresented their prior experiences and/or current behaviors.

NOTE: Excerpt is from *Childhood Sexual Abuse and Family Functioning Linked with Eating and Substance Misuse: Mediated Structural Models* (pp. 92–94, 114–119) by K. J. Mitchell, 1998, unpublished doctoral dissertation, University of Rhode Island, Kingston. Reprinted with permission.

PRACTICAL APPLICATION Analyzing Data in a Quantitative Study

You can gain a clearer understanding of statistics and statistical procedures by reading about them in research reports and using them in actual practice. If your research project involves quantitative data, the following checklist can help you clarify which statistical analyses might be most appropriate for your situation.

CHECKLIST

Questions to Consider When Choosing a Statistical Procedure

CHARACTERISTICS OF THE DATA

- _____ 1. Are the data _____ continuous or _____ discrete?
- _____ 2. What scale do the data reflect? Are they _____ nominal, _____ ordinal, _____ interval, or _____ ratio?
- _____ 3. What do you want to do with the data?
 - _____ Calculate central tendency? If so, with which measure? _____
 - _____ Calculate variability? If so, with which measure? _____
 - _____ Calculate correlation? If so, with which measure? _____
 - _____ Estimate population parameters? If so, which ones? _____
 - _____ Test a null hypothesis? If so, at what confidence level? _____
 - _____ Other? (specify) _____
- _____ 4. State your rationale for processing the data as you have just indicated you intend to do.
 - _____
 - _____
 - _____

INTERPRETATION OF THE DATA

- _____ 5. After you have treated the data statistically to analyze their characteristics, what will you then have?
 - _____
 - _____
- _____ 6. From a research standpoint, what will your interpretation of the data consist of? How will the statistical analyses help you solve any part of your research problem?
 - _____
 - _____
- _____ 7. What remains to be done before your problem (or any one of its subproblems) can be resolved?
 - _____
 - _____
- _____ 8. What is your plan for carrying out this further interpretation of the data?
 - _____
 - _____
 - _____
 - _____

For Further Reading

- Agresti, A., & Finlay, B. (2009). *Statistical methods for the social sciences* (4th ed.). Upper Saddle River, NJ: Pearson.
- Arthur, W., Jr., Bennett, W., Jr., & Huffcutt, A. I. (2001). *Conducting meta-analysis using SAS*. Mahwah, NJ: Erlbaum.
- Azen, R., & Walker, C. M. (2011). *Categorical data analysis for the behavioral and social sciences*. New York: Routledge.
- Bechhofer, R. E., Santner, T. J., & Goldsman, D. M. (1995). *Design and analysis of experiments for statistical selection, screening, and multiple comparisons*. New York: Wiley.
- Bruning, J. L., & Kintz, B. L. (1997). *Computational handbook of statistics* (4th ed.). Boston: Allyn & Bacon.
- Coladarchi, T., Cobb, C. D., Minium, E. W., & Clarke, R. C. (2011). *Fundamentals of statistical reasoning in education with CD* (3rd ed.). New York: Wiley.
- Coolidge, F. L. (2006). *Statistics: A gentle introduction* (2nd ed.). Thousand Oaks, CA: Sage.
- Cooper, H. (2009). *Research synthesis and meta-analysis: A step-by-step approach* (4th ed.). Thousand Oaks, CA: Sage.
- Cramer, D. (1998). *Fundamental statistics for social research: Step-by-step calculations and computer techniques Using SPSS for Windows*. New York: Routledge.
- Cummings, G., & Finch, S. (2005). Inference by eye: Confidence intervals and how to read pictures of data. *American Psychologist*, 60, 170–180.
- Fennessy, J. (1968). The general linear model: A new perspective on some familiar topics. *American Journal of Sociology*, 74, 1–27.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Thousand Oaks, CA: Sage.
- Field, A., & Miles, J. (2010). *Discovering statistics using SAS*. Thousand Oaks, CA: Sage.
- Fowler, F. J., Jr. (2008). *Survey research methods* (4th ed.). Thousand Oaks, CA: Sage. [See Chapter 10.]
- Gonzalez, R. (2009). *Data analysis for experimental design*. New York: Guilford.
- Gravetter, F. J., & Wallnau, L. B. (2011). *Essentials of statistics for the behavioral sciences* (7th ed.). Belmont, CA: Wadsworth/Cengage.
- Haskins, L., & Jeffrey, K. (1990). *Understanding quantitative history*. New York: McGraw-Hill.
- Heiman, G. W. (2006). *Basic statistics for the behavioral sciences* (5th ed.). Boston: Houghton Mifflin.
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Thousand Oaks, CA: Sage.
- Kirk, R. E. (2008). *Statistics: An introduction* (5th ed.). Belmont, CA: Wadsworth/Cengage.
- Kline, R. B. (2010). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford.
- Kranzler, J. H. (2011). *Statistics for the terrified* (5th ed.). Upper Saddle River, NJ: Pearson.
- Lind, D. A., Marchal, W. G., & Wathen, S. A. (2010). *Basic statistics for business and economics* (14th ed.). New York: McGraw-Hill.
- Phillemer, D. B. (1994). One- versus two-tailed hypothesis tests in contemporary educational research. *Educational Researcher*, 20(9), 13–17.
- Phillips, J. L., Jr. (2000). *How to think about statistics* (6th ed.). New York: Henry Holt.
- Rosner, B. (2011). *Fundamentals of biostatistics* (7th ed.). Monterey, CA: Brooks/Cole/Cengage.
- Rowntree, D. (2004). *Statistics without tears: A primer for non-mathematicians*. Boston: Allyn & Bacon.
- Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford, England: Oxford University Press.
- Spector, P. E. (2001). *SAS programming for researchers and social scientists* (2nd ed.). Thousand Oaks, CA: Sage.
- Sweet, S. A., & Grace-Martin, K. (2012). *Data analysis with SPSS: A first course in applied statistics* (4th ed.). Upper Saddle River, NJ: Pearson.
- Thompson, B. (2008). *Foundations of behavioral statistics*. New York: Guilford.
- Vogt, W. P., & Johnson, R. B. (2011). *Dictionary of statistics and methodology: A nontechnical guide for the social sciences* (4th ed.). Thousand Oaks, CA: Sage.
- Wilkinson, L., & the Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 594–604.
- Wood, P. (2000). Meta-analysis. In G. M. Breakwell, S. Hammond, & C. Fife-Schaw (Eds.), *Research methods in psychology* (2nd ed., pp. 414–425). Thousand Oaks, CA: Sage.
- Wright, D. B. (2006). The art of statistics: A survey of modern techniques. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 879–901). Mahwah, NJ: Erlbaum.

Now go to MyEducationLab at www.myeducationlab.com to take assessments to evaluate your mastery of chapter concepts. Assignments and Activities exercises and Building Skills in Reading Research exercises are also available to help you master concepts and skills. Feedback for many of these exercises is provided so that you can see why your answers are correct or incorrect.

12

Writing the Final Research Report

Ultimately, what you put on paper and how you put it there reveals your knowledge, the quality of your thinking, and your standards of excellence more eloquently than anything else you do.

Bringing a research effort to its rightful conclusion involves writing a report that is faithful to the data but also finds meaning in those data. The research report is a straightforward document that sets forth clearly and precisely what the researcher has done to resolve the research problem. In structure, it is factual and logical. Like the research proposal, it makes no pretense at being a work of fine literature. It must, however, be comprehensible so that readers can easily grasp what the researcher has done and found. It must also be flawless in its sentence and paragraph structure, punctuation, and spelling. The research document you write is a clear reflection of your scholarship as a researcher, and for this reason it is often used as a culminating measure of a student's educational achievements.

Getting Started

If you are writing a thesis or dissertation, be sure to check first with your university's graduate school office to ascertain whether it has a prescribed set of guidelines for writing theses. Check such matters as paper quality, width of margins, size and style of font, and heading format. What is permitted at one institution may be unacceptable at another. Ask whether your university has a style manual for writing research documents or whether it recommends that you adhere to a particular published style manual.

University guidelines aside, different disciplines tend to adhere to different styles in research reports; for example, psychologists typically use *American Psychological Association (APA)* style, whereas historians tend to use *Chicago* style. Differences among the styles are most noticeable in the formats used for citations and reference lists. For example, the dissertation near the end of Chapter 11 includes citations within parentheses, consistent with *APA* style. In contrast, the dissertation near the end of Chapter 7 uses footnotes, consistent with *Chicago* style. Table 12.1 lists four commonly used styles, along with sources of information about each one.

Surfing the Internet for Writing Assistance



Numerous sites on the Internet offer assistance on the nitty-gritty details of different styles. In the rightmost column of Table 12.1 are websites that are active as this book goes to press (helpful websites for CSE style are limited and spotty at best). Another strategy is to use a search engine such as Google or Yahoo! and type such keywords as "style manuals," "MLA style," and the like in the search box; doing so will lead you to many potentially useful sites at universities and elsewhere.

Some sites on the Internet offer more general suggestions for writers. A good example is the Online Writing Lab (OWL) at Purdue University (owl.english.purdue.edu). If you go to OWL's

TABLE 12.1

Commonly used styles in research reports

Style	Manual	Online Assistance
APA Style: American Psychological Association	<i>American Psychological Association (APA) (2010). Publication manual of the American Psychological Association (6th ed.).</i> Washington, DC: Author.	www.apastyle.org/ www.psychwww.com/resource/apacrib.htm
Chicago Style: University of Chicago	<i>Chicago manual of style (16th ed.) (2010).</i> Chicago: University of Chicago Press.	www.chicagomanualofstyle.org
CSE Style: Council of Science Editors	Council of Science Editors (2006). <i>Scientific style and format: The CSE manual for authors, editors, and publishers (7th ed.)</i> . Reston, VA: Author.	
MLA Style: Modern Language Association	Modern Language Association (2008). <i>MLA style manual and guide to scholarly publishing (3rd ed.)</i> . New York: Author.	www.mla.org/style

website and click on “Site Map,” you can find links to discussions of a wide range of topics related to writing—how to write persuasively, how to enhance sentence clarity, when to use various pronouns and verb tenses, in what circumstances to use hyphens, and so on. OWL also provides guidance on APA, *Chicago*, and MLA styles.

Learn by Looking

Perhaps the best way to understand and appreciate the nature of research reports—and to prepare yourself for writing one—is to look at existing reports. Any university library should have a collection of graduates’ theses and dissertations on its shelves. You can find theses and dissertations from other universities in the ProQuest Dissertations & Theses database, an online resource to which many university libraries subscribe. Many federal agencies and professional associations disseminate research reports on numerous topics and often post these reports online.

Keep in mind that not all published research reports provide good models for novice researchers. Some, in fact, are poorly written. If you have trouble reading and understanding a report concerning a topic about which you have considerable knowledge, you might very well conclude that the author of the report is *not* a writer whose style you want to emulate!

Planning a Research Report

A research report has a relatively simple format. In general, it should achieve four objectives:

1. It should give readers a clear understanding of the research problem and why it merited an in-depth investigation.
2. It should describe exactly how data were collected in an attempt to resolve the problem.
3. It should present the obtained data precisely and completely. The data presented in the report should substantiate all the interpretations and conclusions that the report contains.
4. It should interpret the data for the reader and demonstrate exactly how the data resolve the research problem. A report that merely presents raw data and uninterpreted facts (in the form of tables, graphs, and other data-summary devices) is of little help to the reader in deriving meaning from those data.

In the following sections, we discuss each of these matters. We then address how best to conclude the report and, in a subsequent section, remind you once again about the importance of being completely truthful and forthcoming about what you have accomplished.

Description of the Research Problem

The statement of the problem and any other information needed to understand it should comprise the first section or chapter of the final report. Readers should be able to comprehend *from the report alone* what the problem is and what its ramifications are. Also, readers should appreciate the setting in which the problem was conceived. In addition, readers should learn why, from both academic and practical standpoints, the study was an important one to conduct.

The first section of the research report has but one purpose: to create a meeting of minds between the writer and readers of the report. Many research reports begin badly because their writers have not set forth the problem clearly and completely for readers. Such omissions get readers off to a confused start, which can impose a cloudy haze over the rest of the report. The writer of a research report must keep in mind that readers are likely to know only those things that the writer has actually put on paper.

After a few introductory comments (perhaps a few sentences or paragraphs) that provide the background and a rationale for the study, the document should set forth clearly and unmistakably the problem that has been researched. Often, an appropriate subheading can draw readers' attention to the research problem. If the problem has been divided into subproblems, these should be presented following the statement of the problem and announced with proper subheadings. And, of course, any preliminary hypotheses should be clearly stated in this opening section. By presenting the problem, its subproblems, and any hypotheses, the writer gives readers a clear and complete understanding of the *principal thrust* of the research effort. With this objective in mind, readers will be in a better position to understand the interpretation of the data and to judge the merits of the research.

This section should also define any terms that may have multiple meanings or in some other way might be ambiguous. For a meeting of minds, it is imperative that the researcher and readers share a common understanding of key concepts around which the research effort has revolved.

Any delimitations should be clearly set forth as well. All who read the research report should know precisely how far the research effort extended and where limits were set. Into what relevant areas did the research effort not inquire? What aspects of the problem have not been studied? Readers want answers to these questions, and they expect to find those answers in the opening pages of the report.

Also, be sure to include the assumptions you have made in designing your study and collecting and interpreting your data. For guidance on pinning down definitions of terms, delimitations, and assumptions, we refer you to Chapter 2.

The extent to which related literature is presented in the first section of the report depends on the nature of the report. In a journal article, the literature immediately relevant to the problem is summarized in the introductory paragraphs before the statement of the research problem. In a thesis or dissertation, only a few key works are identified in the first chapter, and the bulk of the related literature is reserved for a separate chapter.

Description of the Method

The method that was used to collect data—including the sample, assessment instruments, and procedures—should be described with the utmost precision. From this description, readers should know exactly what was done, to the point where the readers could replicate the study and, presumably, get similar results.

More generally, the *design* of the study should be clear. In particular, the researcher should state whether qualitative or quantitative methods (or both) were used and what particular research traditions were followed—for example, whether the study was an ethnography, a grounded theory study, a longitudinal study, a survey, a single-group time-series study, a 2-by-2 factorial design, or some combination of approaches.

Qualitative researchers also engage in *reflexivity*: Because their data collection has inevitably been influenced by their own assumptions and values, they openly acknowledge their biases and speculate on how these may have affected what they did, what data they collected, and how they interpreted their results.

Presentation of the Data

After readers fully understand just what the problem was and the manner in which it was investigated, the next question is, What is the evidence? For the most part, the data are presented *as they relate to the problem and its subproblems*. You have gathered a mass of data. You have then codified, arranged, and organized the data sets into subsets, each of which corresponds to a particular part of the problem being studied. You describe these various subsets in a logical sequence. Before beginning a discussion of each subproblem and its attendant data, it is helpful to restate the subproblem, perhaps in the same words that were used in the first section of the report. Doing so will help to keep readers' attention focused on the specific aspect of the research problem under discussion.

One logical approach is to devote a separate section (each with its own heading or subheading) to each subproblem and its pertinent quantitative and/or qualitative data. Present the subproblem, present the data germane to it, analyze and interpret those data, and draw conclusions that the data warrant. Each section might end with a brief summary in which the findings of that particular section are shown in relationship to the general problem and any relevant subproblems.

Yet you need not limit your discussion of results only to the problem and subproblems you have identified at the beginning of your report. For instance, if you have designed a rating-scale instrument to assess people's attitudes or beliefs regarding a particular issue, you should, if applicable, report the internal consistency reliability of the instrument (see the discussion of internal consistency reliability in Chapter 4, as well as the footnote on page 199 in Chapter 8). And in a later subsection in your results section, you might describe additional findings of interest—for instance, unexpected gender differences in participants' performance or intriguing questionnaire responses that raised issues you had not initially considered.

So that readers don't get lost in the data presentation, it is often helpful to begin the discussion of results with an *advance organizer* in which you lay out the overall organization of how the results will be presented. We refer you to the guidelines "Writing to Communicate" in Chapter 1 for a description of advance organizers.

The data should be presented thoroughly and, of course, accurately. In many cases, it is helpful to organize some of the data in tables, figures, and other concise presentations. A **table** is usually an arrangement of words, numbers, or combinations of these elements in a two-dimensional matrix for the purpose of exhibiting certain information in compact and comprehensive form. A **figure** is any kind of graphic illustration other than a table: a graph, chart, map, flowchart, photograph, drawing, or other device to convey an idea—sometimes a picture really is worth a thousand words! Many computer software programs can create tables and figures for you; for example, see Microsoft Word and Microsoft Excel (for tables and graphs), Inspiration (for flowcharts), and Mapititude or Mapland (for maps). All tables and figures should be specifically labeled (e.g., "Table 1," "Figure 3") and have captions that describe their contents, and you should refer to all of them in the text of your report. If you are writing a thesis or dissertation, most universities ask that you present tables and figures as soon after the in-text reference as possible; however, check with your own university's graduate school office for its own requirements about such things.¹

When the data have been subjected to statistical analysis, present your rationale for employing the particular statistical approach(es) you have used. It is important to know not only that you used a particular technique but also *why* you used it. In fact, throughout the entire research process, you should keep in mind that, generally, the answer to the question "Why?" is just as important as the answer to the question "What?" One of the weakest links in many research reports is the failure to substantiate what one has done with a solid rationale as to why one has done it.

When statistical analyses have *not* been conducted (as is often true in qualitative studies) the data must be presented in such a way that they speak for themselves. As previously mentioned

¹When writing manuscripts for publication, tables and figures often appear at the end of the document, and a notation within the running text indicates where they should be inserted in the published version of the report. For more specific guidance, consult the "Instructions to Authors" guidelines for the journal or other publication in question; such instructions are often found in each issue of the publication or on the publisher's website.

in Chapters 4 and 6, qualitative researchers often engage in *thick description*, presenting the data in such detail that readers can see for themselves what is going on. One well-known ethnographer takes this approach:

In striking the delicate balance between providing too much detail and too little, I would rather err on the side of too much; conversely, between overanalyzing and underanalyzing data, I would rather say too little. (Wolcott, 1994, p. 350)

Descriptions of data in quantitative studies are typically written in an objective, “scientific” style. Those in qualitative studies vary from the objective and aloof, on the one hand, to the more subjective and personal, on the other. Qualitative researchers frequently include dialogues and participants’ statements to illustrate their findings. They may also use metaphors and analogies to make a point. We see a simple yet effective “anti-metaphor”—an example of what something was *not*—in Matthew McKenzie’s (2003) dissertation about the Boston Marine Society (previously excerpted at length in Chapter 7):

No mere gentleman’s club, common work experiences defined the society as a community, set aside from the rest of the town. (McKenzie, 2003, p. 20; emphasis added)

Regardless of how you organize your presentation of the data, it is imperative that you present them as evidence for the conclusions you draw. If the data are extensive and you choose to present them only in summary form in the main body of the report, you might present them in their entirety in an appendix; this strategy is especially common in qualitative research reports. In this way, anyone wishing to replicate the results of the research effort should be able to reach essentially the same conclusions.

Do not forget that you may also be testing one or more hypotheses. Somewhere—probably in the closing paragraphs of each section or chapter—you should indicate whether the data did or did not support the hypotheses being tested. Restate the hypotheses and say explicitly whether they were supported or not supported by the data.

Interpretation of the Data

All too frequently, researchers believe that, having once presented the facts and figures, they have done all that needs to be done. To display the data is certainly important, but as we have said so many times before, the *interpretation of the data* is the essence of research. Without inquiring into the intrinsic meaning of the data, no resolution of the research problem or its subproblems is possible.

One common error that many researchers make is to fail to exploit the data fully. One cannot turn over the facts too often, look at them from too many angles, or chart, graph, and arrange them in too many ways. Ask simple questions of those data. This is not to suggest that you should analyze the data to the point of virtually guaranteeing yourself a Type I error. But sometimes simple questions and naive approaches will afford you startling insights. Have you thought of plotting the data? What has caused the plotted data to peak? To reach a plateau? To dip or plummet? Do dynamics within the data have relevance to events that lie beyond them? Questions like these may sometimes crack the shell of the data and reveal the meaning within.

On the other hand, a researcher must not go *too* far beyond the data. Beginning researchers often lose sight of what they have actually found; so enthusiastic are they about their topic that they make extravagant claims and unwarranted inferences. As an example, one of us authors once sat on a doctoral dissertation committee for a student who had been studying the use of regional dialects in children’s literature. Although the student drew many appropriate conclusions from her data, one of her conclusions was that literature that incorporates a regional dialect can help school children develop “an understanding and acceptance of sociocultural groups other than their own.” The student had collected no data whatsoever about children who were reading such literature, let alone data specifically related to their understanding and acceptance of diverse sociocultural groups. The student’s “conclusion” was, in reality, merely her strongly held conviction about the value of literature written in various dialects, and she should have presented it as such.

Be especially careful that you don't draw conclusions about causation or influence when the design of your study doesn't warrant such conclusions. A point we previously made in Chapters 8 and 11 bears repeating: *Correlation does not, in and of itself, indicate causation.* Certainly you can speculate that there *might* be a cause-and-effect relationship between two correlated variables, but you should never state or imply that there *definitely* is one. Make the speculative nature of your conclusion crystal clear, and back up your speculation with contemporary theory, other researchers' findings, or qualitative data in your own study. And stay away from words and phrases that imply a causal relationship—words and phrases such as *influence*, *affect*, *bring about*, *help students develop*, and *lead people to believe*. All of these inappropriately communicate that *one thing leads to another* in a causal manner.

And by all means, avoid the word *prove*. Research data rarely prove that a particular hypothesis is true. Remember, inferential statistics are based on probabilities: If a particular finding is statistically significant, it probably *wasn't* a fluke, a result that one might get strictly by chance two or three times in a thousand. Even so, it *might* be a fluke due strictly to the researcher's (unlucky) luck of the draw. A good researcher always makes this point clear to readers, for instance by saying something such as "The results support the hypothesis that . . ." or "The significant difference in means for the two groups is consistent with the premise that . . ."

Research is indeed an exciting quest, but researchers must never let their enthusiasm interfere with their objectivity in interpreting and drawing conclusions from the data. The answer to the research question should rest solidly and completely on its own empirical foundation.

Look the data steadfastly in the face. Report honestly what those data reveal to you. Ferret out every conclusion you have drawn, underscore it in red, and then be sure that the data in the tables, graphs, and other exhibits solidly support what those words underlined in red declare. That is good research.

What if the data *don't* support your predictions? Does this mean your hypotheses were wrong? Not necessarily. Look once again at your methodology and statistical analyses to see if you can identify one or more weaknesses in what you have done. Perhaps one of your measurement instruments had lower validity or reliability than you had anticipated and therefore was not yielding accurate and dependable measures of a critical variable in your study. Perhaps you gave participants misleading instructions or asked them misleading questions. Perhaps your statistical analyses lacked power—maybe your sample was too small or your measures too unreliable—and so you made a Type II error. You should report any weaknesses and flaws in your study that may have influenced its outcome.

At the same time, maybe your hypotheses *were* wrong. In the interest of advancing the frontiers of knowledge, you must be sufficiently objective to admit when your thinking was flawed and offer reasonable explanations—perhaps in the form of alternative hypotheses that future research efforts might test—for the results you obtained.

In the final analysis, *the data must speak for themselves*. The researcher is only the mouthpiece. You may not like what the data say. They may not confirm your fondest hopes or support your preconceived opinions, but the researcher is the servant of the scientific method. That method looks at evidence squarely and without prejudice; it reports candidly and precisely what the impersonal data affirm.

Defending one's research effort is an academic tradition. *Defend* in this sense means "to justify one's conclusions, to support one's statements with the backing of solid data that have been presented in the document." Nothing short of this will suffice.

Concluding the Report

Any research report should end by bringing closure to the interpretation of the data. In a thesis or dissertation, this discussion is often in a separate section or chapter, perhaps one titled "Summary, Conclusions, and Recommendations."

In this section, all loose threads should be gathered together. This is the place for looking backward, for distilling into a few paragraphs precisely what has been accomplished in each phase of the research activity. One should be able to see the research endeavor as through the

wrong end of a telescope or set of binoculars: clearly, in miniature, with all significant aspects brought together in proper perspective. Here the researcher should clearly summarize the findings and conclusions pertaining to the problem and the subproblems. The conclusions should be entirely supported by the data presented. Then, the researcher is ready for several final steps: (a) stating whether or not the hypotheses have been supported (quite possibly repeating what was previously said in the data presentation section; (b) identifying weaknesses and limitations in the study as it was designed or carried out; (c) identifying possible practical implications of the results; and (d) making recommendations for further research as a follow-up to the present study.

A few words should be said about summaries. The whole research project—the problem, the data and their organization, the relationships and interrelationships—is so clear in the mind of the researcher that he or she may forget that readers are not so intimately acquainted with the project. Through lengthy and intensive involvement in a study, the researcher has an acute awareness of the master plan, the relation of each component to the total study, the parts as they fit into the whole. Readers, however, are not so fortunate. As they proceed through the report, they need to stop occasionally to consider and reconsider how the text fits into the total investigation.

To facilitate readers' "journey" through a research report, a good researcher provides regular guideposts in the form of headings, transitional words and paragraphs, and other means of helping readers follow the train of thought. But in addition, the researcher provides a summary at the close of each extended discussion. By pausing long enough to summarize what has been said and what relevance such discussion has to the overall research effort, the researcher maintains the unity of the whole. Discussions that ramble on and on tend to produce psychological numbness, bewilderment, and confusion. Frequent summaries prevent such reader disorientation.

Maintaining Your Academic Integrity

By **academic integrity**, we mean conducting and writing about research with utmost honesty and a desire to learn and convey the truth—and *nothing but* the truth—about a topic of investigation. In writing a research report, academic integrity includes all of the following:

- Appropriately crediting the words and ideas of other people (see the discussion of plagiarism in Chapter 3)
- Maintaining confidentiality and protecting participants' right to privacy (in some cases, this may require using pseudonyms or altering demographic information for certain participants, in which case you should specifically state that you have made these changes)
- Explicitly identifying any biases in your sample selection—for instance, by reporting low return rates in mailed surveys or high attrition (drop-out) rates in longitudinal studies
- Describing any participants you dropped from your research sample and explaining why you dropped them
- Describing the limitations of your measurement instruments—for instance, by reporting any evidence of poor validity or reliability
- Describing any procedures you may have used to fill in missing data points in order to increase the number of participants for which you could conduct various statistical analyses
- Providing a comprehensive report of your research findings, including those findings that do *not* support your hypotheses
- Explicitly identifying any potential confounding variables that may cast doubt on conclusions about cause-and-effect relationships

We must emphasize a point we previously made in Chapter 4: *Researchers must report their findings in a complete and honest fashion, without misrepresenting what they have done or intentionally misleading others about the nature of their findings.* Only by being honest with one another can researchers truly advance the frontiers of knowledge. To misrepresent or mislead others in any way—no matter how well-intentioned those actions might be—is to potentially lead a community of scholars astray in their quest for knowledge.

Front Matter and End Matter

In addition to the essential sections described earlier, many lengthy research reports, including theses and dissertations, also contain **front matter** (content that precedes the introductory first chapter) and **end matter** (content that follows the final chapter). We take a few pages to describe this material.

Preliminary Pages

The preliminary pages include all the introductory material that precedes the discussion of the research problem and study. The title page comes first; this also includes the author and, typically, a university affiliation and date. In a thesis or dissertation, the title page is followed by a page for signatures of the faculty advisor and research committee. Next are an abstract, a page for the dedication (if any), an acknowledgment of indebtedness to individuals who have assisted in or in some other way supported the research, a table of contents, lists of any tables and figures, and, if desired, a preface.

In some instances, copyright information is included on the title page or the page immediately after it. **Copyright** is the protection given by law to the authors of literary, dramatic, musical, artistic, and other intellectual works. In the United States, this is U.S. Code, Title 17. Under current U.S. law, which applies to works created on or after January 1, 1978, copyright protection lasts for 70 years following the author's death. A thesis or dissertation is protected by copyright law even if you do not register it with the United States Copyright Office. Nevertheless, registering it often provides reassurance and piece of mind. As this book goes to press in 2011, the fee for filing a research report with the copyright office is \$35 or \$65 for online and paper applications, respectively. You can get more information about U.S. copyright laws and procedures at the Copyright Office's website (www.copyright.gov).

The **abstract** provides a summary of the entire research effort in a paragraph or two. For a journal article, the length of the abstract is usually 100 to 250 words, depending on the journal. For a dissertation, the abstract should be 350 words or less. The abstract should include sufficient information about the research problem, methodology, results, and interpretations to give potential readers an idea as to whether the study addresses a topic of concern to them and therefore merits their further attention. The abstract you write is likely to be included in one or more abstract collections in either paper collections (e.g., *Dissertation Abstracts International* and/or published collections of abstracts specific to your academic discipline) or online databases (e.g., ProQuest Dissertations & Theses) that then become available at many research libraries around the world. It is essential, therefore, that you take seriously the task of writing the abstract and describe your project as clearly, precisely, and succinctly as possible.

To get a sense of what abstracts entail, we urge you to look at numerous examples in professional journals and doctoral dissertations. As you peruse each example, ask yourself questions such as these:

- What was this research project intended to accomplish—in other words, what research problem did it set out to address?
- What was the general design and methodology of the study?
- What were the results, and what conclusions can reasonably be drawn from them?

Find several abstracts that enable you to answer all of these questions easily, and use those abstracts as models as you write your own abstract.

We cannot overemphasize the importance of writing a good abstract. It should summarize your project and findings clearly and succinctly—enough so that readers gain a concrete, stand-alone, take-away understanding of what you have done. But it should also pique readers' interest sufficiently that they follow up by reading your full research report. The more that people read your report, the more that your research project and report will have an impact on your field and, as a result, will enhance the world's knowledge of your research topic.

The **acknowledgments** page graciously recognizes the assistance of those people through whose kindness the research effort has been possible. These people may include those who introduced the researcher to data sources that aided in completing the research or those who guided the study and gave counsel or support—perhaps an academic dissertation committee, a faithful typist and proofreader, and family members who encouraged and assisted in the research effort. One hallmark of a true scholar is to say “Thank you” to those who have given their time and assistance to support one’s efforts and aspirations. The acknowledgments page is the proper place for the expression of such indebtedness.

The remainder of the front matter indicates the content and organization of the text. The most important of this material is the **table of contents**. The table of contents is a bird’s-eye view of what the document contains, how it is organized, and where each part can be found. Following the main table of contents are often two more specific ones, one for the tables and another for the figures that appear throughout the report.

Endnotes and Footnotes

Generally, **endnotes** (appearing at the end of the text) and **footnotes** (appearing at the bottom of relevant pages within the text itself) are used for three purposes. First, depending on the style manual being followed, such notes may be used to indicate sources of information and ideas (e.g., see the footnotes in the sample dissertation at the end of Chapter 7). Second, endnotes and footnotes are occasionally used to acknowledge permission to quote or reproduce something from a copyrighted document. When you quote extensively or use a table or other graphic representation from a copyrighted work in a report you intend to publish or distribute widely, you must secure permission to reprint the material (in writing) from the owner of the copyright (the publisher or author). After you use the material, an endnote or footnote (rather than a citation) may be used to indicate the exact source from which the material was taken, followed by the words “Reprinted by permission” or other wording stipulated by the copyright holder.

A third important function of endnotes or footnotes is to supplement information in the text of the report with additional information that strengthens the discussion. This type of note should be used sparingly and should not be used to explain complicated concepts. Keep such notes short and to the point. If you find your endnotes or footnotes becoming overly long and involved, sharpen your ideas and integrate them into the body of the report.

Reference List

A reference list at the end of your report allows readers to locate and use the sources you have cited. For this reason, it is imperative that reference information be complete—it should include references for *all* of your citations—and accurate. The reference list is not a bibliography, however; that is, it should *not* include references that you have *not* specifically cited in your report. Tempting as it might be to list all of the many books, journal articles, and other resources you have examined in an effort to better understand your topic, resources that you don’t specifically cite in your literature review or elsewhere have no place in your reference list.

Each entry in the reference list should contain information about the author, year of publication, title of the work, and publication information. To some extent, researchers in different academic disciplines format their reference lists differently, and you should follow the format that your institution or your discipline requires. Furthermore, you should apply that format consistently throughout your reference list.

In Chapter 3 we suggest the use of a database to keep track of sources you use in your literature review. Software programs specifically designed for this purpose—known as *bibliographic software*—can also assist you in completing your reference list. Commercially available programs include Biblioscape (www.biblioscape.com), EndNote (www.endnote.com), OneNote (office.microsoft.com/en-us/onenote), and ProCite (www.procite.com). Other programs, such as

FIGURE 12.1**Sample reference list
(APA style)**

From *Effects of Training in Self-Generation on the Quality of Students' Questions, Class Notes and Examination Scores* (p. 130) by D. L. Jackson, 1996, unpublished doctoral dissertation, University of Northern Colorado, Greeley. Adapted with permission.

NOTE: At the time Jackson wrote her dissertation, APA prescribed underlining the titles of books, periodicals, and paper presentations, as well as underlining periodical volume numbers. We have changed her underlines to italics in order to adhere to contemporary APA style.

- Jonassen, D. H., Beissner, K., & Yacci, M. (1993). *Structural knowledge: Techniques for representing, conveying, and acquiring structural knowledge*. Hillsdale, NJ: Erlbaum.
- Kardash, C. M., & Amlund, J. T. (1991). Self-reported learning strategies and learning from expository text. *Contemporary Educational Psychology*, 16, 117–138.
- Keppel, G. (1982). *Design and analysis: A researcher's handbook* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- Kiewra, K. A. (1985a). Learning from a lecture: An investigation of notetaking, review, and attendance at a lecture. *Human Learning*, 4, 73–77.
- Kiewra, K. A. (1985b). Students' notetaking behaviors and the efficacy of providing the instructor's notes for review. *Contemporary Educational Psychology*, 10, 378–386.
- Kiewra, K. A., Dennison, R. S., & Benton, S. L. (1995, April). *How studying text supplements affects prose production*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Kiewra, K. A., & Fletcher, H. J. (1984). The relationship between notetaking variables and achievement measures. *Human Learning*, 3, 273–280.

BiblioExpress (www.biblioscape.com/biblioexpress.htm) and Bibliography Writer (available on many Internet websites; search for "Bibliography Writer software") are freeware that can be downloaded without charge from the Internet. Virtually all of these software packages allow you to choose the particular style manual to apply in creating your reference list.

One widely used style in reference lists is that of the American Psychological Association (APA). APA style is described in detail in the association's *Publication Manual* (2010). We illustrate this format in Figure 12.1, which presents an excerpt from the reference list of Jackson's (1996) dissertation. Let's briefly look at how entries in an APA-style reference list are formatted with respect to author, date, title, and publication information. We'll then consider additional information required for sources obtained on the Internet.

Author. In an APA-style reference list, the author's name appears with the surname first, followed by the author's first and, if applicable, middle initials. When multiple authors are involved, the names are separated by commas. For example, the first and last references in Figure 12.1 have three and two authors, respectively. Commas are always used between the names (even between only two names), and an ampersand (&) is used before the last name in the list.

Date of publication. Following the author's name is the year of publication in parentheses, followed by a period. Magazines, newsletters, newspapers, and presentations also include the month (and, when necessary to pin down the particular source, the day).

In Figure 12.1, notice that two articles written by Kiewra as sole author were published in 1985. These have been listed alphabetically by article title, and the two articles are distinguished by designating the years as "1985a" and "1985b." Any citations in the text would then be either "Kiewra, 1985a" or "Kiewra, 1985b."

Title of the work. In APA style, the title of the article, book, or other source follows the publication year. If you are referencing an article using APA style, the title of the article is *not* italicized, but the title of the journal in which it appears *is* (see, for example, the reference for Kardash and Amlund in Figure 12.1). The title of a book is always italicized. So, too, is the title of a paper presentation (see the reference for Kiewra, Dennison, and Benton in Figure 12.1) or doctoral dissertation.

Be sure to pay attention to the rules for capitalization in whatever style manual you are using. Can you determine what APA's rules are from the entries in the sample? Did you notice that the first word in a book, article, or presentation title is the only one capitalized (unless the word is a proper noun or proper adjective or it follows a colon) but that all major words of a journal title are capitalized?

Publication information. For journal articles, publication information usually includes the volume number (which is italicized), issue, and page numbers. (If separate issues within each volume begin with sequentially numbered pages—for example, if the first issue of a particular volume ends on page 96 and the second issue begins on page 97—then the issue number can be omitted.) Publication information for a book includes the location and name of the publishing company or agency. In the case of a paper presented at a conference, the name of the conference and its location are provided.

Notice how this information is formatted in the sample entries in Figure 12.1. All redundancy is eliminated; there are no extra words such as *volume*, *issue*, and *pages*. By their specific location in a citation, readers understand their meaning. This practice eliminates many extra words; such a reduction means fewer manuscript pages, which translates into lower printing costs.

Notice, too, that references to a publisher are short and succinct, excluding such words as "Publishing Company" and "Publishers, Inc." These words add no new information and thus can be eliminated.



Referencing sources obtained on the Internet. Sources found on the Internet require additional information. Typically this includes either (a) the Internet address at which the document was found or (b) the Digital Object Identifier (DOI). You are almost certainly familiar with Internet addresses—in computer lingo, an Internet address is called a **Uniform Resource Locator**, or **URL**—and we have been using such addresses throughout the book to direct you to various Internet websites. For example, in APA style, an article in an online journal would be referenced using the following format:

Amrein, A. L., & Berliner, D. C. (2002, March 28). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, 10 (18). Retrieved from <http://epaa.asu.edu/epaa/v10n18>

Information found on a public or private organization's website—in this case, one that has no specific date attached—would be referenced using this format:

Statistics Canada (n.d.). *Census of Agriculture counts 3,795 farms in Nova Scotia*. Retrieved from http://www.statcan.ca/english/agcensus2006/media_release/ns.htm

The "n.d." in parentheses in the preceding reference means that the document cited has *no date*, year or otherwise, associated with its posting on the Internet.

Whereas a document's Internet address can change over time, its **Digital Object Identifier (DOI)** is more permanent. DOIs are a fairly recent development for Internet-based documents, and thus they may require some explanation. In particular, DOIs provide a system through which people can register online postings and, later, other people can find them. Even if the particular URL for an online research report changes—for instance, this might happen if a university or other research institution reorganizes its website—the DOI remains constant for the foreseeable future. DOIs are especially helpful when research reports and other scholarly works are available only in electronic form. For instance, they are used to identify journal articles that appear online before they appear in paper. Following is an example:

Wiers, R. W., Eberl, C., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2011). Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. *Psychological Science*. Advance online publication. doi:10.1177/0956797611400615.

People who see this entry in the researcher's reference list can track down the publication by going to the International DOI Foundation's website (www.doi.org) and then entering the article's DOI—in this case, 10.1177/0956797611400615—in the "Submit" box.

Potential resources in a literature review take a wide variety of forms—books, articles, government reports, presentations, posters, videos, website pages, blog postings, and so on—and the various style manuals have prescribed formats for each one of them. Thus, we strongly urge you to obtain an up-to-date version of the manual appropriate for your discipline and follow its prescriptions to the letter.

Appendix Content

Following the main report may be supplementary **appendices** (sometimes instead pluralized as *appendixes*) that may be helpful in understanding the research study more completely but are not absolutely essential to the comprehension of the body of the report. A rule of thumb is that *the material appearing in the appendix enables the reader to go further in understanding the method and/or results if so desired*. For instance, an appendix may include informed consent letters, questionnaires and other measurement instruments, response sheets, field notes, statistical computations, or extensive data tables.

In reporting research, nothing is hidden. All of the data are laid before the reader. The researcher's integrity is thereby preserved, and the results and conclusions of the study can be readily verified.

Organizing the Research Report

Research reports for most quantitative studies are similar in their organizational format. After their preliminary pages (title page, acknowledgments, table of contents, etc.), they typically have five major sections: an introduction (which includes the statement of the problem, assumptions, definitions of terms, etc.), a review of the related literature, a description of the methodology, a discussion of results, and conclusions (including implications and suggestions for future research). They will, of course, also have a list of references and, sometimes, one or more appendices.

Reports of qualitative studies are less predictable, and their specific organization may be highly dependent on the nature and design of the studies themselves.

As illustrations, we present the outlines for two of the dissertations from which we have presented excerpts in previous chapters. The first is a traditional outline, used for a quantitative, quasi-experimental study; the second is less traditional, used for a qualitative, grounded theory study. In the interest of space, we omit any subheadings that appear under the major headings.

Effects of Training in Self-Generation on the Quality of Students' Questions, Class Notes and Examination Scores (Jackson, 1996)

Front Matter

Copyright Notice
Title Page
Signature Page
Abstract
Acknowledgments
Table of Contents
List of Tables
List of Figures

Body of the Report

Chapter I. INTRODUCTION
Statement of the Problem
Purpose of the Study
Research Questions and Hypotheses
Limitations of the Study
Definitions and Terms

Chapter II. REVIEW OF THE LITERATURE

Part I: Adjunct Questioning Research

Part II: Self-Generated Questioning Theories and Models

Part III: Note-Taking

Summary

Chapter III. METHOD

Synopsis of the Pilot Study

Comparison of the Pilot Study with the Present Study

Subjects

Materials and Procedures

Examinations

Maintaining Confidentiality

Criteria for Coding Data

Chapter IV. RESULTS

Hypothesis One

Hypothesis Two

Hypothesis Three

Hypothesis Four

Hypothesis Five

Summary

Chapter V. DISCUSSION

Study Design Based on Contemporary Literature

Review of the Findings

Other Findings

Limitations

Educational Implications

Future Research

Summary and Conclusions

End Matter

References

Appendix A: Sample Test Questions

Appendix B: Table of Specifications for Unit 2

Appendix C: Sample of Questions from Treatment Group

Appendix D: Sample of Questions from Control Group

Appendix E: Syllabus with Initial Training Handout

Appendix F: Mid-semester Training Handbook for Treatment Group

Appendix G: Mid-semester Training Handbook for Control Group

Appendix H: Question Coding Form

Appendix I: Sample Page of Notes from Treatment Group

Appendix J: Sample Page of Notes from Control Group

Appendix K: Note Coding Form

Uncovering the Conceptual Representations of Students with Reading Disabilities (Zambo, 2003)**Front Matter**

Title Page

Signature Page

Abstract

Acknowledgments

Table of Contents

List of Tables

List of Figures

Body of the Report

Chapter I. REVIEW OF THE LITERATURE

Introduction [here the author presents the research problem]

	Conceptual Representations
	The Relationship Between Representation and Student Outcomes
	Evidence-Based Models of Reading
Chapter II.	METHODOLOGY
	Research Standards
	Pilot Study
	Design of the Study
	Data Analysis and Reporting of Data
Chapter III.	TEXT, CONTEXT, AND METACOGNITION
	Introduction
	Reasons and Context for Reading
	The Reading Process
	The Knowledge of Memory Strategies
Chapter IV.	AFFECTIVE AND MOTIVATIONAL FACTORS OF THE READER
	Reflections, Present Situation, and Future Expectations
	Motivation
	Theory of Intelligence and Epistemological Views
	The Teacher's Views of the Students as Readers
Chapter V.	DYSLEXIA AND THE DYSLEXIC BRAIN
	Thinking about the Brain
Chapter VI.	CONCLUSIONS AND FUTURE DIRECTIONS
	Future Directions

End Matter

References

Appendix A: Teacher Interview Questions

Appendix B: Alignment Between Research Questions and Student Interview Questions

Appendix C: Sample of Student Questions and Activities

Appendix D: Sample Drawings Completed by the Students

Appendix E: Internal Review Board Approval Letter and Permission to Copy Ruddell and Unrau's Model from the International Reading Association

Preparing the Report

The research report is precisely that—a *report*. The researcher is reporting on what he or she has done over the course of the research effort. In the process, the researcher is acquainting the reader with the problem, the data brought to bear on the resolution of the problem, the means employed in gathering those data, the ways in which those data were analyzed, and the conclusions reached.

Style of Prose

As noted earlier, various academic disciplines and their preferred style manuals have different standards with regard to style. For instance, the *APA Publication Manual* (2010) prefers that authors write in *active voice* (e.g., "A research assistant interviewed the participants," "Participants completed the survey") rather than *passive voice* (e.g., "The participants were interviewed by a research assistant," "Participants were given the survey"). It also recommends that *past tense* be used for the literature review, the description of research methods, and the presentation of the results; *present perfect tense* (e.g., "Theorists have suggested") is sometimes appropriate in the literature review, especially when making generalizations. When *interpreting* the results and presenting the conclusions, however, *present tense* is appropriate. If you are using a word processor,

the grammar checker in your software may allow you to check for inconsistencies in your use of active versus passive voice or past versus present tense.

Most research reports use a somewhat formal and impersonal style and format. There are exceptions to this rule, however; for example, ethnographic studies are sometimes described in a personal, literary, storytelling fashion (see Chapter 6), and historical researchers often tell a story as well (see Chapter 7). Perusing research reports in your own discipline can give you a good sense of the writing style that is most prevalent (and thus most acceptable) in your field.

Although writing in the third person ("The researcher analyzed the data") was the preferred style only a few years ago, increasingly researchers are using the first person to describe their procedures (e.g., "I instructed participants" in the case of a single author, "We analyzed the data" in the case of multiple authors). Use of the third person can be ambiguous; for instance, if you talk about yourself as "the researcher," it may give some readers the impression that you did not take part in your own study! Whichever style you use, however, it is critical that your point of view remain consistent throughout the report. If you are writing a thesis or dissertation, consult with your university's graduate school office about its preferred or required style. If you are writing a manuscript for publication in a particular journal, look for "Instructions to Authors" guidelines in the journal itself or on the publisher's website.

General Principles of Writing

Remember that your report is *your*. Whether or not you intend it to do so, a report can say a great deal about you to your readers. In preparing your report, therefore, several qualities should remain foremost in your mind: (a) neatness, consistency, and clarity; (b) precision; and (c) logical structure.

Neatness, consistency, and clarity. Your pages should be easy to read, with double-spaced lines and clean letters. In addition, you should format the text in a consistent manner, setting tabs for paragraph indents, bulleted lists, and the like, and setting the margins to control for line and page length. Most word processing software allows you to insert footnotes that will appear on the appropriate pages and to create tables that present numbers and text in perfect columns.

The way you set your words on paper, however, is much more than moving your fingers around on the keyboard. Most important is that you express your thoughts clearly and logically. Show how your data resolve your subproblems and how the subproblems resolve the main problem. Lead your readers through your own thought processes step by step. Such tactics provide evidence that you have approached the entire research endeavor in a thoughtful and systematic manner.

These types of things—some trivial and picayune, some more consequential—leave lasting impressions.

Precision. In addition to neatness, consistency, and clarity, a research report must be crisp with precision. There is no place in it for "sort-of's" and "I-guess-so's." This is not the time for ambiguous or foggy terms or half-stated conclusions. Your document should be clean-cut and present your ideas in a straightforward manner.

Every fact has its precise place within the research report structure. You may recall the weeks, the months, perhaps the years ago when you garnered a particular datum or cluster of data related to your study. You did not gather these as a child gathers wildflowers in a field—willy-nilly, without a guiding purpose. You gathered each datum because you envisioned it as a single building block that would support and advance the basic intent of all research: to enlighten a corner of the darkness related to a specific problem. We have come almost full circle, close to the place this book began. Everything you have done has been for a single purpose: to take one problem, one speck of darkness in human experience, and cause it to glow with the light of data and their interpretation, crisply and precisely.

Logical structure. A neat, consistent, clear, and precise report builds on the basic structure of your original research proposal. You may have wondered why we have recommended that you be

so fastidious and particular in drafting a research proposal. Our purpose has been to discipline your thinking and help you construct a logical plan upon which you can later build a clear, logical research report.

This construction of both a plan and a report cannot be done haphazardly. Every step must progress logically toward cracking open the shell of the enigma you chose at the beginning as the goal of your quest. You may understand better now why we titled Chapter 2 “The Problem: The Heart of the Research Process.”

After you have written several pages of your research report, go back and read your headings and subheadings. Do they form a logical whole? Do the various levels of heads accurately depict how different sections of text are interrelated? For instance, do they appropriately show that some sections are smaller subparts of other, larger sections? And overall, do the headings show a logical progression? They should. Recall the *inverted pyramid* idea in Chapter 3, “Review of the Related Literature.” There we counsel you to “begin your discussion . . . from a comprehensive perspective, like an inverted pyramid—broad end first. Then, as you proceed, you can deal with more specific ideas and studies and focus in more and more on your own particular [central research] problem.”

Such strategies are logic in practice. Logic comes from the Greek word *logos*, which means “reason, order, speech, word.” Think of your research proposal as a skeleton that holds the body of the report together, the framework on which everything else depends. Now, in your research report, adorn that skeleton with muscle: clear, convincing data. March through each of your subproblems, laying datum upon datum with as deliberate a precision as a mason cements masonry blocks in a skyscraper’s wall. Reason, order, and data, carefully placed and cemented with precise words, are the logic of the research report.

PRACTICAL APPLICATION Developing a Writing Schedule

Make no mistake about it, writing a research report—especially writing one *well*—takes considerable time and effort. A research report is not something you can whip up in a few days’ time. In the case of a lengthy report, such as a dissertation, you should plan on taking not several days, not several weeks, but several *months* to complete the report-writing process.

We authors can recall too many sad cases in which aspiring doctoral students completed all of the required coursework for their doctoral degrees, passed their written and oral comprehensive exams with flying colors, got approval for their dissertation projects, collected and analyzed their data, and then became “stuck” indefinitely in the process of writing their final dissertations. Some never got unstuck: They never finished their dissertations and therefore never received their doctoral degrees. Such a waste, we think! And so unnecessary!

To help you start *and finish* a lengthy research report—to grease your wheels and keep them greased so that you don’t get stuck somewhere along the way—we offer two pieces of advice. First, *develop a reasonable writing schedule for yourself*. Second, *stick to it!* The guidelines that follow can help you do both of these things.

GUIDELINES Developing a Workable Schedule

At this point, we urge you to go back and review the Practical Application “Writing Your Proposal” in Chapter 5. Most of the guidelines in “Writing the First Draft” and “Revising Your Proposal” are equally relevant to writing a final research report.

We now offer several additional suggestions. These have emerged from our own experiences in writing research reports (including our own doctoral dissertations) and other lengthy documents (including this textbook!).

1. *Identify small, easily accomplishable goals within the overall project.* A large research report, such as a dissertation, will seem less overwhelming if you break it into small, manageable pieces. These pieces might have such labels as “revision of the methods section,” “data analysis related to the first subproblem,” “implications section,” or “suggestions for future research.” Make each piece small enough that you can complete it within a week’s time or less.

2. *Set reasonable target dates for achieving each goal.* We strongly emphasize the word *reasonable* here. To get an idea of how much you can write in any given day or week, think about how long it has taken you to finish other lengthy writing projects. For instance, how long did it take you to complete your research proposal? How many pages could you write—and write *well*—in a day? (For instance, one of us authors has learned from experience that she can usually write, at most, about 8 to 10 double-spaced manuscript pages a day. After that, she’s essentially brain-dead until the following morning.)

Consider personal matters when you establish your schedule. Do you have a part-time or full-time job to consider? Do you have responsibilities to other family members that will take some of your time and energy? Have you built in adequate time for health and fitness, meals, shopping, home and car repairs, and occasional relaxation? You need to “get real” about how quickly you can complete various aspects of your writing project. Otherwise you will never stick to your schedule; you will be doomed to failure before you even start.

3. *Reward yourself each time you reach one of your goals.* Give yourself a treat of some sort after you successfully finish each piece of your report. Watch a movie, read a magazine or short novel, clean the house, play a few games of solitaire on your computer, surf the Internet—whatever you need to do to get refreshed and ready to tackle the next goal on your schedule.

4. *Seek regular feedback.* We’ve said it before and we’ll say it again: Ask others to give you honest feedback about what you say and how clearly you say it. Honest feedback now can save you more serious criticism—and, we might add, it can save you considerable aggravation and heartache—later on.

5. *Build time into your schedule for at least two or three revisions.* Most research reports are reviewed by others before they ever see the light of day. A committee of university faculty looks closely at any thesis or dissertation. An editorial review board carefully scrutinizes any manuscript submitted to a professional journal. The review process ensures that all approved research reports meet basic standards of scholarship, accuracy, and scientific rigor.

In the case of dissertations, *we have yet to see a report that has not had to undergo at least two revisions.* In fact, a doctoral student often completes four or five rewrites before defending a dissertation before a doctoral committee. Let’s face it: The researcher is very close to his or her research project and, at the end, is equally close to the report that he or she has written about the project. So close, sometimes, that omissions, errors, and logical inconsistencies that may be blatantly obvious to others are not at all obvious to the researcher. Furthermore, other people might have useful ideas about better ways to organize a discussion, suggestions about additional statistical analyses that may shed further light on the data, or new sources of literature that may be relevant to unexpected findings. The recommendations that others make, as well as the revisions that occur as a result of these recommendations, have one primary purpose: to make a research report the very best it can possibly be.

Furthermore, you must remember that any report bearing a stamp of approval from other individuals—whether that “stamp” takes the form of the signatures of a doctoral committee or the publication of a report in a scholarly journal—reflects not only on the author of the report but also on those who have approved the report. A poorly written research report makes a lot of people look bad.

The final stages of the writing project, especially the revisions, may seem to go on interminably. But persist! You have expended a great deal of time and effort in conducting your research project, and perhaps others have devoted considerable time and effort to it as well. It is only by completing your report that your project will ultimately contribute to the world’s knowledge about the topic you have studied.

PRACTICAL APPLICATION Critiquing a Research Report

Beck (1990) has developed a list of insightful questions that every researcher should answer satisfactorily before submitting a final version of a research report. Although she created it for students and professional colleagues in the field of nursing, we have adapted her list of questions to apply to research report writers in general. The following checklist, based on Beck's original list, should both help you evaluate the reports you read and serve as a guide as you assess your own writing.



CHECKLIST

Criteria for Critiquing a Research Report

STEP 1. THE PROBLEM

	YES	NO
Is the problem clearly and concisely stated?	___	___
Is the problem adequately narrowed down into a researchable problem?	___	___
Is the problem significant enough to warrant a formal research effort?	___	___
Is the relationship between the identified problem and previous research clearly described?	___	___

STEP 2. LITERATURE REVIEW

Is the literature review logically organized?	___	___
Does the review provide a critique of the relevant studies?	___	___
Are gaps in knowledge about the research problem identified?	___	___
Are important previous research studies relevant to the topic included in the literature review?	___	___
Are all cited works included in the reference list?	___	___
Are all works included in the reference list cited in the literature review or elsewhere in the report?	___	___

STEP 3. THEORETICAL OR CONCEPTUAL FRAMEWORK

Is the theoretical framework clearly applicable to the problem (as opposed to being a "stretch" in which the theoretical framework is only marginally relevant to the problem)?	___	___
If a conceptual framework is used, are the concepts adequately defined, and are the relationships among these concepts clearly identified?	___	___

STEP 4. RESEARCH VARIABLES

Are the independent and dependent variables operationally defined?	___	___
Are any confounding variables present? If so, are they identified?	___	___

STEP 5. HYPOTHESES

Are the hypotheses clear, testable, and specific?	___	___
Does each hypothesis describe a predicted relationship between two or more variables included in each hypothesis?	___	___
Do the hypotheses flow logically from the theoretical or conceptual framework?	___	___

STEP 6. SAMPLING

	YES	NO
Is the sample size adequate?	___	___
Is the sample representative of the defined population?	___	___
Is the method for selection of the sample appropriate?	___	___
Is any sampling bias in the method acknowledged?	___	___
Are the criteria for selecting the sample clearly identified?	___	___

STEP 7. RESEARCH DESIGN

Is the research design adequately described?	___	___
Is the design appropriate for the research problem?	___	___
Does the research design address issues related to the internal and external validity of the study?	___	___

STEP 8. DATA COLLECTION METHODS

Are the data collection methods appropriate for the study?	___	___
Are the data collection instruments adequately described?	___	___
Do the measurement tools have reasonable validity and reliability?	___	___

STEP 9. DATA ANALYSIS

Is the results section clearly and logically organized?	___	___
Is the type of analysis appropriate for the measurement scale (nominal, ordinal, interval, ratio) for each variable?	___	___
Are the tables and figures clear and understandable?	___	___
Is each statistical test an appropriate one for answering the research question?	___	___

STEP 10. INTERPRETATION AND DISCUSSION OF THE FINDINGS

Does the investigator clearly distinguish between actual findings and interpretations?	___	___
Are the interpretations based on the data obtained?	___	___
Are the findings discussed in relation to previous research and to the conceptual/theoretical framework?	___	___
Are all generalizations warranted and defended?	___	___
Are the limitations of the results identified?	___	___
Are implications of the results discussed?	___	___
Are recommendations for future research identified?	___	___
Are the conclusions justified?	___	___

Beyond the Unpublished Research Report: Presenting and Publishing

If you have completed an unpublished research report, such as a master's thesis or doctoral dissertation, consider this: In most cases, only a few people will ever read your report in its current form. If your research project has uncovered new information, new ideas, and new understandings that can make a significant contribution to the world's body of knowledge about a particular topic, then we urge you to seek a broader audience. Two excellent ways to do this are conference presentations and journal articles.

Conference Presentations

Many researchers present their research findings at regional, national, or international conferences. Some conferences are annual or biennial meetings sponsored by societies related to particular academic disciplines (e.g., American Sociological Association, European Association for Research on Learning and Instruction, Modern Language Association). Others are more specific to particular interest areas (e.g., family violence, Piaget's theory of child development). The organizers of many of these conferences eagerly seek presentations (often called *papers*) from new researchers as well as from more experienced ones. Some conferences also include *poster* sessions, which (as the name "poster" implies) involve visual displays of research projects on large (perhaps 4-foot-by-6-foot) bulletin boards. Typically, one or more of each poster's authors is present at the poster session to describe the project undertaken and answer questions.

If you would like to present a paper or poster at a professional conference, you will probably need to submit a *paper proposal* several months in advance to the association or institution sponsoring the conference. These proposals are usually much shorter than the research proposals described in Chapter 5. Furthermore, their purpose is different: You are submitting a proposal to present a research project that you either (a) have already completed or (b) are currently conducting but will definitely have completed before the conference.

Proposals for paper and poster presentations are often only two or three pages in length. Their specific format varies considerably from one professional group to another, and we urge you to consult the *call for papers* that invites proposals for conference presentations. Regardless of the format, one thing is true for all paper proposals: They need to be written with the same clarity and academic rigor required for any research proposal or research report. Furthermore, they need to adhere faithfully to the guidelines that conference sponsors specify.

PRACTICAL APPLICATION Presenting Your Research at a Professional Conference

Many novice researchers find a conference presentation to be a highly anxiety-arousing experience. It's quite common to have some stage fright, especially when presenting a paper to a large audience. You might find it a bit reassuring to know that even many renowned and well-respected scholars still get nervous when they must speak in front of a large group of peers. Knowing you're in good company won't make your public-speaking jitters go away, of course, but it will at least help you realize that you're simply feeling as most people do in such a situation.

The best way to keep your presentation jitters under control is to be well prepared for your paper or poster. Here we offer a few guidelines that can help you put your best foot forward—and thus can give you a confidence boost—in presenting your research project.

GUIDELINES

Presenting an Effective Paper or Poster

Drawing on recommendations by Munter and Paradi (2009) and Nicol and Pexman (2010), as well as on our own experiences, we offer several suggestions for presenting papers and posters at professional conferences:

1. *Be concise and to the point.* If you give a paper, you are likely to have only 10 to 20 minutes to describe what you have done. If you present a poster, your text (including font size) and graphics should be sufficiently large that other people can readily see them from at least two feet away. In either situation, you won't have the time (in the case of a paper) or the space (in the case of a poster) to describe every detail of what you have done and learned. Instead, present those aspects of your project that are key to your audience's understanding of what you have accomplished, including:

- The title of your presentation, plus your name, affiliation, and contact information
- Your research problem and, if applicable, your hypotheses
- A general rationale and context for your study
- A general description of your design and methodology (including the nature and size of your sample)
- Results that are most central to your research problem and hypotheses
- Your interpretations of and conclusions from your data

Many posters also include a one-page abstract immediately after the title page, plus a short list of cited references at the end.

2. *Prepare polished, professional-looking visuals in advance.* One widely used tool for both papers and posters is Microsoft PowerPoint, with which you can create a wide variety of eye-catching visuals—PowerPoint calls them *slides*—including bulleted lists, charts, and graphs. PowerPoint also allows you to incorporate photographs, scanned documents, and (in the case of paper presentations) short videos into your presentation. If you are presenting a paper, you simply hook up your laptop computer to an LCD (liquid crystal display) projector, which the conference organizers typically provide in each presentation room, and click on a mouse or wireless clicker to advance to successive slides.² If you are presenting a poster, you print out the slides, either on individual sheets of paper or on a single large poster sheet; an Internet search of “poster presentation template” can give you numerous companies that can convert PowerPoint slides into a high-quality poster.

However, we urge you not to clutter up your presentation or poster with too many visual effects, such as distracting and irrelevant images and animations. Simple images and animations—for instance, having individual bullet points “march” across the screen as you introduce them (which PowerPoint lets you do)—can catch and hold people's attention and so are quite appropriate.

3. *Practice ahead of time, but don't overdo it.* Especially if you are giving a paper, it's a good idea to rehearse it at least once, if only in order to time yourself to make sure you can keep your presentation to the prescribed time limits. It's helpful, too, to review your speaking notes within an hour or two of your presentation. However, we do *not* recommend that you either read or memorize your presentation; by doing these things, you will come across as a mindless robot. Instead, you want to convey the impression that you know your project and subject matter *very well*—something you can do only if you talk somewhat extemporaneously about what you have done.

4. *Prepare handouts that summarize or complement your presentation.* Handouts can take a variety of forms. For instance, your handout might provide small versions of your PowerPoint slides (in the “Print” feature, PowerPoint lets you specify how many slides you print on a single page;

²Most PC laptops have a standard outlet for hooking up to an LCD projector. In our experience, however, Macintosh laptops sometimes require special adaptors; check with your local Apple store to be sure you have the appropriate adaptor for your own laptop.

for readability's sake, we suggest three to six slides per page). If your paper or poster is based on a manuscript you have submitted for possible publication, your handout might be a copy of that manuscript. Sometimes presenters instead hand out a slip of paper giving the address of a website at which audience members can download the paper or poster. And in some instances—this is most often the case when someone hasn't planned ahead!—a presenter takes people's e-mail addresses and sends the desired materials after returning home.

5. *Anticipate and be prepared to answer questions.* By their very nature, poster sessions give your audience a chance to ask questions, and most paper sessions also include time for audience members to ask questions. To the extent that you can do so, you should try to anticipate questions and bring any supplementary materials that might help you answer them. But you shouldn't expect that you will be able to answer every question someone might ask. It's quite acceptable—in fact, it's a sign of a candid and open-minded researcher—to respond to some questions by saying, "You raise a good point that I hadn't considered" or "Unfortunately, my study wasn't able to address that particular concern."

6. *Make connections with your audience, including connections you can follow up on after the conference.* Regardless of whether you are presenting a paper or a poster, present *yourself* as someone who is approachable and eager to exchange ideas. Smile, make eye contact, and in other ways convey the message that you want to hear other people's ideas, concerns, and suggestions. And if you don't already have them, get business cards printed that include, at a minimum, your name, credentials, affiliation, and e-mail address.

Journal Articles

An even more effective—and certainly more permanent—way to disseminate your findings and interpretations is to submit a research report to an academic journal. The guidelines we have presented in this book should get you well on your way to writing a manuscript for submission to a research journal, but once again we urge you to *be concise*. As a rule, journal space is at a premium, so journal editors have little tolerance for researchers who say in 100 words what they could have said in 10.

Before you submit a manuscript to a particular journal, read several recent issues of the journal to make sure it is the right place for your article. Determine whether the journal includes research reports, including reports about your general topic, among its articles. Also look at the style of writing that is typical in the journal; you will want to use a similar style in any manuscript you submit. And (forgive us for saying this one more time) seek critical feedback from others about your manuscript, including from people who have published in that journal or similar ones, and use their suggestions to revise and strengthen what you have written.

Sharing Authorship

Whether you are presenting a paper at a conference or submitting a manuscript to a research journal, you must, of course, determine whether you should be sole author or share authorship with one or more other individuals. For example, in the case of master's theses and doctoral dissertations, students often share authorship with their major advisors and perhaps with one or two other faculty members as well.

A general rule of thumb is this: *Individuals who have made significant intellectual contributions to the work should share in its authorship.* Typically, any co-authors have been actively involved in the conceptualization, design, execution, and/or in-depth analysis of the research project. Multiple authors are usually listed in an order indicating which individuals have made the most substantial contributions.

People who have assisted with data collection, coding, computer programming, simple statistical analysis, typing, minor editing, and so on, but who have not contributed *intellectually* to the work, usually do not warrant authorship (Elliott & Stern, 1997; McGue, 2000). Nor do people who have reviewed a paper or manuscript and given their suggestions for how the author(s) might improve it. Such minor contributions are more appropriately acknowledged in a footnote or endnote.

Sharing authorship with others who have contributed in important ways to your research project and listing coauthors in an order that acknowledges their relative contributions are two additional dimensions of the “honesty with professional colleagues” issue we mentioned in our discussion of ethics in Chapter 4. Not only must researchers be honest with their colleagues about what they have done and what they have found but they must also be honest about who has helped them with their research endeavors.

Responding to Reviewers’ Critiques

Throughout the book we have offered many guidelines and checklists that, in one way or another, should strengthen either your methodology or your writing and thus, we hope, should help you successfully present and publish your research. Yet not every proposal gets accepted for a conference presentation, and not every manuscript gets accepted for publication in a professional journal. Many proposals and manuscripts get rejected for good reasons. But in some instances, conference program chairs and journal editors simply don’t have the space for every good research report that comes their way.

Rejection letters are always disappointing, but we urge you not to let them discourage you. (Even very experienced researchers get rejection letters on occasion.) Typically program chairs and journal editors have had one or more people in the field review your submission, and they are apt to include the reviewers’ comments when they give you the bad news. Put these reviews aside for a few days—at least long enough that your frustration and disappointment have dissipated a bit—and then try to look objectively and dispassionately at what the reviewers have had to say. Occasionally a reviewer will be downright nasty, but more often reviewers will have constructive criticisms that can help you strengthen what you have written. You can then resubmit your proposal to another conference or send your manuscript to another journal.

In general, we urge you to *persist* in your efforts to get out the word—and to as broad an audience as possible—about what you have done!

A Closing Thought

As you prepare to write your research report, you might do well to read two books: *The Elements of Style* (Strunk & White, 2009) and *The Art of Readable Writing* (Flesch, 1974). Avoid the overly exaggerated expression; look sharply at your ill-advised and thoughtlessly chosen adjectives. Stick to the facts. Report them accurately but, in so doing, enliven your prose with variety in sentence structure, sentence length, and precision of verb and noun.

More do’s and don’ts at this point are probably “TMI”—too much information. We leave you with this last thought. Distilled into a brief stanza by an anonymous hand is a broad guideline for all your writing. Follow it.

The written word
Should be clean as bone:
Clear as light,
Firm as stone;
Two words are not
As good as one.

For Further Reading

- American Psychological Association (APA). (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.
- American Psychological Association. (2010). *Concise rules of APA style* (6th ed.). Washington, DC: Author.
- Barzun, J. (2001). *Simple and direct: A rhetoric for writers* (4th ed.). New York: HarperCollins.
- Boice, R. (1993). Writing blocks and tacit knowledge. *Journal of Higher Education*, 64(1), 19–54.
- Chicago manual of style* (16th ed.). (2010). Chicago: University of Chicago Press.
- Cone, J. D., & Foster, S. L. (2006). *Dissertations and theses from start to finish: Psychology and related fields* (2nd ed.). Washington, DC: American Psychological Association.
- Council of Science Editors. (2006). *Scientific style and format: The CSE manual for authors, editors, and publishers* (7th ed.). Reston, VA: Author.
- Day, R. A. (1995). *Scientific English: A guide for scientists and other professionals* (2nd ed.). Westport, CT: Greenwood.
- Day, R. A., & Gastel, B. (2006). *How to write and publish a scientific paper* (6th ed.). Westport, CT: Greenwood.
- Gopen, G. D., & Swan, J. A. (1990). The science of scientific writing. *American Scientist*, 78, 550–558.
- Hacker, D. (2007). *A writer's reference* (6th ed.). New York: Bedford/St. Martin's.
- Hacker, D. (2008). *A pocket style manual* (5th ed.). New York: Bedford/St. Martin's.
- Hacker, D. (2008). *Rules for writers* (6th ed.). New York: Bedford/St. Martin's.
- Henson, K. T. (2005). *Writing for scholarly publication*. Boston: Allyn & Bacon.
- Locke, D. (1992). *Science as writing*. New Haven, CT: Yale University Press.
- Madigan, R., Johnson, S., & Linton, P. (1995). The language of style: APA style as epistemology. *American Psychologist*, 50, 428–436.
- Matthews, J. R., & Matthews, R. W. (2007). *Successful scientific writing: A step-by-step guide for the biological and medical sciences* (3rd ed.). New York: Cambridge University Press.
- McGue, M. (2000). Authorship and intellectual property. In B. D. Sales & S. Folkman (Eds.), *Ethics in research with human participants* (pp. 75–95). Washington, DC: American Psychological Association.
- McQuain, J. (1996). *Power of language: Getting the most out of your words*. Boston: Houghton Mifflin.
- Meloy, J. M. (2001). *Writing the qualitative dissertation: Understanding by doing* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Modern Language Association. (2008). *MLA style manual and guide to scholarly publishing* (3rd ed.). New York: Author.
- Morris, D. (1994). *Guidelines for writing a qualitative research report*. Chicago: American Marketing Association.
- Munter, M., & Paradi, D. (2009). *Guide to PowerPoint*. Upper Saddle River, NJ: Pearson/Prentice Hall.
- Nicol, A. A. M., & Pexman, P. M. (2010). *Presenting your findings: A practical guide for creating figures, posters, and presentations* (6th ed.). Washington, DC: American Psychological Association.
- Seely, J. (2005). *Oxford guide to effective writing and speaking*. Oxford, England: Oxford University Press.
- Tuckman, B. W. (1990). A proposal for improving the quality of published educational research. *Educational Researcher*, 19(9), 22–25.
- U.S. Government Printing Office. (2008). *Style manual* (30th ed.). Washington DC: U.S. Author.
- Westmeyer, P. M. (1994). *A guide for use in planning, conducting, and reporting research projects* (2nd ed.). Springfield, IL: Charles C Thomas.
- Wilkinson, A. (1991). *A scientist's handbook for writing papers and dissertations*. Upper Saddle River, NJ: Prentice Hall.
- Wolcott, H. F. (2009). *Writing up qualitative research* (3rd ed.). Thousand Oaks, CA: Sage.

Now go to MyEducationLab at www.myeducationlab.com to take assessments to evaluate your mastery of chapter concepts. Assignments and Activities exercises and Building Skills in Reading Research exercises are also available to help you master concepts and skills. Feedback for many of these exercises is provided so that you can see why your answers are correct or incorrect.

Appendix A

Using a Spreadsheet: Microsoft Excel

On the surface, an electronic spreadsheet looks like nothing more than a table with numerous rows and columns. But the typical features of spreadsheet software enable you to do many things with the data you put in the table. For example, you can recode your data, reorganize it in various ways, and perform simple calculations on subsets of data that you designate.

Here we look at one widely used spreadsheet, Microsoft Excel. We should point out that Excel's format may differ slightly depending on whether it is used with Microsoft Windows or a Macintosh operating system. Also, each new update of Excel tends to be slightly different from its predecessor in appearance and function. We are basing this discussion on the 2008 version of Excel for Macintosh computers.

Using Excel to Keep Track of Literature Resources



In any literature review, you are likely to draw on a variety of resources, probably including books, journal articles, and Internet websites. You need to keep track of and report different information about each kind of resource. For a book, you need to know the author(s) or editor(s), title, publication date, publisher, and the publisher's location; in order to find the book in the library stacks, you also need its call number. For a journal article, you need to know the author(s), titles of both the article and the journal, publication date, volume number (and perhaps issue number), and page numbers. The information you need for an Internet website is apt to vary depending on the nature of its content, but at a minimum you need to record the Internet address (uniform resource locator, or URL) and date on which you retrieved the document; or, for very recently posted documents, the document's digital object identifier (DOI).

Let's organize such information with Excel by going to the "File" menu and creating a "New Workbook." An empty two-dimensional table appears on the screen, with tabs labeled "Sheet 1" and "+" at the very bottom. We'll use Sheet 1 to keep track of books. By clicking on the top left-hand cell in the table, we can insert the word "BOOKS" in uppercase letters. Then, by hitting the *down* arrow key on the keyboard, we move to the cell just below, where we insert the words "Authors/Editors" (never mind for now that the words may appear to spill over into the second column—appearances to the contrary, all words typed in any single cell remain in that cell). Then, we hit the *right* arrow key on the keyboard, move to the cell to the right, and insert the word "Title." We continue moving to the right four more times, inserting the words and phrases "Date," "Call Number," "Publisher," and "Pub. Loc." (short for "Publisher's Location"). The words and phrases we have just entered in Row 2 will be our headings for the columns.

At this point some of our headings are too long for the cells, so let's do two things. First, let's go to the "File" menu and then to "Page Setup" and click on "Landscape" and "OK." By doing this, we turn the page sideways and give ourselves more room across the page. Second, let's now move the cursor to the very top of the screen, where we see alphabet letters labeling the columns. If we move the cursor to the line separating the A and B columns, a cross-with-arrow-points icon appears. By clicking on the mouse at this point, we can drag the line to the

right to make the “Authors/Editors” column wider. We can do the same thing for the other columns as well, in each case adjusting column width to accommodate the column heading or kind of information we expect to insert in the column cells. To make our headings more visible, we’ll also put them in boldface by going to the “Format” menu, then “Cells,” and then “Font” and clicking on “Bold” and “OK.” With such steps we’ve set up our list for keeping track of books.

Let’s now click on the “+” tab at the very bottom of the page. Doing so gives us “Sheet 2,” where we can follow a similar procedure for journal articles. Again let’s set up the page in *landscape* mode. In Row 1 we can insert “ARTICLES” and then in the first six cells of Row 2 we can insert the headings “Authors,” “Article Title,” “Journal Title,” “Date,” “Vol/Iss” (for “volume and issue”), and “Pp.” (for “page numbers”). As we did on Sheet 1, we can adjust the column widths and boldface our headings. If we create a Sheet 3 for Internet websites, we need columns labeled “Address” and “Date Retrieved,” plus possibly additional columns in which to insert names of authors or organizations, titles, posting dates, DOIs (if available), and other pertinent information.

Our workbook of three spreadsheets is now ready for us to enter information about our various library and Internet resources. We can print out the sheets and add the necessary information in pen or pencil or, better still, we can take a laptop or tablet computer with us to the library and insert the information directly into a computer document. Figure A.1 shows how the three spreadsheets might look for a few resources on the topic of schizophrenia. Notice that some of

BOOKS					
Authors/Editors	Title	Date	Call Number	Publisher	Pub. Loc.
Noll, R.	Encyclopedia of Schizophrenia and Other Psychotic Disorders	2007	RC514 .N63 2007	Facts on File	New York
Walker, E. F. (Ed.)	Schizophrenia: A Life-Course Developmental Perspective	1991	RC514 .S3342 19	Academic Press	San Diego
Frith, C., & Johnston	Schizophrenia: A Very Short Introduction	2003	RC514 .F755 200	Oxford University Press	Oxford, England

ARTICLES					
Authors	Article Title	Journal Title	Date	Vol/Iss	Pp.
Lublin, H., & Eberhard, J.	Content versus delivery: Challenges	European Neuropsychopharmacology	2008	18(Suppl 3)	v-vi
Tabarés-Seisdedos, R.	Neurocognitive and clinical predictors	Journal of Affective Disorders	2008	109(3)	286-299
Schwab, S. G., & Wilden	Research on causes for schizophrenia	Schizophrenia Research	2008	102(1-3)	29-30

WEBSITES					
Address	Date Ret'd	Author/Org.	Title	Date Posted	Other Info.
www.nimh.nih.gov/health/to	9/15/08	NIMH/	Schizophrenia	4/2/08	
www.nimh.nih.gov/medlineplus	9/17/08	NIM/NIH	Schizophrenia	no date	
www.schizophrenia.com/diac	9/18/08	NARSAD	Schizophrenia symptoms	no date	

FIGURE A.1

Using Excel to keep track of library and Internet resources

the entries (e.g., some book titles) are too long for the column width. No matter, because the entries *are* recorded in their entirety in the spreadsheet document, and clicking on their particular cells will bring them into full view. Notice, too, that the entries in the “Address” column for the WEBSITES spreadsheet are in blue. When you type an Internet address into a cell, Excel automatically makes it a *hyperlink*: If your computer is currently online and you click on the cell, your computer will take you to that website.

Once you have entered various library and Internet resources into your spreadsheets, you can organize them in various ways, perhaps by call number (for books), journal title (for articles), or date retrieved (for websites). We will look at how to organize spreadsheet entries in an upcoming section. Although our focus at that point will be on organizing data, the same general organizational strategies apply to *any* kind of information.

Using Excel to Record and Recode Data

As you have seen in the preceding section, the data you enter in the cells of an electronic spreadsheet can take a variety of forms: text, numbers, dates, and so on. Thus you can use a spreadsheet to keep track of the information you collect from a qualitative study (provided that the text entries are relatively short), a quantitative study, or a mixed-methods design.

For illustrative purposes, we’ll use hypothetical data from a descriptive quantitative study. We return to the four rating-scale items for risk taking presented in Chapter 8:

		Not at All True	Somewhat True	Very True	
11.	I would prefer to teach in a way that is familiar to me rather than trying a teaching strategy that I would have to learn how to do.	1	2	3	4 5
16.	I like trying new approaches to teaching, even if I occasionally find they don’t work very well.	1	2	3	4 5
39.	I would choose to teach something I knew I could do, rather than a topic I haven’t taught before.	1	2	3	4 5
51.	I sometimes change my plan in the middle of a lesson if I see an opportunity to practice teaching skills I haven’t yet mastered.	1	2	3	4 5

As you may recall from our discussion in that chapter, the researchers included these items in a longer list of items designed to assess a variety of traits in college education majors who were completing their teaching internship year (Middleton, Ormrod, & Abrams, 2007). Let’s consider how we might create a spreadsheet to enter the data for participants’ responses to the entire survey. The general convention is to assign each *row* in the spreadsheet to a particular participant and to assign each *column* to a particular variable that we have assessed for each participant. In this research project Middleton and his colleagues included several demographic variables (e.g., age, gender), supervisor ratings of teacher effectiveness, and participants’ responses to 69 rating-scale items designed to measure several personality and motivational characteristics. For simplicity’s sake, we’ll limit ourselves to the 4 rating-scale items just presented plus 4 additional rating-scale items designed to measure perfectionism, as follows:

		Not at All True		Somewhat True		Very True
19.	It is very important that I always appear to be "on top of things."	1	2	3	4	5
27.	It does not bother me if I occasionally make mistakes in the classroom.	1	2	3	4	5
38.	I do not want people to see me teaching unless I am very good at it.	1	2	3	4	5
60.	I always try to present a picture of perfection in my teaching.	1	2	3	4	5

We'll create a spreadsheet for a sample of 10 hypothetical respondents to the questionnaire and their responses to the 4 risk-taking and 4 perfectionism items (see Figure A.2). Note that the labels "RISK-TAKING" and "PERFECTIONISM" are only in cells B1 and F1, respectively, but because cells to their immediate right are blank, we see the content of these cells in their entirety.

Can we add up the responses to the four risk-taking items to create an overall risk-taking score and, similarly, add up responses to the four perfectionism items to create an overall perfectionism score? No, not yet. If you look at the wordings for the eight items, you should notice that the self-descriptions in Items 16 and 51 are indicative of high risk taking but the self-descriptions in Items 11 and 39 are indicative of *low* risk taking. Similarly, Items 19, 38, and 60 reflect a desire for perfection, but Item 27 reflects comfort with *imperfection*. In order to have responses to all items for a particular characteristic reflect a high degree of that characteristic, we need to reverse, or *recode*, people's responses to Items 11, 39, and 27, changing 1s into 5s, 2s into 4s, 4s into 2s, and 5s into 1s, but leaving 3s as they are (e.g., in their recoded form, higher-number responses to Item 11 indicate high rather than low risk taking). The following simple formula makes this conversion for us:

$$6 - \text{Original response} = \text{Recoded response}$$

For example, if we want to recode a response of 5, then

$$6 - 5 = 1$$

Similarly, if we want to recode a response of 2, then

$$6 - 2 = 4$$

FIGURE A.2

Hypothetical data for 10 people responding to eight rating-scale items related to risk taking and perfectionism

	RISK-TAKING				PERFECTIONISM			
Person #	#11	#16	#39	#51	#19	#27	#38	#60
1	2	5	3	4	3	5	3	2
2	4	1	5	2	4	2	5	4
3	5	2	3	3	5	3	4	4
4	2	4	2	4	4	4	3	3
5	1	5	3	5	2	5	1	2
6	3	2	5	1	5	2	5	3
7	5	1	3	2	4	3	4	3
8	1	3	1	5	4	5	2	3
9	4	1	3	2	5	3	5	4
10	3	3	2	5	3	5	3	2

In Column J of our spreadsheet, we will make a new column, which we'll label "Rev11" (for "Reverse of Response to #11"). We're going to use a *formula* to create the values in this column. In particular, let's click on the first cell below our "Rev11" column heading (this is the cell for Person #1). We type an equals sign (=), followed by a 6 and a minus sign. Before doing anything else, we move the cursor to the cell containing Person #1's response to Item #11 (where we see a response of "2" for the item) and click on that cell. What we see in the Rev11 cell for Person #1 is the following:

$$= 6 - B3$$

We immediately press the Enter or Return button on the keyboard—we must press this button before we do anything else—and Excel executes the formula to give us the desired value of 4. Now here's the cool part: We can click on the cell in which we've just entered a formula, "copy" its contents, and then "paste" the contents into the nine cells immediately below in the same column. What appears in each cell is the result of the same calculation *using the appropriate value for each person in our sample*. For example, Person #2's response of "4" has been recoded as "2," and Person #3's response of "5" has been recoded as "1."

Items 39 and 27 need to be recoded as well. Let's label Columns K and L "Rev39" and "Rev27" (for Items 39 and 27, respectively) and use the same procedure we used in the "Rev11" column. This time, however, after typing "=6-" in the cell below the new column heading, we click on the cell immediately below the heading "#39" or "#27," depending on which item responses we're recoding. The spreadsheet with the three new columns is shown in Figure A.3

We are now ready to compute overall scores for our risk-taking and perfectionism items. Let's create yet another column in the spreadsheet and label it "RtScore" (for "Risk-Taking Score"). We can again use the *formula* tool, this time adding together each person's responses in the #16, #51, Rev11, and Rev39 columns. We take the following steps:

1. Click on Person #1's cell in the new column.
2. Hit the equals sign key (=) on the keyboard.
3. Click on the first cell below the "#16" label.
4. Hit the plus sign key (+) on the keyboard.
5. Click on the first cell below the "#51" label.
6. Hit the plus sign key (+) on the keyboard.
7. Click on the first cell below the "Rev11" label.
8. Hit the plus sign key (+) on the keyboard.
9. Click on the first cell below the "Rev39" label.

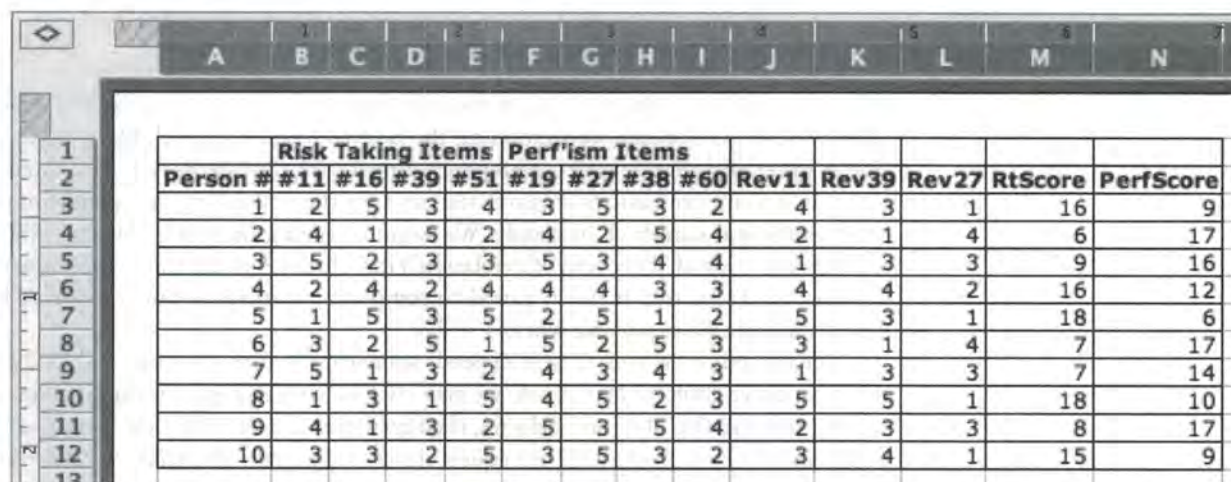
At this point, the entry in the cell you're creating should look like this:

$$= C3 + E3 + J3 + K3$$

FIGURE A.3

Adding three columns for reversing people's responses to certain rating-scale items

	Risk Taking Items				Perf'ism Items						
Person #	#11	#16	#39	#51	#19	#27	#38	#60	Rev11	Rev39	Rev27
1	2	5	3	4	3	5	3	2	4	3	1
2	4	1	5	2	4	2	5	4	2	1	4
3	5	2	3	3	5	3	4	4	1	3	3
4	2	4	2	4	4	4	3	3	4	4	2
5	1	5	3	5	2	5	1	2	5	3	1
6	3	2	5	1	5	2	5	3	3	1	4
7	5	1	3	2	4	3	4	3	1	3	3
8	1	3	1	5	4	5	2	3	5	5	1
9	4	1	3	2	5	3	5	4	2	3	3
10	3	3	2	5	3	5	3	2	3	4	1



	Risk Taking Items				Perf'ism Items								
Person #	#11	#16	#39	#51	#19	#27	#38	#60	Rev11	Rev39	Rev27	RtScore	PerfScore
1	2	5	3	4	3	5	3	2	4	3	1	16	9
2	4	1	5	2	4	2	5	4	2	1	4	6	17
3	5	2	3	3	5	3	4	4	1	3	3	9	16
4	2	4	2	4	4	4	3	3	4	4	2	16	12
5	1	5	3	5	2	5	1	2	5	3	1	18	6
6	3	2	5	1	5	2	5	3	3	1	4	7	17
7	5	1	3	2	4	3	4	3	1	3	3	7	14
8	1	3	1	5	4	5	2	3	5	5	1	18	10
9	4	1	3	2	5	3	5	4	2	3	3	8	17
10	3	3	2	5	3	5	3	2	3	4	1	15	9

FIGURE A.4

Adding two columns with overall scores for risk taking and perfectionism

Immediately hit the Enter or Return button and—*voilà!*—the value 16 should appear. Now copy the contents of this cell into the nine cells below it in the column, and you'll see totals ranging from 6 to 18 for the risk-taking items.

We can follow essentially the same procedure to create a total (which we'll call "PerfScore") for the perfectionism items, this time using the values in the #19, #38, #60, and Rev27 columns. The results of our calculations are shown in Figure A.4.

Reorganizing Data in Excel

An additional feature of virtually all spreadsheets is an ability to organize the data by one or more variables. Our current spreadsheet for responses to risk-taking and perfectionism items is organized by person number. But perhaps, instead, we want to organize it by risk-taking score, with the greatest risk takers listed first and the relatively nonrisk-taking people listed last. We first need to use the cursor to highlight all of the data we want to reorganize—in this case the 10 rows and 14 columns of numbers. We move the cursor to the third cell in the first column (for Person #1), click on the mouse, and then drag the mouse down and to the right until the 140 cells with numbers are all highlighted. We then move the cursor to the "Data" menu and select "Sort." A box appears in which we can sort by several variables in order of priority, but in this situation we want to sort only by risk-taking score. We type "RtScore" in the first box and, because we want to have the high risk takers appear at the top, we click "Descending" (for descending order). When we click on the "OK" button, the data rearrange themselves, with Persons #5 and #8 (with risk-taking scores of 18) appearing first and Person #2 (with a risk-taking score of 6) appearing last.

A word of caution, however. Be sure that you highlight *all* of the data columns in your spreadsheet before hitting the "OK" button. If you highlight only some of them (or perhaps only one or two), you will reorganize the data *only* in those columns, leaving the data in other columns untouched. The result will be a scrambled mess, with some numbers for, say, Person #8 moving to a new row and others staying where they were originally.

The *sort* tool isn't limited to numerical data. For example, let's return to the spreadsheets we created for the books, journal articles, and websites in our literature review. We could easily sort our books by call number or our journal articles alphabetically by journal title, thereby making our search for them in the library stacks more efficient.

Using Excel to Perform Simple Statistical Analyses

When we used formulas to recode some item responses and to compute overall scores for risk taking and perfectionism, we were using the *function* feature of Excel. Many functions are available in Excel, including numerous preprogrammed statistical analyses. For example, let's say that we want to compute basic descriptive statistics for the risk-taking and perfectionism scores for our hypothetical sample of 10 people. We begin by typing the labels "Mean," "SD" (for "Standard Deviation"), and "Corr" (for "Correlation") in Column A in the three cells immediately below our data set. (This step is not required to complete our mission, but it helps us keep track of which statistics we're putting where.)

The procedure we follow next depends somewhat on the particular version of Excel we are using. In Excel 2008 for Macintosh, we now click on the cell representing the intersection of the "Mean" row and the "RtScore" column, then go to the "Insert" pull-down menu at the top of the screen and click on "Function." An equals sign (=) appears in the table cell we've selected and a function box appears on the screen; this box includes many possible calculations we might perform. In the function box, we scroll down to "AVERAGE" (we may possibly have to scroll a long way until we reach the category "Statistical") and *double-click* on AVERAGE. At this point we need to tell Excel *which* numbers—which in this case Excel calls "arguments"—to use in calculating the average (mean). The bottom portion of the function box presents two places where we can indicate the range of numbers we want to use in calculating the mean; for the mean, we want to use only the first of these two places. Excel may also "suggest" one or more table cells with a colored box; if it doesn't, we can create a box by clicking on one of the cells in our spreadsheet. Then, by clicking on various sides and/or corners of the box and dragging the box in appropriate directions, we can capture the numbers to be averaged—and *only* those numbers—at which point we again hit the Enter or Return button on the keyboard. In the example here, we capture the 10 RtScore values for our 10 people, and the mean risk-taking score (12) for our sample appears in the designated cell. We follow a similar procedure for the "PerfScore" column to obtain a mean Perfectionism score (12.7).

We do essentially the same thing to obtain a standard deviation for our two sets of scores, this time clicking on the appropriate cells in the "SD" row of our spreadsheet and double-clicking "STDEV" in the right-hand column in the function box. This procedure gives us standard deviations of 4.9889 and 4.056545 for the Risk-Taking and Perfectionism scores, respectively.

Finally, let's calculate a Pearson product-moment correlation coefficient between the Risk-Taking and Perfectionism scores. This calculation requires a slightly different procedure. We must first click on the cell where we want the *r* value to appear, so let's use the cell representing the intersection of the "Corr" row and the "PerfScore" column. We choose the function feature as we did for means and standard deviations, then double-click on "CORREL." At this point a box appears that asks for "Array1" and "Array2." With the Array1 subbox highlighted, we highlight the 10 data cells in the "RtScore" column of the spreadsheet (we must be sure *not* to highlight the mean and standard deviation we've already calculated). We then move the cursor to the Array2 subbox, click on it and then highlight the 10 data cells in the "PerfScore" column of the spreadsheet. What we will see in the two subboxes are the following:

M3:M12

N3:N12

We immediately hit the Enter or Return key, and a correlation coefficient of -0.91139 appears. In our hypothetical data set, then, risk taking and perfectionism are strongly and negatively correlated.

The statistics we've just calculated include more decimal places than we need and communicate a precision that really isn't warranted from such a small sample size. We can limit the number of decimal places to 2 by going to the "Format" menu, then to "Cells," then to "Number," and then, under "Category," to "Number" again. Our final calculations are shown in Figure A.5, along with the data as we previously reorganized them by risk-taking scores.

	Risk Taking Items				Perf'ism Items								
Person #	#11	#16	#39	#51	#19	#27	#38	#60	Rev11	Rev39	Rev27	RtScore	PerfScore
5	1	5	3	5	2	5	1	2	5	3	1	18	6
8	1	3	1	5	4	5	2	3	5	5	1	18	10
1	2	5	3	4	3	5	3	2	4	3	1	16	9
4	2	4	2	4	4	4	3	3	4	4	2	16	12
10	3	3	2	5	3	5	3	2	3	4	1	15	9
3	5	2	3	3	5	3	4	4	1	3	3	9	16
9	4	1	3	2	5	3	5	4	2	3	3	8	17
6	3	2	5	1	5	2	5	3	3	1	4	7	17
7	5	1	3	2	4	3	4	3	1	3	3	7	14
2	4	1	5	2	4	2	5	4	2	1	4	6	17
Mean												12	12.7
SD												4.99	4.06
Corr													-0.91

FIGURE A.5

The data set as reorganized, with descriptive statistics calculated

You can find other simple statistical tests in Excel, including *t*-tests and chi-square (χ^2) tests. For more sophisticated analyses, however, you will need statistical software such as SPSS, described in Appendix B.

Appendix B

Using SPSS

A complete explanation of how to use SPSS—short for Statistical Package for the Social Sciences—is well beyond the scope of a short appendix. However, a brief explanation of some of the basics can get you started. The version of SPSS we describe is PASW Statistics Student Version 18.0 for Macintosh.¹

Creating a Data Set



Once you have loaded SPSS onto the hard drive of your computer (see the directions that come with the program for details), open the program. On your screen you will see a two-dimensional table that looks very much like a spreadsheet. Each row in the table designates a specific individual (human participant, animal subject, artifact, etc.) in your data set. Each column designates a specific variable in the data set. Once filled in, this table will provide the basis for your data analyses.

As an example, we use data from a pilot study that Dinah Jackson conducted in preparation for her dissertation study (1996; excerpts from her dissertation appear in Chapters 1 and 12). The data include the following information for 15 students in a college psychology class; (a) their gender; (b) their scores on three exams given during the semester; (c) the total of the three exam scores; (d) the quantity of class notes (i.e., number of pages) they took during the semester; and (e) the quality of their class notes. The last of these variables—quality of notes—is based on content analyses of students' notes; the numbers are proportions of notes that reflect an integration of two or more ideas rather than a single, isolated fact. In Jackson's study, better-integrated notes (reflected in higher numbers, such as .406 or .496) were theorized to facilitate better learning—and thus to be of better quality—than relatively non-integrated notes (reflected in lower numbers, such as .166 or .040). Jackson's pilot data are shown in Figure B.1

Notice that the seven columns in the table in Figure B.1 have short labels that tell us what each variable is. To insert such labels, we go down to the bottom of the screen, where there are two "buttons" called "Data View" and "Variable View." If we click on "Variable View," we get another table, which looks like Figure B.2. In this table, we have entered information about each of the variables in the data set. Here the variables are the rows (rather than the columns, as they are in the "Data" table), and the things we want to say about the variables are the columns. To keep our discussion simple, we describe only some of these columns:

- **Name:** Indicates the label that will appear for the variable in the "Data View" table. This label can include alphabet letters, numbers, and a few other meaningful symbols (e.g., "\$").

¹At the instructor's request, this book can be packaged with the Student Version of SPSS at a discount; the CD for the software provides versions for both Windows and Macintosh users. Please contact your local Pearson representative if you are an instructor who is interested in setting up such a package for your students.

FIGURE B.1

The "Data" table

	Gender	Exam1	Exam2	Exam3	TotalExam	NoteQuan	NoteQual	Var
1	1	35.00	36.50	33.50	105.00	31	.315	
2	2	39.50	41.00	41.50	122.00	28	.384	
3	1	45.00	45.50	39.00	129.50	37	.381	
4	1	34.50	33.00	29.00	96.50	31	.251	
5	2	31.00	43.00	37.00	111.00	42	.305	
6	2	38.00	30.25	32.00	100.25	27	.190	
7	2	40.50	43.00	42.00	125.50	43	.350	
8	1	44.00	45.00	43.00	132.00	26	.166	
9	2	43.00	38.00	40.00	121.00	52	.406	
10	1	38.00	32.00	30.00	100.00	33	.208	
11	1	43.00	47.00	44.00	134.00	43	.496	
12	2	43.50	34.75	41.00	119.25	24	.201	
13	2	45.00	43.00	42.00	130.00	50	.321	
14	1	39.50	44.00	43.00	126.50	23	.179	
15	2	40.00	36.00	33.00	109.00	14	.040	
16								

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Gender	Numeric	8	0	Gender	{1, Male}...	None	8	Right	Nominal	Input
2	Exam1	Numeric	8	2	Exam 1 Score	None	None	8	Right	Scale	Input
3	Exam2	Numeric	8	2	Exam 2 Score	None	None	8	Right	Scale	Input
4	Exam3	Numeric	8	2	Exam 3 Score	None	None	8	Right	Scale	Input
5	TotalExam	Numeric	8	2	Exam Score Total	None	None	8	Right	Scale	Input
6	NoteQuan	Numeric	8	0	Quantity of No...	None	None	8	Right	Scale	Input
7	NoteQual	Numeric	8	3	Quality of Notes	None	None	8	Right	Scale	Input
8											
9											

FIGURE B.2

The "Variables" table

- **Type:** Indicates the type of data the variable represents, perhaps a number (numeric data), a letter string, a dollar amount, a date, or something else altogether.
- **Decimals:** Indicates an upper limit on the number of digits that will appear to the right of a decimal point.
- **Label:** Indicates the labels the variables will have when we create a table or graph—perhaps one to be included in a dissertation or research report.
- **Values:** Indicates labels that might be attached to particular values of a variable. For example, one of our variables is gender, a nominal scale. If we click on this "values" cell in the "Gender" row, a little button appears at the right side of the cell. We click on the button, and a box appears that allows us to tell the computer that a value of 1 means "male" and a value of 2 means "female." In Figure B.3, we show this box midway through the process: We've already told the software that a value of 1 means "Male," and we're in the process of telling it that 2 means "Female"; at this point, we click on "Add" and then on "OK" to say that we have labeled all possible values of the Gender variable.

FIGURE B.3

The "Value Labels" box for the "Variable View" table



- **Measure:** Indicates whether the variable reflects a nominal scale or an ordinal scale; the category "ordinal scale" also encompasses interval and ratio scales (Chapters 4 and 11 describe the four kinds of scales). As you can see in Figure B.2, our sample data set consists of one variable (Gender) on a nominal scale and six variables that are on interval or ratio scales—hence also on an ordinal scale, which in the Variables table is simply called "scale."

Computing Basic Descriptive Statistics

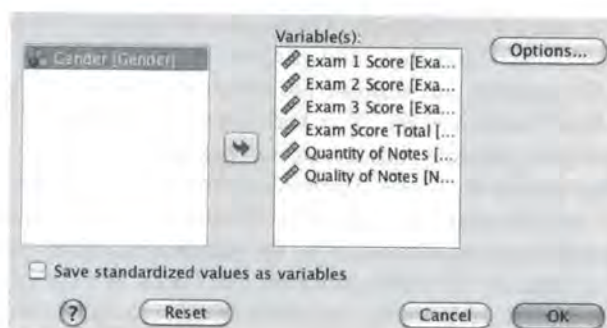
Now that we have our data set, let's conduct some simple analyses. First, let's compute basic descriptive statistics for six of the seven variables (computing a mean and standard deviation for the "gender" variable would, of course, be meaningless). We move the cursor to the word "Analyze" at the top of the screen and click on the mouse. A pull-down menu appears, and we move the mouse down until the term "Descriptive Statistics" is highlighted, at which point another menu appears to its right. We click on "Descriptives" in the right-hand box. A new box appears in front of our data set. This box contains two smaller boxes, with all seven of our variables listed in the left box. To calculate descriptive statistics for the last six variables, we want to move them into the right box. We do this by highlighting each one and then clicking the right-arrow button between the two boxes. After we've moved the six variables, we click on the "OK" button (see Figure B.4). At this point, a table appears that lists the number of observations (N), minimum and maximum values, mean, and standard deviation for each variable. The final row in the table, "Valid N (listwise)," simply means that SPSS found all 15 numbers for each variable to be appropriate ones; in other words, it didn't omit any scores in doing the calculations.

Now let's suppose that we want to see how overall exam performance (Exam Score Total), quantity of notes (Quantity of Notes), and quality of notes (Quality of Notes) are intercorrelated. To do this, we can calculate Pearson r correlation coefficients for each possible pairing of these three variables. Once again, we go up to "Analyze" at the top of the screen and click on the mouse. When the pull-down menu appears, we move the cursor down until the word "Correlate" is highlighted, then move the cursor to the right to highlight "Bivariate," and then click on the mouse. Once again, the two-box box appears, and we must move the three variables we want to analyze to the right box and then click on "OK." We now have a table that gives us the intercorrelations among these variables, which we can print out by going to the "File" pull-down menu and then to "Print" (see Figure B.5). The first number in each cell of the table tells us the Pearson r for a particular pair of variables (this number is 1 when a variable is correlated with itself), and the third number tells us the number of people for whom the r has been calculated. The middle number tells us the probability (p) that we would obtain an r that high if the two variables were *not* correlated in the overall population from which the sample has been drawn.

The table in Figure B.5 marks with two asterisks (**) all r s that are significant at an α level of .01. But we don't necessarily have to use that alpha level. Imagine, instead, that we decide to

FIGURE B.4

Identifying variables for which we want basic descriptive statistics to be calculated

**FIGURE B.5**

Correlations among exam score total, quantity of notes, and quality of notes

		Correlations		
		Exam Score Total	Quantity of Notes	Quality of Notes
Exam Score Total	Pearson Correlation	1	.323	.425
	Sig. (2-tailed)		.241	.114
	N	15	15	15
Quantity of Notes	Pearson Correlation	.323	1	.777**
	Sig. (2-tailed)	.241		.001
	N	15	15	15
Quality of Notes	Pearson Correlation	.425	.777**	1
	Sig. (2-tailed)	.114	.001	
	N	15	15	15

** . Correlation is significant at the 0.01 level (2-tailed).

use a significance level (α) of 0.05 for all of our analyses. Any p value in the table that is *smaller* than 0.05 indicates that the variables probably *are* correlated in the population from which our sample has been drawn. For example, the correlations between Exam Score Total and the Quantity and Quality of Notes are .323 and .425, respectively. Although these correlations are in the low-to-moderate range, the p values associated with them (.241 and .114) tell us that we might get correlations this high *simply by chance* when the two variables are actually unrelated in the overall population. (With a much larger sample size, such correlations would be statistically significant. Our small sample size may be leading us to make Type I errors here.) Now let's look at the correlation between Quantity of Notes and Quality of Notes. This correlation is .777, which has an associated probability of 0.001. This r is statistically significant: Students who take more notes also take better notes. We must be careful, however, that we don't conclude that there is a causal relationship here: Taking more notes does not necessarily cause a student to take better ones, nor does taking better ones cause a student to take more of them. Correlational data alone *never* allows us to draw clear-cut conclusions about cause-and-effect relationships.

Computing Inferential Statistics

In the preceding section we already ventured into inferential statistics a bit. When we looked at the probabilities that our correlation coefficients occurred by chance for a set of possibly unrelated variables, we were drawing inferences. But now let's do so intentionally. Let's see if there are any gender differences in the test performance of males and females. To find out, we need to perform

a t -test between the two groups. Once again, we go up to “Analyze,” and this time we highlight “Compare Means” and then “Independent Samples T Test.”² A box similar to that shown in Figure B.4 appears, but this one has three boxes within it. We move our dependent variable (Exam Score Total) into the “Test Variable(s)” box and our independent variable (Gender) into the “Grouping Variable” box. Next, we click on the “Define Groups” button and tell the computer that a value of “1” puts a person in Group 1 (the males) and a value of “2” puts a person in Group 2 (the females). We click on the “Continue” button and then click on “OK.” We get tables that provide descriptive statistics for the two groups, information about whether the variances of the two groups are equivalent, and results of t -tests. We can, of course, print out these tables (see Figure B.6). The program has calculated two t s, one based on the assumption of equal variances and another based on the assumption of unequal variances. Given the unequal variances for the two groups (the F value for Levene’s test has a probability of .013), we’ll look at the second t , which is .055. This value indicates that the two groups are probably not different in their overall exam performance (the p value is .957). (You can find explanations for the other numbers in this table in many statistics textbooks or through an Internet search.)

We have room for one final statistical analysis. Let’s say we want to know whether the students performed differently on the three exams they took during the semester. To compare three means for the same group of students, we would ideally want to conduct a repeated-measures analysis of variance. Unfortunately, the version of SPSS we are using here performs only between-subjects ANOVAs, so we will have to settle for three paired-samples t -tests.

To conduct our t -tests, we go back up to “Analyze,” move the mouse down to highlight “Compare Means,” and then move it to the right to highlight “Paired-Samples T Tests.” We release the mouse. Once again, we see a two-box box, but in this one the second box includes three columns labeled “Pair,” “Variable 1,” and “Variable 2.” When we click on Exam 1 in the left box and then click on the arrow, and then subsequently do the same thing for Exam 2, we get an Exam 1–Exam 2 pair in the right box. In a similar manner, we can form Exam 1–Exam 3 and Exam 2–Exam 3 pairs. We now have three pairs of variables in the right-hand box. We click on “OK” and print out the three tables that the analysis generates (Figure B.7). The first table gives us descriptive statistics; we’ve seen most of these before, but the column for standard error

Group Statistics

Gender	N	Mean	Std. Deviation	Std. Error Mean
Exam Score Total Male	7	117.6429	16.38524	6.19304
Female	8	117.2500	9.76418	3.45216

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Exam Score Total	Equal variances assumed	8.335	.013	.057	13	.955	.39286	6.85139	-14.40868	15.19439
	Equal variances not assumed			.055	9.520	.957	.39286	7.09022	-15.51373	16.29945

FIGURE B.6

Computing t to determine if males and females have different total exam scores

²As noted in Table 11.5 in Chapter 11, a t -test can take either of two basic forms. An *independent-samples t -test* enables a comparison of means for two separate, independent groups. For instance, an independent-samples t -test enables a comparison of males versus females, as in the example presented here. In contrast, a *dependent-samples t -test*—also known as a *paired samples t -test*—enables a comparison of means for a single group of individuals or, instead, for two related groups. For example, a researcher might obtain measures of two characteristics of a single group of students or, alternatively, might obtain measures of one particular characteristic both for a group of fathers and for their first-born sons.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Exam 1 Score	39.9667	15	4.15102	1.07179
	Exam 2 Score	39.4667	15	5.42388	1.40044
Pair 2	Exam 1 Score	39.9667	15	4.15102	1.07179
	Exam 3 Score	38.0000	15	5.15128	1.33006
Pair 3	Exam 2 Score	39.4667	15	5.42388	1.40044
	Exam 3 Score	38.0000	15	5.15128	1.33006

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Exam 1 Score & Exam 2 Score	15	.388	.153
Pair 2	Exam 1 Score & Exam 3 Score	15	.622	.013
Pair 3	Exam 2 Score & Exam 3 Score	15	.814	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Exam 1 Score – Exam 2 Score	.50000	5.40089	1.39450	-2.49091	3.49091	.359	14	.725
Pair 2	Exam 1 Score – Exam 3 Score	1.96667	4.14241	1.06956	-.32732	4.26065	1.839	14	.087
Pair 3	Exam 2 Score – Exam 3 Score	1.46667	3.23329	.83483	-.32387	3.25720	1.757	14	.101

FIGURE B.7

Computing *t*s to determine if students performed differently on the three exams

of the mean is new. We also see Pearson *r*s for the three pairs. We are most interested in the *t* values for three pairs of exam scores, which are shown in the seventh column in the bottom table. None of these *t*s is statistically significant at our significance level of .05 (see the rightmost column), although the Exam 1–Exam 3 pair comes close, with a *p* value of .087.

We have merely scratched the surface of what SPSS can offer. We have ignored some of the values in the statistical tables we've presented. And we haven't even touched on SPSS's graphing capabilities. We urge you to explore SPSS for yourself to discover the many analyses it can perform and the many graphical displays it can create.

References

- Abraham, E. P., Chain, E., Fletcher, C. M., Gardner, A. D., Heatley, N. G., Jennings, M. A., et al. (1941). Further observations on penicillin. *Lancet*, 2, 177–188.
- Acheson, K. A., & Gall, M. D. (2003). *Clinical supervision and teacher development: Preservice and inservice applications* (5th ed.). New York: Wiley.
- Allen, E. M. (1960, November). Why are research grant applications disapproved? *Science*, 132, 1532–1534.
- Altheide, D. L., & Johnson, J. M. (1994). Criteria for assessing interpretive validity in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 485–499). Thousand Oaks, CA: Sage.
- American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.
- Anderson, C. A., Lindsay, J. J., & Bushman, B. J. (1999). Research in the psychological laboratory: Truth or triviality? *Current Directions in Psychological Science*, 8, 3–9.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Bartholomew, D. J. (2004). *Measuring intelligence: Facts and fallacies*. Cambridge, England: Cambridge University Press.
- Beck, C. T. (1990, January–February). The research critique: General criteria for evaluating a research report. *Journal of Gynecology and Neonatal Nursing*, 19, 18–22.
- Becker, H. S. (1970). Whose side are we on? In W. J. Filstead (Ed.), *Qualitative methodology* (pp. 15–26). Chicago: Markham.
- Bender, G. (2001). Resisting dominance? The study of a marginalized masculinity and its construction within high school walls. In J. N. Burstyn, G. Bender, R. Casella, H. W. Gordon, D. P. Guerra, K. V. Luschen, et al., *Preventing violence in schools: A challenge to American democracy* (pp. 61–77). Mahwah, NJ: Erlbaum.
- Beyer, B. K. (1985). Critical thinking: What is it? *Social Education*, 49, 270–276.
- Brainerd, C. J., & Reyna, V. F. (2005). *The science of false memory*. Oxford, England: Oxford University Press.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school* (expanded ed.). Washington, DC: National Academy Press.
- Breisach, E. (1994). *Historiography: Ancient, medieval, and modern* (2nd ed.). Chicago: University of Chicago Press.
- Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research*, 6(1), 97–113.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105.
- Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 171–246). Chicago: Rand McNally.
- Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridge line of optimal retention. *Psychological Science*, 19, 1095–1102.
- Chain, E., Florey, H. W., Gardner, A. D., Heatley, N. G., Jennings, M. A., Orr-Ewing, J., et al. (1940). Penicillin as a chemotherapeutic agent. *Lancet*, 2, 226.
- Charmaz, K. (2000). Grounded theory: Objective and constructivist methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 509–535). Thousand Oaks, CA: Sage.
- Charmaz, K. (2002). Qualitative interviewing and grounded theory analysis. In J. F. Gubrium & J. A. Holstein (Eds.), *Handbook of interview research: Context and method*. Thousand Oaks, CA: Sage.
- Charmaz, K. (2006). *Constructing grounded theory*. Thousand Oaks, CA: Sage.
- Chatterjee, B. B., & Srivastava, A. K. (1982). A systematic method for drawing sociograms. *Perspectives in Psychological Researches*, 5(1), 1–6.
- Chicago manual of style (16th ed.) (2010). Chicago: University of Chicago Press.
- Chipman, S. F. (2005). Research on the women and mathematics issue: A personal case history. In A. M. Gallagher & J. C. Kaufman (Eds.), *Gender differences in mathematics: An integrative psychological approach* (pp. 1–24). Cambridge, England: Cambridge University Press.
- Cizek, G. J. (2003). *Detecting and preventing classroom cheating: Promoting integrity in assessment*. Thousand Oaks, CA: Corwin Press.
- Cochran-Smith, M., & Lytle, S. (1993). *Inside out: Teacher research and knowledge*. New York: Teachers College Press.
- Coghill, R. D. (1944). Penicillin: Science's Cinderella. The background of penicillin production. *Chemical and Engineering News*, 22, 588–593.
- Coghill, R. D., & Koch, R. S. (1945). Penicillin: A wartime accomplishment. *Chemical and Engineering News*, 23, 2310.
- Cole, D. B., & Ormrod, J. E. (1995). Effectiveness of teaching pedagogical content knowledge through summer geography institutes. *Journal of Geography*, 94, 427–433.
- Corallo, G., Sackur, J., Dehaene, S., & Sigman, M. (2008). Limits on introspection: Distorted subjective time during the dual-task bottleneck. *Psychological Science*, 19, 1110–1117.
- Corbin, J., & Strauss, A. C. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Thousand Oaks, CA: Sage.
- Council of Science Editors. (2006). *Scientific style and format: The CSE manual for authors, editors, and publishers* (7th ed.). Reston, VA: Author.
- Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Upper Saddle River, NJ: Pearson/Allyn and Bacon.
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.
- Cuca, J. M., & McLoughlin, W. J. (1987, May). Why clinical research grant applications fare poorly in review and how to recover. In *Preparing a research grant application to the*

- National Institutes of Health: *Selected articles* (a bulletin from the research branch of the Department of Health and Human Services). Washington, DC: Department of Health and Human Services. (Reprinted from *Clinical Investigation*, 5, pp. 55–58, 1987.)
- Davitz, J. R., & Davitz, L. L. (1996). *Evaluating research proposals: A guide for the behavioral sciences*. Upper Saddle River, NJ: Prentice Hall.
- Deaver, C. M., Miltenberger, R. G., & Stricker, J. M. (2001). Functional analysis and treatment of hair twirling in a young child. *Journal of Applied Behavior Analysis*, 34, 535–538.
- Do, S. L., & Schallert, D. L. (2004). Emotions and classroom talk: Toward a model of the role of affect in students' experiences of classroom discussions. *Journal of Educational Psychology*, 96, 619–634.
- Dowson, M., & McInerney, D. M. (2001). Psychological parameters of students' social and work avoidance goals: A qualitative investigation. *Journal of Educational Psychology*, 93, 35–42.
- Eisenhart, M., & DeHaan, R. L. (2005). Doctoral preparation of scientifically based education researchers. *Educational Researcher*, 34(4), 3–13.
- Eisner, E. W. (1998). *The enlightened eye: Qualitative inquiry and the enhancement of educational practice*. Upper Saddle River, NJ: Merrill/Prentice Hall.
- Elliott, D., & Stern, J. E. (Eds.). (1997). *Research ethics: A reader*. Hanover, NH: University Press of New England.
- Espelage, D. L., & Swearer, S. M. (Eds.). (2004). *Bullying in American schools: A social-ecological perspective on prevention and intervention*. Mahwah, NJ: Erlbaum.
- Ferguson, D. L., & Rosales-Ruiz, J. (2001). Loading the problem loader: The effects of target training and shaping on trailer-loading behavior of horses. *Journal of Applied Behavior Analysis*, 34, 409–424.
- Fleming, A. (1929). On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. *British Journal of Experimental Pathology*, 10, 226–236.
- Flesch, R. (1974). *The art of readable writing*. New York: Harper & Row.
- Fossey, D. (1983). *Gorillas in the mist*. Boston: Houghton Mifflin.
- Freeman, L. C. (2004). *The development of social network analysis: A study in the sociology of science*. North Charleston, SC: BookSurge, LLC.
- Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction* (8th ed.). Upper Saddle River, NJ: Merrill/Pearson Education.
- Gay, L. R., Mills, G. E., & Airasian, P. (2009). *Educational research: Competencies for analysis and application* (9th ed.). Upper Saddle River, NJ: Merrill/Pearson Education.
- Glaser, B. G. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory*. New York: Sociology Press.
- Glaser, B. G. (1992). *Basics of grounded theory analysis*. Mill Valley, CA: Sociology Press.
- Glaser, B. G., & Strauss, A. (1967). *The discovery of grounded theory*. Chicago: Aldine.
- Good, R. (1993). More guidelines for reviewing research. *Journal of Research in Science Teaching*, 30(1), 1–2.
- Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior*. Cambridge, MA: Harvard University Press.
- Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust Web-based studies? A comparative analysis of six preconceptions about Internet questionnaires. *American Psychologist*, 59, 93–104.
- Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, 11, 255–274.
- Guba, E., & Lincoln, Y. S. (1988). Do inquiry paradigms imply inquiry methodologies? In D. M. Fetterman (Ed.), *Qualitative approaches to evaluation in education* (pp. 89–115). New York: Praeger.
- Halpern, D. E. (1998). Teaching critical thinking for transfer across domains. *American Psychologist*, 53, 449–455.
- Halpern, D. E. (2008). Is intelligence critical thinking? Why we need a new definition of intelligence. In P. C. Kyllonen, R. D. Roberts, & L. Stankov (Eds.), *Extending intelligence: Enhancement and new constructs* (pp. 349–370). New York: Erlbaum/Taylor & Francis.
- Harwell, M. R., & Gatti, G. G. (2001). Rescaling ordinal data to interval data in educational research. *Review of Educational Research*, 71, 105–131.
- Haskins, L., & Jeffrey, K. (1990). *Understanding quantitative history*. New York: McGraw-Hill.
- Heck, A., Collins, J., & Peterson, L. (2001). Decreasing children's risk taking on the playground. *Journal of Applied Behavior Analysis*, 34, 349–352.
- Heine, S. J. (2007). Culture and motivation: What motivates people to act in the ways that they do? In S. Kitayama & D. Cohen (Eds.), *Handbook of cultural psychology* (pp. 714–733). New York: Guilford.
- Hesse-Biber, S. N. (2010). *Mixed methods research: Merging theory with practice*. New York: Guilford.
- Historical statistics of the United States: *Colonial times to 1970* (Part 1, bicentennial edition). (1975). Washington, DC: U.S. Department of Commerce.
- Holmbeck, G. N. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology*, 65, 599–610.
- Howe, K., & Eisenhardt, M. (1990). Standards for qualitative (and quantitative) research: A prolegomenon. *Educational Researcher*, 19(4), 2–9.
- Jaccard, J., & Jacoby, J. (2010). *Theory construction and model-building skills*. New York: Guilford.
- Jackson, D. L. (1996). *Effects of training in self-generation on the quality of students' questions, class notes and examination scores*. Unpublished doctoral dissertation, University of Northern Colorado, Greeley.
- Janos, P. M., & Robinson, N. M. (1985). Psychosocial development in intellectually gifted children. In F. D. Horowitz & M. O'Brien (Eds.), *The gifted and talented: Developmental perspectives*. Washington, DC: American Psychological Association.
- Johnson, B. (2001). Toward a new classification of nonexperimental quantitative research. *Educational Researcher*, 30(2), 3–13.
- Kahn, P. G. K., & Pompea, S. M. (1978, October 19). Nautiloid growth rhythms and dynamical evolution of the earth-moon system. *Nature*, 275, 606–611.
- Karabenick, S. A., Woolley, M. E., Friedel, J. M., Ammon, B. V., Blazewski, J. B., Bonney, C. R., et al. (2007). Cognitive processing of self-report items in educational research: Do they think what we mean? *Educational Psychologist*, 42, 139–151.
- Kearns, K. C. (1994). *Dublin tenement life*. Dublin, Ireland: Gill & MacMillan.
- Keeter, S., Dimock, M., Christian, L., & Kennedy, C. (2008, January 31). *The impact of "cell-onlys" on public opinion polls: Ways of coping with a growing population segment*. Washington, DC: Pew Research Center. Retrieved February 4, 2008, from <http://pewresearch.org/pubs/714/>
- Kellogg, R. T. (1994). *The psychology of writing*. New York: Oxford University Press.

- Kim-Cohen, J., Moffitt, T. E., Caspi, A., & Taylor, A. (2004). Genetic and environmental processes in young children's resilience and vulnerability to socioeconomic deprivation. *Child Development*, 75, 651-668.
- Kime, N. (2008). Children's eating behaviours: The importance of the family setting. *Area*, 40, 315-322.
- Kinnick, V. (1989). *Learning fetal monitoring under three conditions of concept teaching*. Unpublished doctoral dissertation, University of Colorado, Boulder.
- Kontos, S. (1999). Preschool teachers' talk, roles, and activity settings during free play. *Early Childhood Research Quarterly*, 14(3), 363-382.
- Kozinets, R. V. (2010). *Netnography: Doing ethnographic research online*. London: Sage.
- Krathwohl, D. R. (1993). *Methods of educational and social science research: An integrated approach*. White Plains, NY: Longman.
- Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., & Couper, M. (2004). Psychological research online: Report of Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet. *American Psychologist*, 59, 105-117.
- Kuhn, D. (1995). Microgenetic study of change: What has it told us? *Psychological Science*, 6, 133-139.
- Kvale, S. (1996). *InterViews: An introduction to qualitative research interviewing*. Thousand Oaks, CA: Sage.
- Lara, L. G. (2009). *A mixed method study of factors associated with the academic achievement of Latino college students from predominantly Mexican American backgrounds: A strengths-based approach*. Doctoral dissertation, University of Northern Colorado, Greeley. (Available through the online database ProQuest Dissertations & Theses: Full Text; publication number 3397099)
- Lauer, J. M., & Asher, J. W. (1988). *Composition research: Empirical designs*. New York: Oxford University Press.
- Leavenworth, P. S. (1998). "The best title that Indians can claim . . .": Native agency and consent in the transfer of Penacook-Pawtucket land in the seventeenth century. Unpublished master's thesis, University of New Hampshire, Durham.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Thousand Oaks, CA: Sage.
- Lind, D. A. (2006). *Statistical techniques in business and economics* (13th ed.). New York: McGraw-Hill.
- Lippa, R. A. (2002). *Gender, nature, and nurture*. Mahwah, NJ: Erlbaum.
- Lipsey, M. W. (1990). *Design sensitivity: Statistical power for experimental research*. Newbury Park, CA: Sage.
- Locke, E. A. (2009). It's time we brought introspection out of the closet. *Perspectives on Psychological Science*, 4, 24-25.
- Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, 13, 585-589.
- Lowes, J. L. (1927). *The road to Xanadu*. Boston: Houghton Mifflin.
- Lowes, J. L. (1955). *The road to Xanadu: A study in the ways of the imagination* (Rev. ed.). Boston: Houghton Mifflin.
- Lundeberg, M., & Mohan, L. (2009). Context matters: Gender and cross-cultural differences in confidence. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 221-239). New York: Routledge.
- Malthus, T. R. (1963). *An essay on the principle of population; or, a view of its past and present effects on human happiness, with an inquiry into our prospects respecting the future removal or mitigation of the evils which it occasions*. Homewood, IL: Irwin. (Original work published 1826)
- Marius, R. (1989). *A short guide to writing about history*. New York: HarperCollins.
- Marsh, H. W., Gerlach, E., Trautwein, U., Lüdtke, O., & Bretschneider, W.-D. (2007). Longitudinal study of preadolescent sport self-concept and performance reciprocal effects and causal ordering. *Child Development*, 78, 1640-1656.
- Maurois, A. (1959). *The life of Alexander Fleming: Discoverer of penicillin*. New York: E. P. Dutton.
- McCallin, R. C. (1988). *Knowledge application orientation, cognitive structure, and achievement*. Unpublished doctoral dissertation, University of Northern Colorado, Greeley.
- McCloskey, M. (1983). Naive theories of motion. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 299-324). Hillsdale, NJ: Erlbaum.
- McCrea, S. M., Liberman, N., Trope, Y., & Sherman, S. J. (2008). Construal level and procrastination. *Psychological Science*, 19, 1308-1314.
- McGibbon, E., Peter, E., & Gallop, R. (2010). An institutional ethnography of nurses' stress. *Qualitative Health Research*, 20, 1353-1378.
- McGraw, K. O., Tew, M. D., & Williams, J. E. (2000). The integrity of Web-delivered experiments: Can you trust the data? *Psychological Science*, 11, 502-506.
- McGrew, K. S., Flanagan, D. P., Zeith, T. Z., & Vanderwood, M. (1997). Beyond g: The impact of Gf-Gc specific cognitive abilities research on the future use and interpretation of intelligence tests in the schools. *School Psychology Review*, 26, 189-210.
- McGue, M. (2000). Authorship and intellectual property. In B. D. Sales & S. Folkman (Eds.), *Ethics in research with human participants* (pp. 75-95). Washington, DC: American Psychological Association.
- McKenzie, M. G. (2003). *Vocational science and the politics of independence: The Boston Marine Society, 1754-1812*. Unpublished doctoral dissertation, University of New Hampshire, Durham.
- Medawar, P. B. (1979). *Advice to a young scientist*. New York: Harper & Row.
- Mehan, H. (1979). *Social organization in the classroom*. Cambridge, MA: Harvard University Press.
- Mehan, H., & Wood, H. (1975). *The reality of ethnomethodology*. New York: Wiley.
- Mertler, C. A. (2009). *Action research: Teachers as researchers in the classroom* (2nd ed.). Thousand Oaks, CA: Sage.
- Middleton, M., Ormrod, J. E., & Abrams, E. (2007, April). Motivation, cognition, and social support: Achievement goals and preservice teacher apprenticeship. In M. Middleton & M. A. Duggan (Chairs), *Motivation of teachers as learners of the teaching craft*. Symposium presented at the American Educational Research Association, Chicago.
- Miksa, F. L. (1987). *Research patterns and research libraries*. Dublin, OH: OCLC.
- Milch-Reich, S., Campbell, S. B., Pelham, W. E., Jr., Connelly, L. M., & Geva, D. (1999). Developmental and individual differences in children's on-line representations of dynamic social events. *Child Development*, 70, 413-431.
- Miller, S. M., Nelson, M. W., & Moore, M. T. (1998). Caught in the paradigm gap: Qualitative researchers' lived experience and the politics of epistemology. *American Educational Research Journal*, 35, 377-416.
- Mills, G. E. (2011). *Action research: A guide for the teacher researcher* (4th ed.). Upper Saddle River, NJ: Allyn & Bacon/Pearson.
- Mitchell, K. J. (1998). *Childhood sexual abuse and family functioning linked with eating and substance misuse: Mediated structural models*. Unpublished doctoral dissertation, University of Rhode Island, Kingston.
- Modern Language Association. (2008). *MLA style manual and guide to scholarly publishing* (3rd ed.). New York: Author.

- Moon, J. (2008). *Critical thinking: An exploration of theory and practice*. London: Routledge.
- Munter, M., & Paradi, D. (2009). *Guide to PowerPoint*. Upper Saddle River, NJ: Pearson/Prentice Hall.
- Murphy, K. R., Myers, B., & Wolach, A. (2009). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests* (3rd ed.). New York: Routledge.
- Myrdal, G. (1973). The beam in our eyes. In D. Warwick & S. Osherson (Eds.), *Comparative research methods* (pp. 89–99). Englewood Cliffs, NJ: Prentice Hall.
- Neuman, W. L. (2011). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Boston: Allyn & Bacon.
- Nicholls, M. E. R., Orr, C. A., Okubo, M., & Loftus, A. (2006). Satisfaction guaranteed: The effect of spatial biases on responses to Likert scales. *Psychological Science*, 17, 1027–1028.
- Nichols, J. D. (1998). Multiple perspectives of collaborative research. *International Journal of Educational Reform*, 7, 150–157.
- Nicol, A. A. M., & Pexman, P. M. (2010). *Presenting your findings: A practical guide for creating figures, posters, and presentations* (6th ed.). Washington, DC: American Psychological Association.
- Nussbaum, E. M. (2008). Collaborative discourse, argumentation, and learning: Preface and literature review. *Contemporary Educational Psychology*, 33, 345–359.
- Ormrod, J. E. (2011). *Our minds, our memories*. Boston: Allyn & Bacon/Pearson.
- Ormrod, J. E. (2012). *Human learning* (6th ed.). Upper Saddle River, NJ: Pearson.
- Ormrod, J. E., Ormrod, R. K., Wagner, E. D., & McCallin, R. C. (1988). Reconceptualizing map learning. *American Journal of Psychology*, 101, 425–433.
- Ormrod, R. K. (1974). *Adaptation in cultural ecosystems: Early 19th century Jamaica*. Unpublished doctoral dissertation, The Pennsylvania State University, University Park.
- Ormrod, R. K., & Trahan, R. G. (1982). Can signs help visitors control their own behavior? *Trends*, 19(4), 25–27.
- Peshkin, A. (1993). The goodness of qualitative research. *Educational Researcher*, 22(2), 23–29.
- Peterson, C. (2009). Minimally sufficient research. *Perspectives on Psychological Science*, 4, 7–9.
- Polkinghorne, D. E. (1989). Phenomenological research methods. In R. S. Valle & S. Halling (Eds.), *Existential-phenomenological perspectives in psychology* (pp. 41–60). New York: Plenum.
- Ramirez, I. L. (2001). *The relation of acculturation, criminal history, and social integration of Mexican American and non-Mexican students to assaults on intimate partners*. Unpublished doctoral dissertation, University of New Hampshire, Durham.
- Rogelberg, S. G., & Luong, A. (1998). Nonresponse to mailed surveys: A review and guide. *Current Directions in Psychological Science*, 7, 60–65.
- Sales, B. D., & Folkman, S. (Eds.). (2000). *Ethics in research with human participants*. Washington, DC: American Psychological Association.
- Schram, T. H. (2006). *Conceptualizing and proposing qualitative research* (2nd ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.
- Schuman, H. (1967). Economic development and individual change: A social-psychological study of the Comilla Experiment in Pakistan. *Occasional Papers in International Affairs*, No. 15. Cambridge, MA: Harvard University, Center for International Affairs.
- Schunk, D. H., & Pajares, F. (2005). Competence perceptions and academic functioning. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 85–104). New York: Guilford Press.
- Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, 54, 93–105.
- Scott, K. M., McGee, M. A., Wells, J. E., & Oakley Browne, M. A. (2008). Obesity and mental disorders in the adult general population. *Journal of Psychosomatic Research*, 64, 97–105.
- Scott-Jones, D. (2000). Recruitment of research participants. In B. D. Sales & S. Folkman (Eds.), *Ethics in research with human participants* (pp. 27–34). Washington, DC: American Psychological Association.
- Senders, V. L. (1958). *Measurement and statistics: A basic text emphasizing behavioral science*. New York: Oxford University Press.
- Shaklee, J. M. (1998). *Elementary children's epistemological beliefs and understandings of science in the context of computer-mediated video conferencing with scientists*. Unpublished doctoral dissertation, University of Northern Colorado, Greeley.
- Shanahan, T. (2004). Overcoming the dominance of communication: Writing to think and to learn. In T. L. Jetton & J. A. Dole (Eds.), *Adolescent literacy research and practice* (pp. 59–74). New York: Guilford.
- Shank, G. D. (2006). *Qualitative research: A personal skills approach* (2nd ed.). Upper Saddle River, NJ: Pearson/Merrill/Prentice Hall.
- Sheehan, K. (2001, January). E-mail survey response rates: A review. *Journal of Computer-Mediated Communication*, 6(2). Retrieved from <http://jcmc.indiana.edu/vol6/issue2/sheehan.html>
- Siebert, J. E. (2000). Planning research: Basic ethical decision-making. In B. D. Sales & S. Folkman (Eds.), *Ethics in research with human participants* (pp. 13–26). Washington, DC: American Psychological Association.
- Silverman, D. (1993). *Interpreting qualitative data: Methods for analysing talk, text and interaction*. London: Sage.
- Silverman, D., Masland, R., Saunders, M. G., & Schwab, R. S. (1970, June). Irreversible coma associated with electrocerebral silence. *Neurology*, 20, 525–533.
- Skagert, K., Dellve, L., Eklöf, M., Pousette, A., & Ahlberg, G. (2008). Leaders' strategies for dealing with their own and their subordinates' stress in public human service organisations. *Applied Ergonomics*, 39, 803–811.
- Smith, R. M. (1999). *Academic engagement of high school students with significant disabilities: A competence-oriented interpretation*. Unpublished doctoral dissertation, Syracuse University, Syracuse, New York.
- Sobel, D. (2000). *Galileo's daughter*. New York: Penguin Books.
- Solomon, R. I. (1949). An extension of control group design. *Psychological Bulletin*, 46, 137–150.
- Sowell, E. R., Thompson, P. M., Holmes, C. J., Jernigan, T. L., & Toga, A. W. (1999). *In vivo* evidence for post-adolescent brain maturation in frontal and striatal regions. *Nature Neuroscience*, 2, 859–861.
- Stake, R. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.
- Steiner, E. (1988). *Methodology of theory building*. Sydney, Australia: Educology Research Associates.
- Stevens, S. S. (1946, June 7). On the theory of scales of measurement. *Science*, 103, 677–680.
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Thousand Oaks, CA: Sage.
- Strunk, W., Jr., & White, E. B. (2009). *The elements of style* (50th anniversary ed.). New York: Pearson/Longman.
- Survey Research Center, Institute for Social Research at the University of Michigan. (1976). *Interviewer's manual* (Rev. ed.). Ann Arbor: Author.
- Tesch, R. (1994). The contribution of a qualitative method: Phenomenological

- research. In M. Langenbach, C. Vaughn, & L. Aagaard (Eds.), *An introduction to educational research* (pp. 143–157). Boston: Allyn & Bacon.
- Thompson, B. (2008). *Foundations of behavioral statistics*. New York: Guilford.
- Thompson, K. R. (2006). Axiomatic theories of intentional systems: Methodology of theory construction. *Scientific Inquiry Journal*, 7(1), 13–24.
- Thraillkill, N. J. (1996). *Imagery-evoking and attention-attracting material as facilitators of learning from a lecture*. Unpublished doctoral dissertation, University of Northern Colorado, Greeley.
- Toynbee, A. (1939–1961). *A study of history* (12 vols.). London: Oxford University Press, Royal Institute of International Affairs.
- Trahan, R. G. (1978). *Social science research: Rocky Mountain National Park* (Contract agreement PX 1520-8-A529). Greeley, CO: Author.
- Triola, M. F. (2008). *Essentials of statistics* (3rd ed.). Reading, MA: Addison-Wesley.
- Use of elegant statistics. (1987). *Research in Nursing and Health*, 10, iii.
- Uziel, L. (2010). Rethinking social desirability scales: From impression management to interpersonally oriented self-control. *Perspectives on Psychological Science*, 5, 243–263.
- Walton, D. N. (2003). *Ethical argumentation*. New York: Lexington Books.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge, England: Cambridge University Press.
- Wennick, A., Lundqvist, A., & Hallström, I. (2009). Everyday experience of families three years after diagnosis of Type 1 diabetes in children: A research paper. *Journal of Pediatric Nursing*, 24, 222–230.
- Witt, E., & Best, J. (April 21, 2008). *How different are people who don't respond to pollsters?* Washington, DC: Pew Research Center. Retrieved April 28, 2008, from <http://pewresearch.org/pubs/807/>
- Wolcott, H. F. (1994). *Transforming qualitative data: Description, analysis, and interpretation*. Thousand Oaks, CA: Sage.
- Wong, P. T. P. (n. d.). *How to write a research proposal*. Retrieved from http://www.meaning.ca/archives/archive/art_how_to_write_P_Wong.htm.
- Zambo, D. (2003). *Uncovering the conceptual representations of students with learning disabilities*. Unpublished doctoral dissertation, Arizona State University, Tempe.

Index

- Abduction, 20
- Abraham, E. P., 22
- Abrams, E., 198, 336
- Abscissa, 186
- Absolute value, 288
- Absolute zero point, 87, 279
- Abstracts, 55, 317
- Academic disciplines
 - codes of ethics, 109
 - conference presentations and, 329
 - identifying tools in, 24–25
 - interdisciplinary nature of, 55, 74, 75
 - mixed-method designs and, 258–259
 - prose style and, 324
 - qualitative research in, 139
 - questionnaires and, 199–200
 - reference lists and, 318
 - research in, 24, 76
 - styles in research reports, 310, 311, 323
- Academic integrity, 33, 69, 80, 217, 219, 316, 321
- Accessible Archives* (online database), 174
- Accidental sampling, 214
- Acheson, K. A., 194
- Acknowledgments, 318
- Action research, 27, 100
- Active voice, 323
- Advance organizers, 13, 68, 313
- Ahlborg, G., 146
- Airasian, P., 215, 259n
- Allen, E. M., 133
- Alpha (α), 297, 299n
- Alpha error (Type I error), 298–299, 299n, 314
- Alternating treatments designs, 239, 241, 247
- Altheide, D. L., 162
- American Psychological Association (APA)
 - Publications Manual*
 - formatting headings and, 122–123
 - reference list sample, 319
 - style guidelines, 310, 311, 319–320, 323
- Analysis. *See also* Data analysis; Dissertation analysis
 - ethnography and, 144, 145
- Analysis of covariance (ANCOVA), 231, 292, 301
- Analysis of variance (ANOVA), 301, 346
- Ancestry Library* (online database), 174
- Anderson, C. A., 103n
- Animal research review boards, 109, 263
- APA. *See* American Psychological Association (APA) *Publications Manual*
- Appendices, 127–128, 314, 321
- Applied research, basic research *versus*, 27, 103–104
- Argument, rules for, 178
- Argument analysis, 17
- Arithmetic average, 283–284
- Arithmetic mean, 284, 286
- The Art of Readable Writing* (Flesch), 332
- Asher, J. W., 139
- Association for Psychological Science, 205n, 250n
- Assumptions
 - definition of, 5
 - historical research and, 173, 177
 - proposal writing and, 46
 - qualitative research and, 162
 - questionnaires and, 197
 - in research report, 312
 - stating, 44
- Attachments, 21, 21n
- Attrition rates, for online studies, 250, 250n
- Authors, in reference lists, 319
- Authorship, sharing, 331–332
- Average, 276, 277
- Average deviation (AD), 288
- Axial coding, in grounded theory studies, 147
- Bakari, Rosenna, 134–138
- Bandura, A., 41
- Bartholomew, D. J., 82n
- Baseline data, 238
- Basic research, applied research *versus*, 27, 103–104
- Beck, C. T., 327
- Becker, H. S., 152
- Bell curve, 280, 281, 284
- Bender, G., 143
- Benton, Arthur, 60–61, 131
- Best, J., 218
- Beta error (Type II error), 298–299, 315
- Beyer, B. K., 17
- Bias
 - acknowledging presence of, 218–219
 - collaboration and, 21
 - confirmation bias, 39
 - definition of, 217
 - historical research and, 173, 174
 - literature review and, 51–52
 - measurement and, 92
 - objectivity and, 9
 - qualitative research and, 162
 - in research report, 312
 - in samples, 206, 207, 294, 316
 - sampling bias, 216–219
- BiblioExpress, 66, 319
- Bibliographic software, 318
- Bibliography Writer, 66, 319
- Biblioscope, 66, 318
- Biserial correlation, 292
- Borg, W. R., 162, 195
- Bracketing, 146
- Brainerd, C. J., 153
- Brainstorming
 - subproblems and, 38–39, 60
 - word processors and, 15
- Brainstorming software, 38–39, 60
- Bransford, J. D., 20
- Breaching experiments, 144
- Breisach, E., 177
- Brettschneider, W.-D., 291
- Brown, A. L., 20
- Browsing library shelves, 54–55
- Bryman, A., 259
- Bulleted lists, 131, 324, 330
- Bushman, B. J., 103n
- Byrd, William, 32, 37–38, 43
- Call numbers, 53–54, 62
- Campbell, D. T., 90, 104, 228–229, 232, 232n
- Campbell, S. B., 252
- Caracelli, V. J., 259
- Case studies, 100, 141–142, 150, 233
- Caspi, A., 41
- Causal-comparative designs. *See* Ex post facto designs
- Causation, correlation distinguished from, 23, 187, 293, 315
- Cause-and-effect relationships
 - assumptions and, 5
 - determining, 248–249
 - experimental design and, 226, 228–229, 232, 293
 - internal validity and, 101, 102
 - pre-experimental designs and, 233
 - quantitative research and, 140
 - in research report, 315
 - variables and, 40–42, 228
- Central tendency, measures of. *See* Measures of central tendency
- Cepeda, N. J., 250, 251
- Chain, Ernst B., 22, 23
- Charmaz, K., 147, 148
- Chatterjee, B. B., 84

Checklists

- for confounding variables, 248–249
- for critiquing research reports, 327–328
- in descriptive research, 192–193, 194
- for evaluating qualitative research studies, 163–164
- for evaluating research, 64–65
- for evaluating research problems, 35
- for evaluating research project, 46–47
- for expert researcher interviews, 25
- for judging feasibility of research project, 110–112
- for mixed-methods designs feasibility, 263–264
- for planning ethical research study, 109–110
- for planning qualitative research, 160–162
- for population analysis, 219–220
- for research proposal effectiveness, 132–133
- Chicago Manual of Style*, guidelines of, 310, 311
- Children, ethical issues and, 76, 105–106, 105n
- Chipman, S. F., 199
- Chi-square (χ^2) goodness-of-fit test, 85, 301
- Christian, L., 217
- Chunk selection, in multistage sampling, 215
- Circular definitions, 44
- Citations
 - format for, 310
 - literature review and, 59–60, 69
 - reference list and, 131, 318
- Cizek, G. J., 68
- Clarity
 - in interviews, 195
 - in proposal writing, 129
 - in research reports, 324
 - in stating research problem, 31–32, 34
 - in writing literature review, 67
- Classification, 11, 12
- Cluster sampling, 211, 213, 215
- Cochran-Smith, M., 27
- Cocking, R. R., 20
- Codes of ethics, 109
- Coding
 - axial coding, 147
 - open coding, 147
 - recoding as spreadsheet function, 274–275
 - schemes for, 14
 - selective coding, 147
- Coefficient of determination (R^2), 291, 292
- Coghill, R. D., 22
- Coherence, as qualitative research evaluation criterion, 162
- Cohort-sequential studies, 189
- Cole, D. B., 210, 218
- Collaboration with others, 21, 36, 47, 119
- Collective case studies, 141
- Collins, J., 239
- Combined experimental and ex post facto designs, 244–245, 247

- Comparative-historical research, 170
- Complementarity, in mixed-methods designs, 259
- Completeness
 - mixed-methods designs and, 259
 - qualitative research and, 162
- Computers. *See also* Internet; Software; Technology; Word processors
 - data collection with, 9, 61–62
 - organizing collected information on, 65–66
 - for questionnaire administration, 201
 - as research tools, 8–9
- Conceptual density, 147
- Conceptual historical research, 177–178
- Conclusions
 - of research report, 315–316
 - summaries in, 14, 316
- Concrete examples, 13–14
- Conferences, professional, 30–31
- Confidence intervals, 296
- Confirmation bias, 39
- Confounding variables, 228, 229–232, 316
- Connelly, L. M., 252
- Consensus, as qualitative research evaluation criterion, 162
- Consent. *See* Informed consent
- Consistency
 - internal consistency reliability, 91, 199, 199n
 - in questionnaires, 198–199
 - in research reports, 324
- Constant comparative method, 147
- Construct validity, 90
- Content analysis, 100, 148–149, 150
- Content validity, 89–90
- Contingency coefficient, 292
- Continuous variables, discrete variables *versus*, 279
- Control
 - confounding variables and, 228, 229–232
 - experimental design and, 227–232, 246
 - as research criteria, 76–77
- Control group, time-series design, 238, 247
- Control groups
 - confounding variables and, 229–230
 - definition of, 229
 - experimental designs and, 232, 246–247
 - pre-experimental designs and, 233–234
 - quasi-experimental designs and, 237–241
 - true experimental designs and, 234–237
- Controlled laboratory studies, 102
- Convenience sampling, 214
- Convergent designs, 260
- Copyrights
 - endnotes and footnotes and, 318
 - permissions and, 63, 69, 69n, 318
 - in preliminary pages, 317
- Corallo, G., 153
- Corbett, K. E., 70–73
- Corbin, J., 146n, 147, 148

Correlation

- causation distinguished from, 23, 187, 293, 315
- definition of, 185
- measures of association and, 291–293
- Correlational research
 - characteristics of, 100, 185–187
 - ex post facto designs contrasted with, 242
 - longitudinal studies and, 188
- Correlational statistics, examples of, 292
- Correlation coefficients
 - correlational research and, 186
 - measures of association and, 291
 - research problems and, 28
 - t*-test for, 301
 - validity and reliability affecting, 293
- Counterbalancing approach, in questionnaires, 199
- Credibility, 17, 101, 104, 176, 262
- Credit
 - in literature review, 69, 108
 - in research report, 316
- Creswell, J. W., 104, 139, 141, 143, 145, 146, 148, 151, 154, 158, 159, 162, 258, 259n, 262, 263, 264–265
- Criterion validity, 90
- Critical thinking, 17
- Cronbach's alpha coefficient, 199n
- Cross-sectional studies, 188–189
- CSE (Council of Science Editors) style, 310, 311
- Cuca, J. M., 133
- Data
 - admissibility of, 79–80
 - definition of, 77
 - dynamics within, 271
 - importing, 274
 - location of, 80
 - nature of, 278–282, 285, 291, 293
 - as plural form, 4n, 77
 - presentation in research report, 313–314
 - primary *versus* secondary data, 78
 - research and, 2, 4, 77–81
 - research methodology linked with, 94–98
 - statistics and, 10–11
 - transient nature of, 77–78
- Data analysis. *See also* Statistics
 - case studies and, 141–142
 - content analysis and, 149
 - dissertation analysis, 304–307
 - ethnography and, 144
 - grounded theory studies, 147–148
 - mixed-methods designs and, 264–265
 - phenomenological study and, 146
 - proposal writing and, 134
 - qualitative research and, 97, 158–160
 - quantitative research and, 97, 270–308
 - questionnaires and, 200–201
 - research methodology and, 5

- software and, 274–276
- word processors and, 15
- Data analysis spiral, 158, 159
- Databases
 - data interpretation and, 159–160
 - definition of, 65
 - importing data and, 274
 - literature review and, 318
 - online databases, 55–58, 174
 - questionnaires and, 201
 - of related literature, 65–66
- Data collection
 - computers and, 9, 61–62
 - data-gathering tools, 61–62
 - descriptive research and, 191, 205–206
 - historical research and, 125–126, 174, 175
 - planning for, 80–81
 - qualitative research and, 75, 97, 98–99, 140, 151–154
 - quantitative research and, 97, 98–99
 - research problems and, 5–6
 - research report and, 312
 - technology used for, 157–158, 205–206
- Data file, 66
- Data interpretation
 - computers and, 9
 - databases for, 159–160
 - descriptive research and, 220
 - descriptive statistics and, 290
 - historical research and, 176–177
 - measurement and, 82
 - mixed-methods designs and, 264–265
 - proposal writing and, 126–127
 - qualitative research and, 158–159
 - quantitative research and, 272–274
 - research problems and, 5–6, 28, 80, 303, 304
 - in research reports, 311, 314–315
 - statistical techniques and, 303–304
 - subproblems and, 37, 80
- Data sets
 - correlation in, 293
 - descriptive statistics and, 282
 - normalizing, 290
 - organization of, 270–276
- Date of publication, in reference lists, 319
- Davitz, J. R., 133
- Davitz, L. L., 133
- Deaver, C. M., 241
- Decile, 287n
- Decision making, 17
- Deductive logic, 17–18, 20, 23, 97
- Definitions
 - delimiting research and, 43–44
 - research proposals and, 45
- DeHaan, R. L., 55
- Dehaene, S., 153
- Delimitations, 43–44, 45, 312
- Dellve, L., 146
- Dependent-samples *t*-test, 301, 346n
- Dependent variables, 40–42, 226–227, 232
- Description
 - ethnography and, 144, 145
 - as purpose of qualitative research, 140
 - thick description, 104, 162, 314
- Descriptive research
 - checklists in, 192–193, 194
 - correlational research, 185–187
 - data collection and, 191, 205–206
 - data interpretation and, 220
 - developmental designs, 188–189
 - Internet and, 250
 - issues related to, 220–221
 - observation studies, 100, 184–185
 - planning and, 191
 - population analysis and, 219–220
 - samples and, 206–219
 - survey research, 189–191
- Descriptive statistics. *See also* Inferential statistics; Statistical techniques; Statistics
 - function of, 10, 277, 282
 - measures of association and, 291–293
 - measures of central tendency and, 283–286
 - measures of variability and, 286–290
 - in SPSS, 344–345
- Descriptive surveys, 189
- Developmental research, 100, 188–189
- Deviation, 287. *See also* Standard deviation
- Dewey decimal (DD) classification system, 53–54
- Dictionaries, 16, 66
- Differentia, 44, 45
- Digital cameras, 174
- Digital Object Identifier (DOI), 59, 320, 334
- Dimock, M., 217
- Direction, correlation coefficients for two variables, 291
- Discrete variables, continuous variables *versus*, 279
- Discriminant sampling, 152
- Dissertation Abstracts International*, 27, 33, 317
- Dissertation analysis
 - data analysis, 304–307
 - data collection, 164–168
 - experimental designs, 252–256
 - historical research, 179–182
 - literature review, 70, 71–72
 - mixed-methods designs, 265–269
 - questionnaires and, 221–224
 - research proposal, 135–138
- Distributions
 - measures of variability and, 287, 290
 - normal distributions, 280–282, 284, 296
 - polymodal distributions, 285
 - of sample means, 295, 296
- Do, S. L., 146
- Documentary theft, 108
- Document delivery service, online, 63
- Dogma, 23
- DOI (Digital Object Identifier), 59, 320, 334
- Double-blind experiments, 102
- Dowson, M., 191
- Dublin Tenement Life* (Kearns), 171
- Duke University, 108
- Dynamics and energy, as research problem category, 75
- Editing. *See also* Writing
 - revisions and, 128–132
 - in statement of research problem, 34–35
 - in word processors, 15, 16
- Editing marks, 34, 128, 129
- Effect size, 302, 304
- Einstein, Albert, 6, 20
- Eisenhardt, M., 162
- Eisenhart, M., 55
- Eisner, E. W., 139, 141, 151, 154, 157, 162
- Eklöf, M., 146
- Electronic spreadsheets, 194, 274–276. *See also* Microsoft Excel
- The Elements of Style* (Strunk & White), 332
- Elliott, D., 332
- E-mail, 21, 21n, 200
- Embedded designs, 260
- Emergent theory. *See* Grounded theory studies
- Emerging design, 151, 260
- Empirical research, 5n
- End matter, for research reports, 317, 318–321
- EndNote, 66, 318
- Endnotes, 318
- Epoché, 146
- Equivalent forms reliability, 91
- Espelage, D. L., 291
- An Essay on the Principle of Population* (Malthus), 284
- Ethical issues
 - children and, 76, 105–106, 105n
 - content analysis and, 148n
 - experimental research and, 230
 - honesty with professional colleagues, 108
 - informed consent and, 105–106, 128n, 250
 - internal review boards and, 106, 108–109
 - Internet-based research studies and, 108, 158, 250
 - mixed-methods designs and, 263
 - online surveys and, 206
 - permissions and, 76, 102, 105–106, 105n, 107, 108, 156, 194, 196, 203, 263
 - placebos and, 230
 - planning and, 104–110
 - protection from harm and, 105, 206, 250
 - qualitative research and, 151
 - right to privacy and, 107–108, 151, 158, 250, 316
 - shared authorship and, 332
 - unobtrusive measures and, 102, 106
 - voluntary and informed participation, 105–106, 128n

- Ethics, codes of, 109
- Ethnography
 - characteristics of, 100, 150, 314
 - prose style and, 324
 - qualitative research design and, 100, 142–145, 150
 - research methodology and, 143–145
- Evaluation
 - as purpose of qualitative research, 140
 - of qualitative research, 162
- Excel. *See* Microsoft Excel
- Exercise in self-enlightenment, 2
- Expectations, 23
- Experimental designs
 - cause-and-effect relationships and, 226, 228–229, 232, 293
 - combined experimental and ex post facto design, 244–245
 - confounding variables and, 228, 229–232
 - control and, 77, 227–232
 - dissertation analysis, 252–256
 - ex post facto designs, 232, 242–243
 - factorial designs, 232, 243–245
 - hypotheses and, 251
 - internal validity and, 102, 227–229
 - meta-analysis and, 302
 - pre-experimental designs, 232, 233–234, 246
 - quasi-experimental designs, 231, 232, 237–241, 246–247
 - summary of, 246–247
 - true experimental designs, 232, 234–327, 246
- Experimental groups, 229, 230
- Experimental research
 - characteristics of, 100
 - correlational research and, 187
 - hypotheses and, 39
 - Internet and, 250–251
- Experts
 - checklist for interviewing, 25
 - judgment by panel of, 91
 - seeking advice of, 30, 36
- Explanatory designs, 260, 263
- Exploratory designs, 260, 263
- Ex post facto designs
 - characteristics of, 232, 242
 - combined experimental and ex post facto design, 244–245
 - hypotheses and, 251
 - meta-analysis and, 302
 - simple ex post facto design, 242–243, 247
 - summary of, 247
- Ex post facto research, 100, 178
- External evidence, in historical research, 176
- External validity, 103–105, 207, 229, 235, 262
- Face-to-face interviews, 190–191, 191n
- Face validity, 89
- Factor analysis, 301
- Factorial designs, 232, 243–245
- Fact transcription, research *versus*, 2
- Fact transportation, research *versus*, 2
- Fahrenheit, Gabriel, 86
- Faust, K., 84
- Feedback
 - fine-tuning research problem and, 48
 - for journal articles, 331
 - literature review and, 70
 - proposal writing and, 132
 - research report writing and, 326
 - validity and, 104
- Ferguson, D. L., 251, 274, 275
- Fieldwork, 104, 143–145
- Figures, in research reports, 14, 313, 313n
- First person, 324
- Fisher's exact test, 301
- Fiske, D. W., 90
- Flanagan, D. P., 291
- Flash drives, 16
- Fleming, Alexander, 22, 23, 238
- Flesch, R., 332
- Florez, Henry W., 22, 23
- Focus groups, 154, 157
- Folkman, S., 105
- Follow-up letters, 205
- Footnotes, 318
- Foreign languages, 12
- Formatting features, in word processors, 15, 124
- Formulas
 - for mean, 284
 - as spreadsheet function, 275–276, 338–339
 - for standard deviation, 288
 - for standard error of the mean, 295–296
- Fossey, D., 95
- Fractional parts, 87
- Freeman, L. C., 84
- Front matter, of research reports, 317–318
- Gall, J. P., 162, 195
- Gall, M. D., 162, 194, 195
- Gallop, R., 143
- Gatekeepers, 143
- Gatti, G. G., 290
- Gay, L. R., 215, 259n
- Genealogical research, 174, 178
- Genera, 44, 45
- Generalizations, 12, 23, 103–104, 142
- Geometric mean, 284–285, 286
- Gerlach, E., 291
- Geva, D., 252
- Glaser, B. G., 146, 147, 148, 162
- Goals
 - articulation of research goal, 2–3
 - writing schedule and, 326
- Good, R., 163
- Goodall, J., 95
- Google Books, 57, 58, 63
- Google Scholar, 27, 57, 58
- Gosling, S. D., 206
- Government publications, 58, 59
- Graham, W. E., 259
- Grammar checkers, 14, 15, 16, 130, 323–324
- Graphing, as spreadsheet function, 276
- GraphPad Software, 303
- Greene, J. C., 259
- Grounded theory studies, 4, 100, 146–148, 150, 259, 262
- Growth, as function of geometric progression, 284, 285
- Guba, E. G., 104, 139
- Hallström, I., 148
- Halpern, D. F., 17
- Handouts, for presentations, 330–331
- Harm, protection from, 105, 206, 250
- Harwell, M. R., 290
- Haskins, L., 173
- Hawthorne effect, 102, 102n, 229
- Headings, formatting, 122–123
- Heck, A., 239
- Heine, S. J., 156, 199
- Helix, 7
- Heraclitus, 190
- Hesse-Biber, S. N., 147, 259, 260, 262, 263
- Historical research
 - characteristics of, 100
 - data collection and, 125–126, 174, 175
 - data interpretation and, 176–177
 - data sources in, 170–174
 - prose style and, 324
 - psychological or conceptual, 177–178
 - purpose of, 170
 - research reports and, 178
- Historiography, 95
- Holmbeck, G. N., 41
- Holmes, C. J., 57
- Homonyms, 130–131
- Honesty, 2, 33, 69, 104, 108, 304, 316, 332
- Housing unit selection, in multistage sampling, 215
- Howe, K., 162
- Human communication
 - communicating through writing, 13–16
 - forms of, 148
- Human mind. *See also* Open-mindedness
 - as research tool, 11, 16–21
- Hypotheses
 - correlational research and, 187
 - definition of, 39
 - experimental designs and, 251
 - ex post facto designs and, 251
 - grounded theory studies and, 147
 - inferential statistics and, 296–300
 - mixed-methods designs and, 259, 261–262
 - null hypothesis, 40, 297, 298, 299–300
 - pitfalls in reasoning and, 23
 - proposal writing and, 45

- research hypotheses, 3–4, 7, 40, 296–297, 298, 299–300
- research problem and, 3–4, 34, 39–42
- in research reports, 312, 314, 315, 316
- scientific method and, 19–20
- stating, 39–42
- statistical hypotheses, 296–297, 298, 299–300
- testing of, 17, 259, 296–300, 314
- theory building and, 20–21
- IACUC (institutional animal care and use committee), 109, 263
- Idiographic research. *See* Case studies
- Independent-samples *t*-test, 301, 346n
- Independent variables, 40–42, 226–227, 232, 243–245
- Indexes, 55
- Inductive reasoning, 18–19, 20, 23, 97
- Inferences, 12, 191, 277, 314
- Inferential statistics. *See also* Descriptive statistics; Statistical techniques; Statistics
 - estimating population parameters and, 294–296
 - examples of procedures, 301
 - function of, 10, 277
 - meta-analysis and, 300, 302
 - probabilities and, 296, 297, 298, 299, 299n, 315
 - in SPSS, 345–347
 - testing hypotheses and, 296–300
- Informed consent
 - data collection and, 151, 158
 - ethical issues and, 105–106, 128n, 250
 - ethnography and, 143
 - mixed-methods designs and, 263
 - placebos and, 230
- Informed consent forms, 106, 107
- Inspiration software, 38, 60–61, 313
- Instant messaging, 167
- Institutional animal care and use committee (IACUC), 109, 263
- Insubstantial phenomena, 81–84, 89, 191, 270, 272
- Interlibrary loan, 36, 63
- Internal consistency reliability, 91, 199, 199n
- Internal evidence, in historical research, 176–177
- Internal review boards (IRBs), 106, 108–109, 263
- Internal validity, 101–102, 103, 189, 227–229, 232, 234, 262
- Internet. *See also* Technology
 - bibliographic software and, 319
 - conducting experiments on, 250–251
 - content analysis and, 148n
 - data collection for descriptive research and, 205–206
 - ethnography and, 143, 143n
 - literature resources and, 334, 335, 336
 - literature review and, 53, 58–59
 - questionnaires and, 202, 217
 - referencing sources obtained on, 320
 - right to privacy and, 108, 158, 250
 - scheduling software and, 114
 - spreadsheet software and, 274
 - writing assistance, 310–311
- Interpretation. *See also* Data interpretation
 - correlational research and, 186, 187
 - ethnography and, 144
 - historical research and, 170
 - as purpose of qualitative research, 140
- Interquartile range, 287
- Interrater reliability, 91, 185
- Interval data, 279, 284, 290
- Interval estimates, point estimates *versus*, 296
- Interval scales, 86, 87
- Interval validity, 188–189
- Intervening variables, 41, 188
- Interviews
 - of expert researchers, 25
 - face-to-face interviews, 190–191, 191n
 - guidelines for conducting, 154–157
 - historical research and, 171
 - phenomenological study and, 145–146
 - qualitative research and, 140, 151, 153–158, 190
 - quantitative research and, 154, 190, 194–196
 - research questions aligned with research questions, 155–156
 - survey research and, 190–191
- Inverted pyramid, 67, 325
- IRBs (internal review boards), 106, 108–109, 263
- Italics, underlining *versus*, 46
- Jaccard, J., 11, 20, 146n, 147
- Jackson, D. L., 13, 14, 18, 252, 319, 321–322, 342
- Jacoby, J., 11, 20, 146n, 147
- Janos, P. M., 291
- Jeffrey, K., 173
- Jernigan, T. L., 57
- John, O. P., 206
- Johnson, B., 242n
- Johnson, J. M., 162
- Journal articles, publishing, 331–332
- Journals, as literature resources, 8, 24, 52, 53, 56, 57, 62, 63, 334, 335
- JSTOR, 57, 63
- Juried research reports, 24
- Kahn, P. G. K., 272
- Karabenick, S. A., 199
- Kearns, K. C., 171
- Keeter, S., 217
- Kellogg, R. T., 12
- Kendall coefficient of concordance, 292
- Kendall's tau correlation, 292
- Kennedy, C., 217
- Kennedy, John F., 5–6
- Key informants, 143
- Keywords, 52, 54, 58, 59, 62
- Kim-Cohen, J., 41
- Kime, N., 146
- Kinnick, V., 252–256
- Koch, R. S., 22
- Kontos, S., 185
- Kozinets, R. V., 143
- Krantz, J., 205
- Krathwohl, D. R., 303
- Kraut, R., 143n, 148n, 206
- Kruskal-Wallis test, 301
- Kuder-Richardson Formula 20, 199n
- Kuhn, D., 188n
- Kurtosis, 281
- Kvale, S., 157
- Language
 - questionnaires and, 197
 - as research tool, 11–12
- Lara, L. G., 265–269
- Latent variables, 82n
- Lauer, J. M., 139
- LCD projectors, 330, 330n
- Leavenworth, P. S., 31, 171, 172, 173
- Leptokurtic curve, 281, 282
- Letter of inquiry, for questionnaires, 202–203, 204
- Lieberman, N., 197
- Librarians, reference, 52, 55, 58
- Libraries
 - guidelines for efficient use of, 61–64
 - as research tools, 8, 24, 52–60
- Library catalogs, for literature review, 52–55
- Library of Celsus, 8
- Library of Congress, 174
- Library of Congress (LC) classification system, 24, 54
- Library of Nineveh, 8
- The Life of Sir Alexander Fleming* (Maurois), 23
- Likert, Rensis, 192
- Likert scales, 192, 193, 197
- Lincoln, Abraham, 176, 187
- Lincoln, Y. S., 104, 139
- Lind, D. A., 216
- Lindsay, J. J., 103n
- Line graphs, 273
- Line of regression, 186
- Lippa, R. A., 29
- Lipsey, M. W., 298n
- List servers, 21
- Literature reviews
 - cohesive review, 66–67
 - computers and, 9
 - conducting, 60–64
 - dissertation analysis, 70, 71–72

- Literature reviews (*continued*)
 evaluating others' research, 64–65
 finding research problems and, 30
 fine-tuning research problem and, 47
 knowing when to quit, 66
 mixed-methods designs and, 262
 organizing collected information, 65–66
 prose style for, 323–324
 role of, 51–52
 strategies for, 52–60
 word processors and, 15
 writing, 67–70
- Locke, E. A., 154
 Loftus, A., 199
 Loftus, E. F., 251
 Logical structure, in research reports, 324–325
 Longitudinal studies, 188, 188n, 189
 Lowes, John Livingston, 95, 178
 Lüdtke, O., 291
 Lundeberg, M., 199
 Lundqvist, A., 148
 Luong, A., 202, 204, 218
 Lytle, S., 27
- Making of America* (online database), 174
 Malthus, Thomas Robert, 284
 Manifest variables, 82n
 Mann-Whitney *U*, 301
 Mapland, 313
 Maptitude, 313
 Marius, R., 171, 177, 178
 Marsh, H. W., 291
 Masland, R., 19
 Matched pairs, 230–231, 237
 Maurois, André, 22, 23
 McCallin, R. C., 106, 107, 244, 261
 McCloskey, M., 20
 McCrea, S. M., 197
 McGee, M. A., 292
 McGibbon, E., 143
 McGraw, K. O., 206
 McGrew, K. S., 291
 McGue, M., 68, 332
 McInerney, D. M., 191
 McKenzie, M. G., 170–171, 172, 176–177, 179–182, 314
 McLoughlin, W. J., 133
 Meaning of events, 170
 Means
 formula for, 284
 geometric mean, 284–285, 286
 interval scales and, 86
 measures of central tendency and, 283–285
 measures of variability and, 287–290
 population means, 294–296
 statistics and, 276
 Measurement
 defining, 81–82
 identifying instruments for, 81–87
 insubstantial phenomena and, 81–84
 reliability of, 89, 91–93, 299, 316
 as research criteria, 77
 as research tool, 9–10
 substantial measurements, 81, 191, 272
 types of scales, 84–87
 validity of, 44n, 89–93, 299, 316
 Measures of association, 291–293
 Measures of central tendency
 curves determine means, 284–285
 descriptive statistics and, 283–286, 290
 population parameters and, 294
 as predictors, 286
 summary of, 286
 Measures of variability
 depiction of, 284
 descriptive statistics and, 286–290
 population parameters and, 294
 spread and, 287–288
 standard scores and, 288–290
 statistics appropriate for, 276
 summary of, 289
 Medawar, P. B., 29
 Median, 283, 284, 286, 287
 Mediating variables, 41–42, 188
 Mehan, H., 143, 144
 Memos, 151
 Merge function, in word processors, 201, 202
 Mertler, C. A., 27, 100
 Meta-analyses, 249–250, 300, 302
 Methodology, research. *See also* Qualitative research; Quantitative research
 case studies, 141
 content analysis and, 149
 data linked with, 94–98
 deciding on approach, 98–101
 definition of, 7
 ethnography and, 143–145
 functions of, 4–5
 grounded theory studies and, 147
 justifying, 112–113
 mixed-method design, 98
 phenomenological study and, 145–146
 planning *versus*, 76
 proposal writing and, 125
 in research reports, 312
 research tools confused with, 7–8
 weaknesses in, 315
 Microforms, 53
 Microgenetic studies, 188n
 Microsoft Excel
 data analysis and, 274
 data interpretation and, 159
 data recording and, 336–339
 formula tool in, 338–339
 literature resources and, 334–336
 reorganizing data in, 339
 simple statistical analyses in, 340–341
 tables and figures created in, 313
 Microsoft PowerPoint, 330
 Middleton, M., 198, 336
 Miksa, F. L., 55
 Milch-Reich, S., 252
 Miller, S. M., 99
 Mills, G. E., 27, 100, 215, 259n
 Miltenberger, R. G., 241
 Mind, human. *See also* Open-mindedness
 as research tool, 11, 16–21
 Mitchell, K. J., 304–307
 Mixed-methods designs
 combined experimental and ex post facto design, 244–245
 combining quantitative and qualitative designs, 98, 139, 195, 258
 content analysis and, 148–149
 data analysis and, 264–265
 data interpretation and, 264–265
 decisions concerning, 258–259, 263–264
 dissertation analysis, 265–269
 ethical issues and, 263
 identifying research questions and hypotheses, 261–262
 planning and, 260–263
 types of, 259–260
 MLA (Modern Language Association) style, 311
 Mode, 283, 286
 Moderate correlation, 291
 Moderating variables, 41–42, 188
 Moffitt, T. E., 41
 Mohan, L., 199
 Moon, J., 17
 Moore, M. T., 99
 Multi-group data, single-group data *versus*, 278
 Multiple baseline designs, 239–240, 241, 247
 Multiple case studies, 141
 Multiple correlation, 292
 Multiple drafts, 14, 48, 70
 Multiple linear regression, 301
 Multiple perspectives, 139
 Multiples, and ratio scales, 87
 Multistage sampling of areas, 215, 216
 Multitrait-multimethod approach, 90
 Munter, M., 330
 Murphy, K. R., 298n
 Myors, B., 298n
 Myrdal, G., 177
- Narrative research, 171
 NCE score, 290n
 Neatness, in research reports, 324
 Negative case analysis, 104, 152
 Negative correlation, 291
 Negatively skewed distribution, 281
 Nelson, M. W., 99
 Neuman, W. L., 30, 147, 154, 204
 Newton, Isaac, 20
New York Times Article Archive (online database), 57, 174
 Nicholls, M. E. R., 199
 Nichols, J. D., 21

- Nicol, A. A. M., 330
- Nominal data, 279
- Nominal scales, 85, 87
- Nonjuried research reports, 24
- Non-normal distributions, 280–282
- Nonparametric statistics, 282, 291, 292, 299, 300, 301
- Nonprobability sampling
 bias in, 218
 forms of, 214–215
- Nonrandomized control group pretest-posttest designs, 237–238, 246
- Normal curve, 280
- Normal distributions, 280–282, 284, 296
- Normative surveys, 189
- Norm-referenced scores, 289
- Note-taking forms, 61, 62
- Novelty effect, 102
- Null hypotheses, 40, 297, 298, 299–300
- Numerical data, 171, 173
- Nussbaum, E. M., 17
- Oakley Browne, M. A., 292
- Objectivity
 data admissibility and, 80
 data interpretation and, 315
 descriptive research and, 184–185
 emotion and, 23
 ethnography and, 145
 grounded theory studies and, 148
 hypotheses and, 39
 measurement and, 9–10
 qualitative research and, 9, 139, 153
- Observations
 qualitative research and, 140, 151, 152–153
 quantitative research and, 194
 simple time-series design and, 238
- Observation studies
 characteristics of, 100
 quantitative research and, 184–185
- Odds ratio, 301
- Okubo, M., 199
- One-group pretest-posttest designs, 233–234, 246
- OneNote, 66, 318
- One-shot experimental case studies, 233, 246
- Online databases
 historical research and, 174
 literature review and, 55–58
- Online document delivery services, 63
- Online journals, 53, 62
- Online library catalogs, 53
- Online Writing Lab (OWL), 310–311
- Open coding, in grounded theory studies, 147
- Open-mindedness
 grounded theory studies and, 147
 hypotheses and, 39
 proposal writing and, 125
 qualitative research and, 151, 162
 in stating research problem, 34
- Operational definitions, 44
- Oral history, 171
- Ordinal data, 279, 282, 284, 290
- Ordinal scales, 85–86, 87
- Ordinate, 186
- Ormrod, J. E., 11, 13, 20, 23, 198, 210, 218, 244, 261, 336
- Ormrod, R. K., 102, 174, 177, 244, 252, 261
- Orr, C. A., 199
- Outliners, in word processors, 15
- OWL (Online Writing Lab), 310–311
- Pajares, F., 41
- Palestrina, Giovanni Pierluigi da, 32, 37–38, 43
- Palmer, J. C., 251
- Paper-and-pencil approach
 to data gathering, 61, 62
 historical research and, 175
 questionnaires and, 191, 200
 to subproblems, 37–38, 60
- Paper copy, editing from, 16
- Paper proposals, 329
- Papers, presentation of, 329
- Paradi, D., 330
- Parallel structure, 131
- Parameters, 278
- Parametric statistics, 282, 291, 292, 299, 300, 301
- Partial correlation, 231, 292
- Participant observation, 143
- Participants, subjects *versus*, 236
- Pashler, H., 250
- Passive voice, 323
- Past tense, 323
- Path analysis, 301
- Pdfs (portable document formats), 27
- Pearson product moment correlation
 interval scales and, 86
 measures of association and, 291, 292
- Pearson *r*, 86, 291
- Pelham, W. E., Jr., 252
- People, as research problem category, 75
- Percentile, 287n
- Percentile ranks, 86, 281–282, 289
- Perfect correlation, 291
- Periodicals, 24, 53
- Peripheral devices, 194
- Permissions
 copyright and, 63, 69, 69n, 318
 ethical issues and, 76, 102, 105–106, 105n, 107, 108, 156, 194, 196, 203, 263
 research proposal and, 119–120
- Persuasiveness, as qualitative research
 evaluation criterion, 162
- Peshkin, A., 140
- Peter, E., 143
- Peterson, C., 98
- Peterson, L., 239
- Pexman, P. M., 330
- Phenomenological studies, 100, 145–146, 148, 150, 259
- Phi coefficient, 292
- Piaget, Jean, 12
- Pilot studies, 92, 112
- Pilot tests, for questionnaires, 195, 199–200
- Placebos, 128n, 229–230
- Plagiarism, 68–69, 108
- Planners, electronic, 114
- Planning
 computers and, 9
 for data collection, 80–81
 descriptive research and, 191
 ethical issues and, 104–110
 interviews, 194–196
 justifying research methodology and, 112–113
 linking data and research methodology, 94–98
 literature review writing and, 67–68
 measurement and, 81–93
 mixed-methods designs and, 260–263
 qualitative research and, 140, 160–162
 research approach and, 98–101
 research design and, 74–77
 research methodology *versus*, 76
 research projects and, 4–5
 research proposals and, 119–120
 research reports, 311–316
 role of data and, 77–81
 scheduling and, 113–116
 scrutinizing overall plan, 110–112
 validity and, 101–104
 word processors and, 15
- Plano Clark, V. L., 258, 259n, 262, 263
- Plato, 77n, 82
- Platykurtic curve, 281, 282
- Poe, Edgar Allan, 175
- Point biserial correlation, 292
- Point estimates, interval estimates *versus*, 296
- Points of central tendency, 283, 286. *See also* Measures of central tendency
- Polkinghorne, D. E., 145, 146
- Polymodal distributions, 285
- Pompea, S. M., 272
- Population characteristics, 214
- Population means, estimating, 294–296
- Population parameters
 inferential statistics and, 294–296
 population analysis, 219–220
 statistical notation for, 278
 statistics as estimates of, 278
 surveys of very large populations, 215
- Positive correlation, 291
- Positively skewed distribution, 281
- Poster sessions, 329, 330–331
- Posttest-only control group designs, 235–236, 246
- Pousette, A., 146

- Power, of statistical test, 298–299
- Practical applications
- cause-and-effect relationship determination, 248–249
 - checklists and rating scales, 192–193
 - communicating through writing, 13–16
 - computerizing observations, 194
 - data collection for descriptive research and, 205–206
 - evaluating others' research, 64–65
 - evaluating qualitative research studies, 163–164
 - historical data and, 175
 - historical research writing, 178
 - interviews in qualitative study, 154–158
 - literature review writing, 67–70
 - literature searches, 60–64
 - mixed-methods designs decisions, 263–264
 - planning and conducting interviews, 194–196
 - planning qualitative research, 160–162
 - population analysis for descriptive research, 219–220
 - presentations at professional conferences, 329–331
 - proposal writing, 45–47, 123–132
 - questionnaires and, 196–205
 - research problem identification, 29–35
 - research report critiquing, 327–328
 - strengthening research proposal, 132–133
 - tools in disciplines, 24–25
 - writing first sections of proposal, 45–47
 - writing schedule development, 325–326
- Practical significance, 44–45, 304
- Precision
- in research reports, 324
 - in stating research problem, 33
- Predetermined sequence, 212
- Predictions. *See also* Hypotheses
- measures of central tendency as, 286
- Pre-experimental designs, 232, 233–234, 246
- Preliminary pages, of research reports, 317–318, 321
- Premises, 17
- Presentations, of research reports, 329–331
- Present perfect tense, 323
- Present tense, 323
- Pretest-posttest control group designs, 234–235, 246
- Pretests, 230–231
- Primary area section, in multistage sampling, 215
- Primary data, 78, 170
- Primary sources, 170–171, 173, 174
- Principal thrust, of research, 312
- Privacy, right to, 107–108, 151, 158, 250, 316
- Probabilistic reasoning, 17
- Probability sampling
- cluster sampling, 211–212, 213
 - population characteristics and, 214
 - proportional stratified sampling, 211, 212
 - random selection and, 207–208
 - simple random sampling, 210–211
 - stratified random sampling, 211
 - systematic sampling, 212–213
- Problems, research. *See also* Subproblems
- context of, 124–125
 - data interpretation and, 5–6, 28, 80, 303, 304
 - delimiting, 43–44
 - delineation of, 39–45
 - finding legitimate problems, 29–31
 - finding research projects and, 27–29
 - fine-tuning, 47–48
 - hypotheses and, 3–4, 34, 39–42
 - identifying, 29–31
 - importance of study and, 44–45, 46
 - mixed-methods designs and, 259
 - qualitative research and, 140
 - questions and, 2, 39–42
 - reappraisal of, 47–48
 - in research reports, 312, 313, 316
 - setting of the problem, 39–45
 - statement of, 31–35, 52, 119, 207, 312
- ProCite, 66, 318
- Professional associations, websites of, 59
- Professional codes of ethics, 109
- Professional conferences, 30–31
- Project management software, 114
- Proofreading, 16
- Proportion, and population parameters, 294
- Proportional stratified sampling, 211, 212
- Proposals, research. *See also* Proposal writing
- characteristics of, 120–121
 - ordering topics in, 45
 - organization of, 121–123, 325
 - planning and, 119–120
 - sample, 134–135
 - strengthening, 132–133
 - weaknesses in, 133, 304
- Proposal writing
- challenges of, 134
 - collaboration with others and, 119
 - first draft, 124–128
 - first sections, 45–47
 - format of, 46, 121–123
 - revising, 128–132
 - strengthening proposal and, 132–133
- ProQuest Digital Dissertations, 63, 311, 317
- Prose style, 323–324
- Protection from harm, 105, 206, 250
- Pseudo-subproblems, subproblems *versus*, 36
- "Psychological Research on the Net" website, 205, 205n, 206, 250, 250n
- PsycINFO, 56–57, 63
- Publication information, in reference lists, 320
- Punctuation, 14, 15
- Purposefulness, as qualitative research evaluation criterion, 162
- Purposive sampling, 215
- Qualitative research. *See also* Mixed-method designs
- criteria for evaluating, 162
 - data analysis and, 97, 158–160
 - data collection and, 75, 97, 98–99, 140, 151–154
 - deciding on approach, 98–101
 - definition of, 95, 139
 - distinguishing characteristics of, 96–97
 - format of, 75
 - objectivity and, 9, 139, 153
 - planning and, 140, 160–162
 - process of, 97
 - proposal writing and, 125
 - purpose of, 96, 140, 150
 - quantitative research compared to, 95–98
 - research questions and, 39, 140
 - research reports and, 97, 312, 314, 321
 - validity in, 104
- Qualitative research designs
- case studies, 100, 141–142, 150, 233
 - content analysis, 100, 148–149, 150
 - distinguishing characteristics of, 150
 - ethnography, 100, 142–145, 150
 - grounded theory studies, 4, 100, 146–148, 150, 259, 262
 - phenomenological studies, 100, 145–146, 148, 150, 259
- Quantitative research. *See also* Descriptive research; Experimental research; Mixed-method designs
- data analysis and, 97, 270–308
 - data collection and, 97, 98–99
 - deciding on approach, 98–101
 - definition of, 95
 - distinguishing characteristics of, 96–97
 - process of, 97
 - proposal writing and, 125
 - purpose of, 96
 - qualitative research compared to, 95–98
 - research reports and, 97, 312, 314, 321
- Quartiles, 287
- Quasi-experimental designs, 231, 232, 237–241, 246–247
- Quasi-experimental research, 100
- Questionnaires
- bias and, 218
 - checklists or rating scales, 192–193
 - guidelines for constructing, 196–200
 - Internet and, 202, 217
 - mixed-methods designs and, 263
 - return rates and, 191, 196, 201–205
 - survey research and, 191
- QuickCalcs, 303
- Quota sampling, 214
- Quotations, 69–70
- Ramirez, I. L., 221–224
- Random numbers, table of, 208, 208n, 209, 210

- Random sampling, 210–211
- Random selection
 - descriptive research and, 230
 - population parameters and, 294, 295
 - posttest-only control group design and, 236
 - pretest-posttest control group design and, 234
 - probability sampling and, 207–208
 - quasi-experimental designs and, 237
 - random numbers table and, 209, 210
- Range
 - measures of variability and, 287
 - statistics and, 276
- Rank order, 85–86
- Rating scales, in descriptive research, 192–193, 194
- Ratio data, 279–280, 284
- Ratio scales, 86–87
- Raw score, 281–282, 288–290
- Reactivity, 102, 153, 229
- Real-life settings, 103, 139
- Realm of the Data, 78, 79
- Reasoning
 - inductive, 18–19, 20, 23, 97
 - pitfalls of, 22, 23
 - probabilistic, 17
 - verbal, 17
- Recoding, as spreadsheet function, 274–275
- Records, as research problem category, 75
- Reference librarians, 52, 55, 58
- Reference lists
 - citations in text and, 131, 318
 - creating computer database of, 65–66
 - format for, 310
 - literature review and, 59–60
 - in research reports, 318–321
 - URLs and, 59
- Reflexivity, 312
- Regression, 301
- Relativity, theory of, 6, 20
- Reliability
 - correlation coefficients and, 293
 - determining, 91–92
 - enhancing, 92–93
 - internal consistency reliability, 91, 199, 199n
 - interrater reliability, 91, 185
 - of measurement, 89, 91–92, 299, 316
 - of questionnaires, 199
- Reliability coefficients, 199n
- Repeated-measures designs, 231, 236–237
- Replication
 - data admissibility and, 79
 - external validity and, 103
 - meta-analyses and, 249–250
 - as research criteria, 76
 - research reports and, 312
- Reports, research
 - case studies and, 142
 - computers and, 9
 - conclusion of, 315–316
 - content analysis and, 149
 - critiquing, 327–328
 - data interpretation in, 311, 314–315
 - data presentation in, 313–314
 - description of method, 312
 - drafts of, 14
 - ethnography and, 144–145
 - feedback on, 326
 - figures and tables in, 14, 313, 313n
 - general principles of writing, 324–325
 - grounded theory studies and, 148
 - historical research and, 178
 - hypotheses and, 312, 314, 315, 316
 - juried, 24
 - language use and, 12
 - nonjuried, 24
 - objectives of, 310, 311
 - organization of, 321–323
 - phenomenological study and, 146
 - planning, 311–316
 - preliminary pages, 317–318, 321
 - preparation of, 323–325
 - presentations of, 329–331
 - prose style for, 323–324
 - published reports as models, 311
 - qualitative research and, 97, 312, 314, 321
 - quantitative research and, 97, 312, 314, 321
 - reference lists in, 318–321
 - related literature in, 312
 - research problems and, 312, 313, 316
 - results section, 297
 - revisions of, 326
 - styles for, 310, 311
 - writing guidelines, 13–15, 310
 - writing schedule for, 114, 325–326
- Representative samples, 103, 294
- Research
 - in academic disciplines, 24, 76
 - characteristics of, 2–7, 5n
 - checklist of evaluating, 64–65
 - delimiting, 43–44
 - fact transcription *versus*, 2
 - fact transportation *versus*, 2
 - formal research, 2
 - meanings of, 1
 - misconceptions about, 1–2
 - schedule for, 113–116
- Research cycle, 6–7, 19, 64
- Research design, 74–77. *See also* Qualitative research designs; *and other specific designs*
- Researcher as instrument, 139, 159
- Research hypotheses, 3–4, 7, 40, 296–297, 298, 299–300
- Research methodology. *See* Methodology, research
- Research problems. *See* Problems, research
- Research process, basic format of, 6–7, 74–75
- Research projects
 - checklist for evaluating, 46–47
 - feasibility of, 32–33, 110–112, 131
 - finding, 27–29
 - general criteria for, 76–77
 - hypotheses in, 3–4
 - planning and, 4–5
- Research proposals. *See* Proposals, research
- Research reports. *See* Reports, research
- Research tools
 - computers as, 8–9
 - development of, 259
 - fine-tuning research problem and, 47
 - human mind as, 11, 16–21
 - identifying, 24–25
 - language as, 11–12
 - libraries as, 8, 24, 52–60
 - measurement as, 9–10
 - research methodology contrasted with, 7–8
 - statistics as, 10–11, 270
- Resolution of puzzling findings, and mixed-method designs, 259
- Respondent validation, 104
- Response cards, for questionnaires, 203, 204
- Response rates, 190, 201–205, 218
- Return rates
 - academic integrity and, 316
 - questionnaires and, 191, 196, 201–205
- Reversal time-series designs, 238–239, 241, 247
- Reviewers' critiques, responding to, 332
- Revisions
 - multiple drafts and, 14, 48, 70
 - proposal writing and, 128–132
 - research reports and, 326
- Reyna, V. E., 153
- Right to privacy, 107–108, 151, 158, 250, 316
- Rigor, as qualitative research evaluation criterion, 162
- Rigorous subjectivity, 144
- The Road to Xanadu* (Lowes), 178
- Robinson, N. M., 291
- Robust statistical procedures, 282
- Rogelberg, S. G., 202, 204, 218
- Rohrer, D., 250
- Rosales-Ruiz, J., 251, 274, 275
- Rules for argument, 178
- Sackur, J., 153
- Sales, B. D., 105
- Sample location selection, in multistage sampling, 215
- Sample means, distribution of, 295, 296
- Samples
 - bias in, 206, 207, 294, 316
 - definition of, 206
 - descriptive research and, 206–219
 - qualitative research and, 152
 - representative, 103, 294
 - research proposals, 134–135
- Sample selection, 217–218

- Sample size, identifying sufficient, 298–299, 298n
- Sample statistics
notation for, 278
population parameters and, 294
- Sampling
bias in, 216–219
identifying sufficient sample size, 215–216
inferential statistics and, 277
nonprobability sampling, 214–215, 218
online studies and, 251
online surveys and, 206
probability sampling, 207–214
qualitative research and, 152
in surveys of very large populations, 215
- Sampling bias, 216–219
- Sampling designs, 207–215
- Saunders, M. G., 19
- Scales of measurement
interval scales, 86, 279
nominal scales, 85, 87, 279
ordinal scales, 85–86, 87, 279
ratio scales, 86–87, 279–280
types of, 84–87, 279–280
- Scanners, for data tabulation, 201
- Scatter plots, 186, 292
- Schallert, D. L., 146
- Schedule
for research, 113–116
for writing, 114, 325–326
- Schram, T. H., 147, 152, 158
- Schuman, H., 154, 156
- Schunk, D. H., 41
- Schwab, R. S., 19
- Schwarz, N., 153, 190
- Schwarzchild, Karl, 20
- Scientific method, 19–20
- Scott, K. M., 292, 293
- Scott-Jones, D., 105, 105n
- Search and replace features, in word processors, 15
- Search engines, 58
- Searching, keeping records of online searches, 62
- Secondary data, 78, 170
- Secondary sources, 170, 173
- Segment selection, in multistage sampling, 215
- Selective coding, in grounded theory studies, 147
- Self-questioning, 13, 18
- Self-report data, 190
- Semistructured interviews, 190
- SEM (structural equation modeling), 187, 188, 231, 301, 305
- Senders, V. L., 87
- Setting of the problem, 39, 45
- Settings, real-life, 103, 139
- Shaklee, J. M., 193
- Shanahan, T., 12
- Shank, G. D., 154, 157
- Sheehan, K., 202
- Sherman, S. J., 197
- Sieber, J. E., 105
- Sigman, M., 153
- Significance level, 297
- Silverman, D., 19, 153, 154
- Simple ex post factor designs, 242–243, 247
- Simple linear sequence, 273
- Simple random sampling, 210–211
- Simple time-series designs, 238, 247, 274
- Single-group data, multi-group data *versus*, 278
- Single-subject designs, 241
- Skagert, K., 146
- Skewed distribution, 281
- Skype, 191n
- Slides, for presentations, 330, 330n
- Smith, R. M., 164–168
- Sobel, D., 30
- Sociograms, 83–84, 85
- Software. *See also* Microsoft Excel; SPSS (Statistical Package for the Social Sciences); Word processors
bibliographic software, 318
brainstorming software, 38–39, 60
data analysis and, 274–276
freeware, 66, 114, 274, 319
project management software, 114
statistical software packages, 302–303
transcription software, 157–158
- Solomon, R. I., 235
- Solomon four-group designs, 235, 243, 246
- Sorting, as spreadsheet function, 274
- Sowell, E. R., 57, 58
- Spearman's rank order correlation (Spearman's rho), 86, 292
- Speculation, 4
- Spell checkers, 14, 15, 16, 130
- Spread, and measures of variability, 287–288
- Spreadsheets, 194, 274–276, 302–303, 338–339. *See also* Microsoft Excel
- SPSS (Statistical Package for the Social Sciences)
computing basic descriptive statistics in, 344–345
computing inferential statistics in, 345–347
creating data set in, 342–344
- Srivastava, A. K., 84
- Srivastava, S., 206
- Stake, R., 141
- Standard deviation
interval scales and, 86
measures of variability and, 288–290, 294
square of, 292
standard scores and, 288–290
statistics and, 276
- Standard error of the mean, 295–296
- Standardization, in measurement, 92
- Standard scores, 289–290, 290n
- Stanines, 290
- Stanley, J. C., 104, 228–229, 232, 232n
- Static group comparisons, 234, 242, 246
- Statistical analyses, Microsoft Excel for, 340–341
- Statistical hypotheses, 296–297, 298, 299–300
- Statistically significant, 297
- Statistical Package for the Social Sciences (SPSS). *See* SPSS (Statistical Package for the Social Sciences)
- Statistical significance, 89, 300, 304
- Statistical software packages, 302–303
- Statistical techniques. *See also* Data analysis; Descriptive statistics; Inferential statistics
data interpretation and, 303–304
in Microsoft Excel, 340–341
nature of data and, 278–282, 285, 291, 293
purpose of, 276–277
rationale for in research report, 313
for testing hypotheses, 300, 301
weaknesses in, 315
- Statistics. *See also* Descriptive statistics; Inferential statistics
choosing appropriate statistics, 276–282
confounding variables and, 231–232
definition of, 270, 278
function of, 277–278, 291
measurement scales and, 87
null hypotheses and, 40
as research tool, 10–11, 270
- Streiner, E., 20
- Stern, J. E., 332
- Stevens, S. S., 84
- Story line, in grounded theory studies, 147
- Stratified populations, 211
- Stratified random sampling, 211
- Strauss, A., 146, 146n, 147
- Strauss, A. C., 146n, 147, 148
- Strength, correlation coefficients for two variables, 291
- Stricker, J. M., 241
- Strong correlation, 291
- Strong Vocational Interest Blank (SVIB), 60–61
- Structural equation modeling (SEM), 187, 188, 231, 301, 305
- Structured interviews, 190
- Strunk, W., Jr., 332
- Student's *t* test, 301, 346n
- A Study of History* (Toynbee), 177–178
- Style of prose, 323–324
- Subheadings
attention to research problem and, 312
formatting, 122–123
- Subjects, participants *versus*, 236
- Subproblems
characteristics of, 36–37
data interpretation and, 37, 80
dividing research problem into, 3, 36–39
hypotheses and, 39

- identifying, 37–39
- literature review and, 60–61, 67
- mixed-methods designs and, 259
- proposal writing and, 45
- pseudo-subproblems *versus*, 36
- in research reports, 312, 313, 316
- Substantial measurements, 81, 191, 272
- Summaries
 - in abstracts, 317
 - in conclusions, 14, 316
 - in literature reviews, 70
- SurveyMonkey, 205–206
- Survey research, 100, 189–191
- Survey Research Center of the University of Michigan's Institute of Social Research, 215, 216–217
- Surveys
 - design of, 189–190
 - Internet and, 205–206
 - sampling in surveys of very large populations, 215
- SVIB (Strong Vocational Interest Blank), 60–61
- Swearer, S. M., 291
- Symbols
 - for ordinal scales, 85
 - for statistics, 278
 - word processors and, 124
- Systematic sampling, 212–213
- Table of contents, 318
- Table of random numbers, 208, 208n, 209, 210
- Table of specifications, 90
- Tables
 - definition of, 313
 - questionnaires and, 200
 - in research reports, 14, 313, 313n
 - two-dimensional tables, 273
 - word processors and, 15, 124
- Tau correlation, Kendall's, 292
- Taylor, A., 41
- Technology. *See also* Computers; Internet; Software
 - brainstorming software and, 38–39, 60
 - collaboration and, 21
 - computer databases facilitating data organization, 159–160
 - computerizing observations, 194
 - conducting experiments on the Internet, 250–251
 - creating computer database of related literature, 65–66
 - data analysis and, 200–201, 274–276
 - database for related literature, 65–66
 - data collection and transcription, 157–158
 - data collection for descriptive research and, 205–206
 - historical research and, 175
 - interviews and, 195
 - literature resources and, 334–336
 - literature review and, 55–59
 - obtaining sources with, 63
 - online databases and, 55–58, 174
 - online library catalogs and, 53
 - for questionnaire administration, 200–201
 - referencing sources obtained on Internet, 320
 - research process and, 8, 9
 - search engines and, 58–59
 - statistical software packages, 302–303
 - writing assistance, 310–311
- Telephone interviews, 190–191, 191n
- Template documents, 194
- Terms, defining, 43–44, 45, 312
- Tesch, R., 145
- Test-retest reliability, 91
- Tew, M. D., 206
- Theoretical sampling, 152
- Theories, defined, 20
- Theory-building process, 20–21
- Theory development, 147
- Theory of relativity, 6, 20
- Thesaurus feature, in word processors, 15, 130
- Thick description
 - qualitative research and, 162, 314
 - validity and, 104
- Things, as research problem category, 75
- Third person, 324
- Thompson, B., 292
- Thompson, K. R., 20
- Thompson, P. M., 57
- Thoughts and ideas
 - as research problem category, 75
 - writing as clarification of, 12
- Thrailkill, N. J., 68, 69, 227, 231
- Title of work, in reference lists, 319–320
- Title page, 317
- Toga, A. W., 57
- Topics, ordering, 45
- Toynbee, Arnold, 95, 177–178
- Track changes feature, in word processors, 15
- Trahan, R. G., 102, 192, 193, 252
- Transcription software, 157–158
- Trautwein, U., 291
- Treatment groups, 229
- Treatments, 226
- Trial-and-error explorations, as spreadsheet function, 276
- Triangulation, 102, 142, 162, 259
- Triola, M. E., 216
- Triserial correlation, 292
- Trope, Y., 197
- True experimental designs, 232, 234–237, 246
- Trustworthiness, 262
- Truth
 - academic integrity and, 316
 - data and, 77, 78, 79, 94, 170, 173
 - mixed-methods designs and, 259
 - objectivity and, 139
- Tufts University, 108
- Two-dimensional tables, 273
- Two-factor experimental design, 243–244, 247
- Two-phase projects, 260, 263
- Type I error (alpha error), 298–299, 299n, 314
- Type II error (beta error), 298–299, 315
- Underlining, italics *versus*, 46
- Understanding Quantitative History* (Haskins and Jeffrey), 173
- Uniform Resource Locators (URLs), 59, 320, 334
- Universality, as research criterion, 76
- Universe (population), 278
- University of New Hampshire, 108
- Unobtrusive measures, 102, 106
- URLs (Uniform Resource Locators), 59, 320, 334
- Usefulness, as qualitative research evaluation criterion, 162
- U.S. Government Printing Office, 59
- U-shaped relationships, 292, 293
- Uziel, L., 153
- Validity
 - considerations of, 101–104
 - correlation coefficients and, 293
 - determining, 90–91
 - enhancing, 92–93
 - external validity, 103–105, 207, 229, 235, 262
 - internal validity, 101–102, 103, 189, 227–228, 232, 234, 262
 - of measurement, 44n, 89–91, 299, 316
 - mixed-methods designs and, 262
 - qualitative research and, 104
 - of questionnaires, 199–200
- Values
 - population parameters and, 278
 - research reports and, 312
 - variables and, 279
- Vanderwood, M., 291
- Variability. *See also* Measures of variability
 - population parameters and, 294, 296
 - qualitative research and, 152
- Variables
 - confounding variables, 228, 316
 - continuous *versus* discrete variables, 279
 - correlational research and, 185–187
 - definition of, 40, 279
 - dependent variables, 40–42, 226–227, 232
 - developmental design and, 188
 - experimental design and, 226–232
 - factorial design and, 243–244
 - identifying, 40–42
 - independent variables, 40–42, 226–227, 232, 243–245
 - latent variables, 82n
 - manifest variables, 82n

- Variables (*continued*)
 - measures of association and, 291–293
 - mediating variables, 41–42, 188
- Variance, 288, 291–292
- Vazire, S., 206
- Verbal reasoning, 17
- Verification, as purpose of qualitative research, 140
- Video conferencing, 157, 191n
- Visuals, for presentations, 330
- Voice recognition software, 157–158
- Voluntary participation, 105–106
- Vul, E., 250
- Vygotsky, Lev, 12

- Wagner, E. D., 244, 261
- Walton, D. N., 20
- Wasserman, S., 84
- Weak correlation, 291
- Web of Science, 57–58
- Web pages, 21, 59, 320, 334, 335, 336
- Wells, J. E., 292
- Wennick, A., 148
- “What ifs,” as spreadsheet function, 276

- White, E. B., 332
- Wikipedia, 59
- Wilcoxon signed-rank test, 301
- Williams, J. E., 206
- Within-subjects designs, 231, 236–237, 246
- Within-subjects variable, 231
- Witt, E., 218
- Wixted, J. T., 250
- Wolach, A., 298n
- Wolcott, H. E., 139, 144, 145, 314
- Wong, P. T. P., 133
- Wood, H., 144
- Word processors
 - copy and paste function, 174
 - footnotes created in, 324
 - guidelines for using, 15–16
 - importing data and, 274
 - for proposal writing, 124
 - prose style and, 323–324
 - questionnaires and, 201
 - saving documents, 15–16
 - search function, 175
 - tables and figures created in, 15, 124, 313, 324

- templates and, 194
- Working memory, 11n
- WorldCat, 58
- World Wide Web (WWW). *See* Internet
- Writing. *See also* Editing; Proposal writing; Reports, research
 - assistance with, 310–311
 - communicating effectively through, 13–16
 - general principles of, 324–325
 - guidelines for, 13–15, 310
 - importance of, 12, 132
 - literature reviews, 67–70
 - multiple drafts and, 14, 48, 70
 - qualitative research and, 100
 - schedule for, 114, 325–326
 - styles for, 310, 311

- χ^2 (chi-square) goodness-of-fit test, 301

- Zambo, D., 14, 154, 155–156, 322–323
- Zeith, T. Z., 291
- Zero point, absolute, 87, 279
- Zoomerang, 205–206
- z-score, 289–290