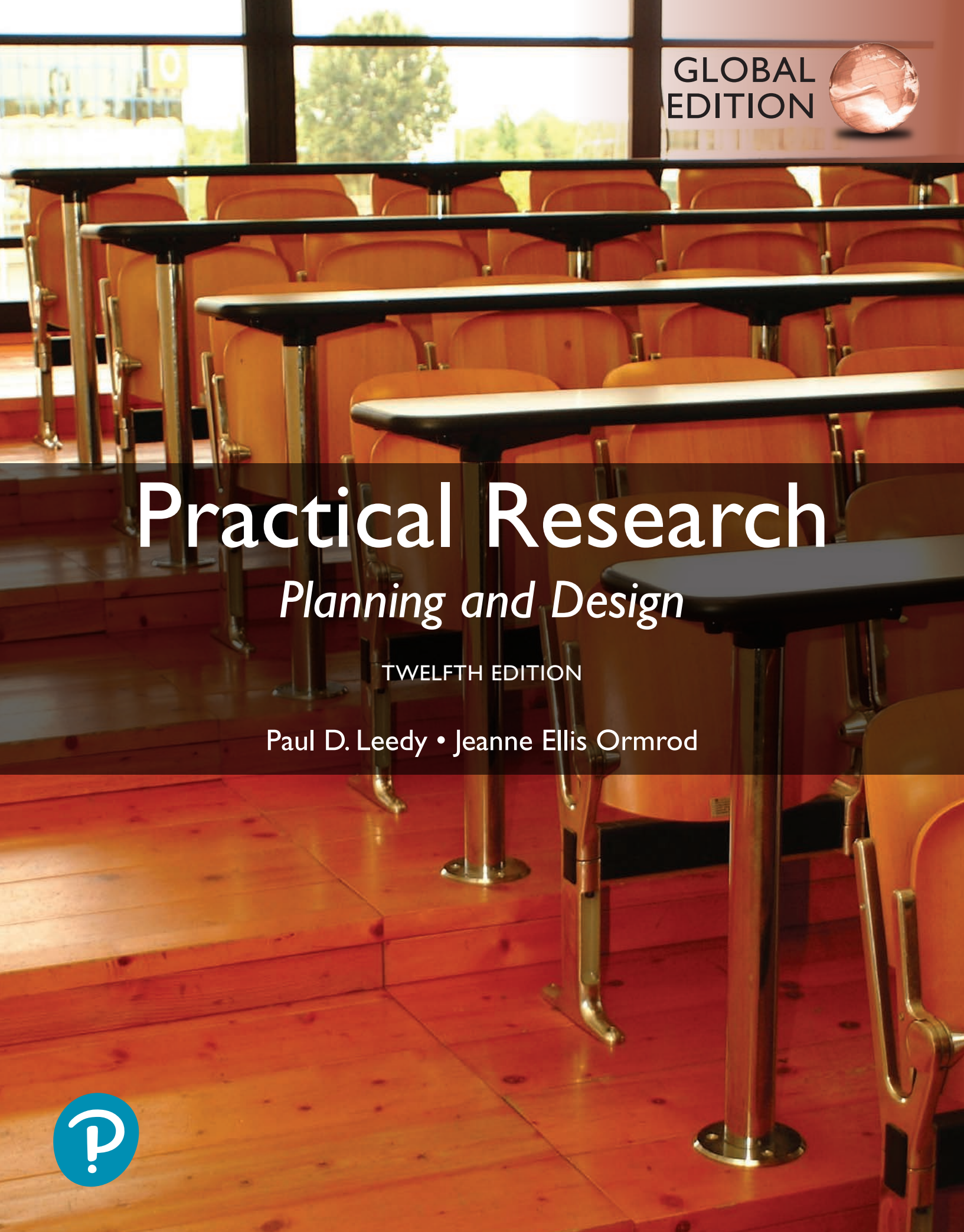


GLOBAL EDITION

Practical Research

Planning and Design

TWELFTH EDITION

Paul D. Leedy • Jeanne Ellis Ormrod

TWELFTH EDITION
GLOBAL EDITION

PRACTICAL RESEARCH

PLANNING AND DESIGN

Paul D. Leedy
Late of American University

and

Jeanne Ellis Ormrod
University of Northern Colorado (Emerita)

with **Laura Ruth Johnson**
Northern Illinois University

Pearson

Harlow, England • London • New York • Boston • San Francisco • Toronto • Sydney • Dubai • Singapore • Hong Kong
Tokyo • Seoul • Taipei • New Delhi • Cape Town • São Paulo • Mexico City • Madrid • Amsterdam • Munich • Paris • Milan

Please contact https://support.pearson.com/getsupport/s/contactsupport with any queries on this content.

Pearson Education Limited
KAO Two
KAO Park
Hockham Way
Harlow
Essex
CM17 9SR
United Kingdom

and Associated Companies throughout the world

Visit us on the World Wide Web at: www.pearsonglobaleditions.com

© Pearson Education Limited, 2021

Authorized adaptation from the United States edition, entitled Practical Research: Planning and Design, 12th Edition, ISBN 978-0-13-477565-4 by Paul D. Leedy and Jeanne Ellis Ormrod published by Pearson Education © 2019.

Acknowledgments of third-party content appear on the appropriate page within the text, which constitutes an extension of this copyright page. Cover Image: © D-VISIONS/Shutterstock

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without either the prior written permission of the publisher or a license permitting restricted copying in the United Kingdom issued by the Copyright Licensing Agency Ltd, Saffron House, 6–10 Kirby Street, London EC1N 8TS. 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 otherwise. For information regarding permissions, request forms, and the appropriate contacts within the Pearson Education Global Rights and Permissions department, please visit www.pearsoned.com/permissions/.

All trademarks used herein are the property of their respective owners. The use of any trademark in this text does not vest in the author or publisher any trademark ownership rights in such trademarks, nor does the use of such trademarks imply any affiliation with or endorsement of this book by such owners.

This eBook is a standalone product and may or may not include all assets that were part of the print version. It also does not provide access to other Pearson digital products like MyLab and Mastering. The publisher reserves the right to remove any material in this eBook at any time.

ISBN 10: 1-292-33924-1
ISBN 13: 978-1-292-33924-5
eBook ISBN 13: 978-1-292-33928-3

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Typeset in Garamond 3 LT Pro 11/12.5 by SPi Global

Preface

NEW TO THE TWELFTH EDITION

Every year brings new strategies for research design, data collection, and data analysis. Accordingly, this twelfth edition of the book has been revised in many ways. Discussions of some topics have been expanded—often with new, illustrative examples—and new topics have been added; meanwhile, sections that few of our readers were finding useful have been either reduced in length or eliminated altogether. Technology-based strategies have been updated to include new software options. And as always, every page has been revisited—every word, in fact—and many minor changes have been made to tighten the prose or enhance its clarity and readability.

Especially noteworthy changes in this edition are (a) a better balance between quantitative and qualitative methods than was true for the eleventh edition; (b) the addition of a new chapter on action research, with room for it being made by the elimination of the chapter on historical research (which reviewers have almost unanimously been telling us they don't assign in their classes); and (c) a reorganization of what were formerly Parts III, IV, and V (Chapters 6 through 12) into a new Part III ("Research Designs," with Chapters 6 through 10) and a new Part IV ("Data Analyses," with Chapters 11 and 12). The last of the changes just listed—the reorganization of chapters—was due in large part to the fact that researchers are increasingly drawing from both quantitative and qualitative traditions in their efforts to address important research problems and questions.

Other significant changes in this twelfth edition are the following:

- **Chapter 1.** Revised discussions of Step 1 and Step 4 in the research cycle, with Figure 1.1 also being revised accordingly; expansion of the section on philosophical assumptions to include *phenomenology* and *action-research* orientations; replacement of the key term *hypothesis* with *research hypothesis* to contrast it with the less formal hypotheses of everyday life; introduction of *purpose statement* as a key term; replacement of the key term *juried* with the term *peer-reviewed*, to reflect more popular terminology.
- **Chapter 2.** Discussion of research problems broadened to include research *questions* (the term more commonly used in qualitative research) and *purpose statements*; updated and expanded discussion of mind-mapping software, with new illustrative example (Figure 2.3); new section on identifying a *theoretical* or *conceptual framework* for a research study, along with (a) a new example and graphic illustrating the nature of a conceptual framework and (b) an additional suggestion to identify or create a theoretical/conceptual framework in the Practical Application feature "Writing the First Section of a Proposal."
- **Chapter 3.** Introduction of key term *open-access journal*; updated and expanded discussion of online databases; movement of what was formerly Table 13.1 ("Commonly Used Styles in Research Reports") to this chapter, where it is now Table 3.3; more specific recommendations for annotating sources during an in-house or online library search; new paragraph regarding the importance of writing an honest, nonbiased literature review; use of excerpts from a more current (2016) doctoral dissertation in the Dissertation Analysis feature.

- **Chapter 4.** Addition of the concept *unit of analysis* as a key term; explicit distinction between the key terms *assessment* and *measurement* (to create a better balance between qualitative and quantitative strategies in this and succeeding chapters); overhaul of the eleventh edition's section "Considering the Validity of Your Method" to give a better balance between qualitative and quantitative approaches, with the new heading "Enhancing the Credibility of Your Findings" (the in-depth discussion of *internal validity* in this section has been moved to Chapter 7); addition of follow-up studies as a strategy for enhancing the credibility of a research project; revision of the eleventh edition's section "External Validity" to offer a better balance between quantitative and qualitative methods, with a new heading "Enhancing the Generalizability of Your Findings"; renaming of the eleventh edition's section "Identifying Measurement Strategies" to "Choosing Appropriate Assessment Strategies," with a reorganization and many revisions to achieve a better quantitative/qualitative balance; new "Examples" column in Table 4.4 ("Contrasting the Four Types of Measurement Scales"); in many instances, replacement of the term *assessment instrument* with the more inclusive term *assessment strategy*; revision of the eleventh edition's section "Validity and Reliability in Measurement" (including revisions of glossary definitions of several key terms) to reflect a better balance between quantitative and qualitative approaches, with a heading change to "Validity and Reliability in Assessment."
- **Chapter 5.** Several minor changes in response to reviewers' concerns regarding (a) possible organizational structures for a research proposal, (b) the need for a reminder about not plagiarizing from other sources, and (c) the distinction between a reference list and a bibliography (in a new footnote).
- **Chapter 6.** Revision of the section on surveys to encompass qualitative as well as quantitative data collection; new section on *experience-sampling methods* (ESMs); in the section "Nonprobability Sampling," a new paragraph regarding the use of social media (e.g., Facebook) as a possible strategy for recruiting participants; significant revision of the eleventh edition's checklist "Analyzing Characteristics of the Population Being Studied" to include nonprobabilistic as well as probabilistic sampling, with the new title "Considering the Nature of the Population When Identifying Your Sampling Procedure"; addition of the *social desirability effect* as a bold-faced key term (with an accompanying hotlinked glossary definition) as a way of increasing its salience within the chapter.
- **Chapter 7.** Addition of a description of double-blind experiments and an in-depth discussion of internal validity (both discussions were previously in Chapter 4); relabeling of Designs 4 and 6 as *control-group pretest–posttest design* and *control-group posttest-only design,* respectively, to make them parallel to the labels for Designs 2 and 8; "down-grade" of term *repeated-measures design* from its previous key-term status (because the term *repeated measures* is increasingly being used in descriptive studies as well, especially in experience-sampling methods); switch from term *single-subject design* to the more contemporary and inclusive term *single-case intervention research,* with a new illustrative example that encompasses collection of data for both a single group and individual members of that group; addition of a reminder to *interpret* one's results regarding their relevance to one's initial research problem.
- **Chapter 8 (formerly Chapter 9).** Expansion of section on ethnographies to include *autoethnographies* and *reflective journals*; expansion of the section on phenomenological studies to include the *three-interview series* strategy for data collection; new section on *narrative inquiry*; expanded discussion of memos to include three types (reflective, methodological, and analytical memos); significant revision of the first Practical Application feature to include the key terms *credibility*, *transferability*, *dependability*, *confirmability*, *member checking*, and *audit trails*, along with the new title "Ensuring That Qualitative Data Collection Yields Credible, Trustworthy Data"; addition of the key terms *theoretical sampling*, *discriminant sampling*, *primary informants* (a.k.a.,

key informants), *extreme case sampling*, *convenience sampling*, and *snowball sampling* in the Practical Application feature "Selecting an Appropriate Sample for a Qualitative Study"; addition of *interview guide* as a key term in the guidelines for "Conducting a Productive Interview."

- **Chapter 9 (formerly Chapter 12).** New section on *longitudinal mixed-methods designs*; elimination of embedded designs as a category distinct from convergent designs (in line with Creswell's recent revisions of design categories); additional illustrative examples of mixed-methods research; expansion of Conceptual Analysis exercise "Identifying Mixed-Methods Research Designs" to include an example of a longitudinal mixed-methods design; movement of sections on "Analyzing and Interpreting Mixed-Methods Data" and "Systematic Reviews of Qualitative and Mixed-Methods Studies" to Chapter 12.
- **Chapter 10 (new chapter).** In-depth discussion of *action research* and *participatory designs*, which includes *teacher research*, *design-based research* (DBR), *participatory action research* (PAR), *youth participatory action research* (YPAR), and three distinct forms of *community-based research* (CBR); new Conceptual Analysis feature "Choosing an Action-Oriented Design"; new Practical Application feature "Deciding Whether to Use an Action Research and/or Participatory Design"; new sections "Data Collection and Analysis in Action Research" and "Disseminating the Findings of Action Research Projects"; new Practical Application feature "Using Community Forums as a Means of Disseminating the Results of Action Research and Participatory Research Projects"; new hypothetical sample research report, with the usual side commentary.
- **Chapter 11 (formerly Chapter 8).** Simplification of Figure 11.10 (formerly Figure 8.10) to enhance its readability; addition of using a repeated-measures variable as a strategy for enhancing the power of statistical analyses; new paragraph regarding *data dredging* (a.k.a. *p-hacking*) as a generally inappropriate and potentially unethical practice; updated list of popular statistical software programs.
- **Chapter 12 (formerly Chapter 11).** Chapter title now "Analyzing Qualitative and Mixed-Methods Data"; substantially updated and expanded discussion of qualitative data analysis strategies; new example illustrating data analysis in an ethnographic study; addition and revision of sections "Analyzing and Interpreting Mixed-Methods Data," "Using Computer Software to Facilitate Mixed-Methods Data Analysis," and "Systematic Reviews of Qualitative and Mixed-Methods Studies" (all of which were previously in the chapter on mixed-methods research).
- **Chapter 13.** Movement of Table 13.1 (regarding style manuals) to Chapter 3 (where it is now Table 3.3); substantial revision of the checklist "Criteria for Critiquing a Research Report" so that it more even-handedly addresses important qualities of qualitative and mixed-methods research as well as those of quantitative research.
- **Appendix A.** Update that describes the use of Microsoft Excel 2016 for Macintosh (rather than the 2008 version used in the eleventh edition of the book); update of Figure A.1 to include more recent literature sources.
- **Application Exercises.** Included at the end of each chapter, these exercises help readers implement their understanding of various research tasks essential to conducting research.

MyLab FOR EDUCATION

One of the most visible changes in the new edition, also one of the most significant, is the expansion of the digital learning and assessment resources embedded in the etext and the inclusion of *MyLab* in the text. *MyLab for Education* is an online homework, tutorial, and assessment program designed to work with the text to engage learners and to improve

learning. Within its structured environment, learners apply what they learn, test their understanding, and receive feedback to guide their learning and to ensure their mastery of key learning outcomes. The *MyLab* portion of the new edition of *Practical Research* is designed to help learners (1) understand the basic vocabulary of educational research and (2) get guided practice in planning and developing a research proposal and in collecting and analyzing research data. The resources in *MyLab for Education* with *Practical Research* include:

- *Self-Check* assessments with feedback throughout the etext help readers assess how well they have mastered content.
- *Application Exercises* allow the reader to practice research tasks from selecting and refining a topic, doing a literature review, collecting and analyzing research data, through writing up a proposal.
- *StatPak*, an easy to use statistical tool, enables students to enter sets of data and calculate common statistics automatically.

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 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 at specimens through a microscope, a psychologist by administering certain tests or systematically observing people's behavior, and an anthropologist by examining artifacts from a particular cultural group and perhaps from an earlier time period. Otherwise, the basic approach to research is the same. Regardless of the discipline, the researcher identifies a problem or question in need of a solution, collects data potentially relevant to the solution, 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, 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 they can effectively and professionally conduct their own research projects. 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 try to answer a question that has never been answered before and, in the process, must discover something that no one else has ever discovered.

Research has one end: the discovery of some sort of "truth." Its purpose is to learn what has never before been known; 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 university degree.

Many students have found *Practical Research* quite helpful in their efforts both to understand the nature of the research process and to 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 a course on 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 research methods course as one of the most rewarding and practical courses in your entire educational experience.

Acknowledgments

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. But some individuals have especially stood out. First and foremost, I must thank Laura Ruth Johnson (Northern Illinois University), whose particular expertise in qualitative and mixed-methods research has enhanced the content of the two qualitative chapters (Chapters 8 and 12) in ways I could not possibly have done myself; Laura has also written most of the new chapter on action research (Chapter 10). Furthermore, Laura offered many helpful suggestions for how I might give a better quantitative-qualitative balance to Chapters 1 through 5. And she has either created or revised each of the hotlinked Self-Check quizzes and Application Exercises that appear in the book—and done so under very tight time constraints. Laura's contributions to this edition have been substantial enough that I have rightly credited her with a "with" status on the book's title page.

I must also thank Hadley Solomon (University of New Hampshire) both for the conceptual framework graphic used in Figure 2.3 and for her wise counsel and suggestions regarding the discussion of philosophical assumptions in Chapter 1.

In addition, those of you who have written in journals and textbooks about research methods and strategies, the generations of graduate and undergraduate students whom we authors have taught and who have also taught *us*, the kindly letters and e-mail messages that so many of you have written to suggest how we might improve on the book—to all of you, I extend my acknowledgment and appreciation wherever you may be. You have had the greater part in bringing this book through its previous 11 editions. I am especially grateful to reviewers of the 11th edition, who have offered many good suggestions for strengthening the book so that it can better assist novice researchers in the 21st century: Natasha V. Christie (University of North Florida), Rachel D. Goodman (George Mason University), Darren Liu (University of Nevada, Las Vegas), and Robert N. Ronau (University of Cincinnati).

I am also indebted to the students whose research proposals, doctoral dissertations, and master's theses have enabled me to illustrate some of the research and writing strategies described in the book. In particular, I extend my gratitude to Rosenna Bakari, Arthur Benton, Douglas Burtorff, Jennifer Chandler, Chelsie Hess, Dinah Jackson, Ginny Kinnick, Laura Lara-Brady, Christy Leung, Kimberly Mitchell, Luis Ramirez, Janie Shaklee, Nancy Thrailkill, and Debby Zambo. And I must certainly mention Hannah Wilson and Ernie Badger, who drew the maps used in Figure 12.3.

Equally important is to say "Thank you, thank you, thank you" to folks at Pearson Education and Cenveo Publisher Services who have been key players in bringing this book to fruition. In particular, I extend my deepest gratitude to Janelle Rogers, Pam Bennett, and Norine Strang, each of whom has expertly overseen execution of the many nitty-gritty details of the publishing process. Finally, I must thank our editor, Kevin Davis, for his guidance throughout this and preceding editions. Throughout its many editions, Kevin has shared Paul's and my vision for the book and struck the ever-so-important balance between providing substantive guidance.

Paul and I have had many hands guiding our pens plus many minds adding richness and depth to our thoughts. For that, I offer my humble and hearty thanks.

Jeanne Ellis Ormrod

Global Edition Acknowledgments

Pearson would like to thank the following people for their work on the Global Edition:

Contributor

Kriti Kaul, Defence Research and Development Organization

Reviewer

Lambros Lazuras, Sheffield Hallam University

This page intentionally left blank

Brief Contents

Preface 3

Acknowledgments 8

PART I The Fundamentals

CHAPTER 1 The Nature and Tools of Research 23

PART II Focusing Your Research Efforts

CHAPTER 2 The Problem: The Heart of the Research Process 52

CHAPTER 3 Review of the Related Literature 82

CHAPTER 4 Planning Your Research Project 106

CHAPTER 5 Writing the Research Proposal 151

PART III Research Designs

CHAPTER 6 Descriptive Research 174

CHAPTER 7 Experimental, Quasi-Experimental, and Ex Post Facto Designs 220

CHAPTER 8 Qualitative Research Methods 258

CHAPTER 9 Mixed-Methods Designs 291

CHAPTER 10 Action Research and Participatory Designs 314

PART IV Data Analyses

CHAPTER 11 Analyzing Quantitative Data 337

CHAPTER 12 Analyzing Qualitative and Mixed-Methods Data 380

PART V Research Reports

CHAPTER 13 Planning and Preparing a Final Research Report 410

Appendices

Appendix A Using a Spreadsheet: Microsoft Excel 436

Appendix B Using SPSS 444

Glossary 450

References 460

Index 467

Contents

Preface 3

Acknowledgments 8

PART I The Fundamentals

Chapter 1

The Nature and Tools of Research 23

What Research Is Not 23
What Research Is 24
Philosophical Assumptions Underlying Research Methodologies 30
Tools of Research 32
The Library and Its Resources 32
Computer Technology 33
Measurement 33
Statistics 33
Language 35
PRACTICAL APPLICATION: Communicating Effectively Through Writing 37
GUIDELINES: Writing to Communicate 37
GUIDELINES: Using the Tools in Word-Processing Software 39
The Human Mind 41
Critical Thinking 41
Deductive Logic 42
Inductive Reasoning 43
Scientific Method 44
Theory Building 45
Collaboration with Other Minds 46
Reflections on Noteworthy Research 48
Exploring Research in Your Field 48
PRACTICAL APPLICATION: Identifying Important Tools in Your Discipline 49
CHECKLIST: Interviewing an Expert Researcher 50
Application Exercises 50
For Further Reading 51

PART II Focusing Your Research Efforts

Chapter 2

The Problem: The Heart of the Research Process 52

Finding Research Projects 52

PRACTICAL APPLICATION: Identifying and Presenting the Research Problem or Question 54

GUIDELINES: Choosing an Appropriate Problem or Question 54

GUIDELINES: Writing a Purpose Statement 56

CHECKLIST: Evaluating a Purpose Statement 60

Dividing the Main Research Problem or Question into Subproblems 61

Subproblems Versus Pseudo-Subproblems 62

Characteristics of Subproblems 62

Identifying Subproblems 63

Taking a Paper-and-Pencil Approach 63

Using Mind-Mapping (Brainstorming) Software 64

Every Research Problem Needs Further Delineation 65

Identifying a Relevant Theoretical or Conceptual Framework 66

Stating Hypotheses 67

Distinguishing Between Research Hypotheses and Null Hypotheses in Quantitative Research 67

Identifying the General Concepts and Possibly Also Specific Variables That Are the Focus of the Investigation 68

CONCEPTUAL ANALYSIS EXERCISE: Identifying Independent, Dependent, Mediating, and Moderating Variables 71

Defining Terms 71

Stating Assumptions 72

Identifying Delimitations and Limitations 72

Importance of the Study 73

Writing the First Chapter or Section of a Research Proposal 73

PRACTICAL APPLICATION: Writing the First Section of a Proposal 74

CHECKLIST: Evaluating Your Proposed Research Project 76

PRACTICAL APPLICATION: Reappraising a Proposed Research Problem 77

GUIDELINES: Fine-Tuning Your Research Problem or Question 77

Application Exercises 78

For Further Reading 79

Chapter 3

Review of the Related Literature. 82

Understanding the Role of the Literature Review 82

Strategies for Locating Related Literature 83

Using the Library Catalog 84

Using Online Databases 86

Consulting with Reference Librarians 89
Surfing the Internet 90
Using Citations and Reference Lists of Scholars Who Have Gone Before You 91
Identifying Your Sources: The Importance of Adhering to an Appropriate Style Manual 91
PRACTICAL APPLICATION: Planning a Literature Search 92
GUIDELINES: Using Your Library Time Efficiently 93
PRACTICAL APPLICATION: Evaluating the Research of Others 96
CHECKLIST: Evaluating a Research Article 97
Knowing When to Quit 98
Organizing and Synthesizing the Literature into a Cohesive Review 98
PRACTICAL APPLICATION: Writing the Literature Review 99
GUIDELINES: Writing a Clear and Cohesive Literature Review 99
A Sample Literature Review 102
Application Exercises 105
For Further Reading 105

Chapter 4
Planning Your Research Project 106

Planning a General Approach 107
Research Planning Versus Research Methodology 107
The Nature and Role of Data in Research 108
Data Are Transient and Ever Changing 108
Primary Data Versus Secondary Data 109
Planning for Data Collection 110
Linking Data and Research Methodology 112
Comparing Quantitative and Qualitative Methodologies 112
Combining Quantitative and Qualitative Designs 114
PRACTICAL APPLICATION: Choosing a General Research Approach 115
GUIDELINES: Deciding Whether to Use a Quantitative or Qualitative Approach 115
Choosing Methods That Enhance the Value of Your Research Project 117
Enhancing the Credibility of Your Findings 117
Enhancing the Generalizability of Your Findings 119
Choosing Appropriate Assessment Strategies 120
Distinguishing Between Substantial and Intangible Phenomena 120
Assessing Intangible Phenomena: An Example 121
Types of Measurement Scales 123
Nominal Scales 123
Ordinal Scales 124
Interval Scales 124
Ratio Scales 125
CONCEPTUAL ANALYSIS EXERCISE: Identifying Scales of Measurement 126
Validity and Reliability in Assessment 127

Validity of Assessment Strategies 128
Reliability of Assessment Strategies 130
Thinking of Validity and Reliability as Reflections of Error in Your Assessments 133
CONCEPTUAL ANALYSIS EXERCISE: Identifying Problems with Validity and Reliability in Assessment Strategies 133
Ethical Issues in Research 135
Protection from Harm 135
Voluntary and Informed Participation 136
Right to Privacy 138
Honesty with Professional Colleagues 139
Institutional Review Boards 139
Professional Codes of Ethics 140
PRACTICAL APPLICATION: Planning an Ethical Research Study 140
CHECKLIST: Determining Whether Your Proposed Study Is Ethically Defensible 140
Critically Scrutinizing Your Overall Plan 141
PRACTICAL APPLICATION: Judging the Feasibility of a Research Project 142
CHECKLIST: Determining Whether a Proposed Research Project Is Realistic and Practical 142
When You Can't Anticipate Everything in Advance: The Value of a Pilot Study 143
PRACTICAL APPLICATION: Developing a Plan of Attack 144
Using Project Management Software and Electronic Planners 144
Keeping an Optimistic and Task-Oriented Outlook 146
Application Exercises 146
For Further Reading 147

Chapter 5
Writing the Research Proposal 151
Characteristics of a Proposal 152
A Proposal Is a Straightforward Document 152
A Proposal Is Not a Literary Production 153
A Proposal Is Clearly Organized 153
Organizing and Writing a Research Proposal 153
Formatting Headings and Subheadings 154
PRACTICAL APPLICATION: Writing Your Proposal 155
GUIDELINES: Writing the First Draft 156
GUIDELINES: Revising Your Proposal 160
PRACTICAL APPLICATION: Strengthening Your Proposal 165
CHECKLIST: Evaluating an Early Draft of a Research Proposal 166
Final Thoughts About Proposal Writing 166
A Sample Research Proposal 167
Application Exercises 171
For Further Reading 173

PART III Research Designs

Chapter 6

Descriptive Research 174

Descriptive Research Designs 174

Observation Studies 175

Correlational Research 176

A Caution About Interpreting Correlational Results 177

Developmental Designs 178

Experience-Sampling Methods 180

Survey Research 181

Face-to-Face and Telephone Interviews 181

Questionnaires 182

Planning for Data Collection in a Descriptive Study 182

PRACTICAL APPLICATION: Using Checklists, Rating Scales, and Rubrics 183

PRACTICAL APPLICATION: Computerizing Observations 186

PRACTICAL APPLICATION: Planning and Conducting Interviews in a Quantitative Study 187

GUIDELINES: Conducting Interviews in a Quantitative Study 187

PRACTICAL APPLICATION: Constructing and Administering a Questionnaire 188

GUIDELINES: Constructing a Questionnaire 189

GUIDELINES: Using Technology to Facilitate Questionnaire Administration and Data Analysis 193

GUIDELINES: Maximizing Your Return Rate for a Questionnaire 194

PRACTICAL APPLICATION: Using the Internet to Collect Data for a Descriptive Study 198

Choosing a Sample in a Descriptive Study 199

Sampling Designs 200

Probability Sampling 200

Nonprobability Sampling 205

Sampling in Surveys of Very Large Populations 206

PRACTICAL APPLICATION: Identifying a Sufficient Sample Size 207

PRACTICAL APPLICATION: Choosing an Appropriate Sample in Descriptive Research 208

CHECKLIST: Considering the Nature of the Population When Identifying Your Sampling Procedure 208

Common Sources of Bias in Descriptive Studies 209

Sampling Bias 209

Instrumentation Bias 211

Response Bias 211

Researcher Bias 212

PRACTICAL APPLICATION: Acknowledging the Probable Presence of Bias in Descriptive Research 212

GUIDELINES: Identifying Possible Sampling Bias in Questionnaire Research 212
CHECKLIST: Identifying Potential Sources of Bias in a Descriptive Study 213
Interpreting Data in Descriptive Research 214
Some Final Suggestions 214
A Sample Dissertation 215
Application Exercises 218
For Further Reading 219

Chapter 7
Experimental, Quasi-Experimental, and Ex Post Facto Designs 220
The Importance of Control 222
Controlling for Confounding Variables 223
Types of Experimental, Quasi-Experimental, and Ex Post Facto Designs 227
Pre-Experimental Designs 228
Design 1: One-Shot Experimental Case Study 228
Design 2: One-Group Pretest–Posttest Design 229
Design 3: Static Group Comparison 229
True Experimental Designs 230
Design 4: Control-Group Pretest–Posttest Design 230
Design 5: Solomon Four-Group Design 230
Design 6: Control-Group Posttest-Only Design 231
Design 7: Within-Subjects Design 231
Quasi-Experimental Designs 233
Design 8: Nonrandomized Control-Group Pretest–Posttest Design 233
Design 9: Simple Time-Series Design 233
Design 10: Control-Group Time-Series Design 234
Design 11: Reversal Time-Series Design 234
Design 12: Alternating-Treatments Design 235
Design 13: Multiple-Baseline Design 235
Using Designs 11, 12, and 13 in Single-Case Intervention Research 237
Ex Post Facto Designs 239
Design 14: Simple Ex Post Facto Design 240
Factorial Designs 240
Design 15: Two-Factor Experimental Design 240
Design 16: Combined Experimental and Ex Post Facto Design 241
CONCEPTUAL ANALYSIS EXERCISE: Identifying Quantitative Research Designs 245
PRACTICAL APPLICATION: Determining Possible Cause-and-Effect Relationships 246
CHECKLIST: Looking for Confounding Variables 246
Meta-Analyses 248

Conducting Experiments on the Internet 248
Testing Your Hypotheses, and Beyond 249
PRACTICAL APPLICATION: Acknowledging the Probable Presence of Bias in Experimental Research 249
CHECKLIST: Identifying Potential Sources of Bias and Potential Threats to External Validity in an Experimental, Quasi-Experimental, or Ex Post Facto Study 250
A Sample Dissertation 251
Application Exercises 255
For Further Reading 255

Chapter 8
Qualitative Research Methods 258
Research Problems and Methodology Choice in Qualitative Research 259
Potential Advantages of a Qualitative Approach 260
Qualitative Research Designs 260
Case Study 260
Ethnography 261
Phenomenological Study 263
Grounded Theory Study 263
Narrative Inquiry 264
Content Analysis 265
CONCEPTUAL ANALYSIS EXERCISE: Choosing a Qualitative Research Design 267
Collecting Data in Qualitative Research 268
PRACTICAL APPLICATION: Ensuring That Qualitative Data Collection Yields Credible, Trustworthy Data 269
PRACTICAL APPLICATION: Selecting an Appropriate Sample for a Qualitative Study 271
PRACTICAL APPLICATION: Making Observations in a Qualitative Study 273
PRACTICAL APPLICATION: Planning and Conducting Interviews in a Qualitative Study 274
GUIDELINES: Conducting a Productive Interview 275
An Example in International Relations 279
Using Technology to Facilitate Collection of Interview Data 280
Criteria for Evaluating Qualitative Research 281
PRACTICAL APPLICATION: Planning the Logistics of a Qualitative Study 282
CHECKLIST: Pinning Down the Methodology of a Qualitative Study 282
A Sample Dissertation 283
Application Exercises 287
For Further Reading 288

Chapter 9
Mixed-Methods Designs 291
When Mixed-Methods Designs Are Most Useful and Appropriate 292

General Categories of Mixed-Methods Research 293
Convergent Designs 293
Exploratory Sequential Designs 295
Explanatory Sequential Designs 296
Longitudinal Mixed-Methods Designs 297
Multiphase Iterative Designs 298
CONCEPTUAL ANALYSIS EXERCISE: Identifying Mixed-Methods Research Designs 299
Planning a Mixed-Methods Study 301
Identifying Research Questions and Hypotheses 301
Conducting the Literature Review 302
Choosing One or More Appropriate Samples 302
Ensuring That the Study Will Have Reasonable Credibility 303
Analyzing and Interpreting the Data 304
Special Ethical Considerations in Mixed-Methods Research 304
PRACTICAL APPLICATION: Deciding Whether to Use a Mixed-Methods Design 305
CHECKLIST: Pinning Down the Logistics and Feasibility of a Mixed-Methods Study 305
A Sample Dissertation 306
Application Exercises 310
For Further Reading 311

Chapter 10
Action Research and Participatory Designs 314
Action-Oriented Models and Designs 316
Action Research/Teacher Research 316
Design-Based Research (DBR) 318
Participatory Action Research (PAR) and Youth Participatory Action Research (YPAR) 319
Community-Based Research (CBR) Models 321
Course-Based Action Research 322
Community-Based Participatory Research 323
Community-Based Qualitative Research 323
CONCEPTUAL ANALYSIS EXERCISE: Choosing an Action-Oriented Design 324
PRACTICAL APPLICATION: Deciding Whether to Use Action Research and/or a Participatory Design 325
CHECKLIST: Pinning Down the Logistics and Feasibility of Action Research, a Participatory Design, or a Combination of the Two 325
Data Collection and Analysis in Action Research 327
Disseminating the Findings of Action Research Projects 328
PRACTICAL APPLICATION: Using Community Forums as a Means of Disseminating the Results of Action Research and Participatory Research Projects 330
GUIDELINES: Planning and Conducting an Effective Community Forum 330
A Sample Research Report 331

Application Exercises 335
For Further Reading 336

PART IV Data Analyses

Chapter 11

Analyzing Quantitative Data 337

Exploring and Organizing a Data Set 337
Organizing Data to Make Them Easier to Think About and Interpret 339
Using Computer Spreadsheets to Organize and Analyze Data 342
Choosing Appropriate Statistics 343
Functions of Statistics 344
Statistics as Estimates of Population Parameters 344
Considering the Nature of the Data 345
Single-Group Versus Multi-Group Data 345
Continuous Versus Discrete Variables 345
Nominal, Ordinal, Interval, and Ratio Data 345
Normal and Non-Normal Distributions 346
Choosing Between Parametric and Nonparametric Statistics 348
Descriptive Statistics 349
Measures of Central Tendency 349
Curves Determine Means 351
Measures of Central Tendency as Predictors 352
Measures of Variability: Dispersion and Deviation 353
How Great Is the Spread? 354
Using the Mean and Standard Deviation to Calculate Standard Scores 355
Keeping Measures of Central Tendency and Variability in Perspective 357
Measures of Association: Correlation 358
How Validity and Reliability of Assessment Strategies Affect Correlation Coefficients 360
A Reminder About Correlations 361
Inferential Statistics 361
Estimating Population Parameters 361
An Example: Estimating a Population Mean 362
Point Versus Interval Estimates 363
Testing Hypotheses 364
Making Errors in Hypothesis Testing 365
Examples of Statistical Techniques for Testing Hypotheses 367
Meta-Analysis 369
Using Statistical Software Packages 370
Interpreting the Data 371
PRACTICAL APPLICATION: Analyzing and Interpreting Data in a Quantitative Study 372

CHECKLIST: Choosing Statistical Procedures 372
A Sample Dissertation 373
Application Exercises 376
For Further Reading 377

Chapter 12
Analyzing Qualitative and Mixed-Methods Data 380
Qualitative Analysis Strategies 381
General Strategies for Organizing and Analyzing Qualitative Data 381
Creswell's Data Analysis Spiral 387
An Example: Data Analysis in a Grounded Theory Study 387
An Example: Data Analysis in a Content Analysis Study 389
An Example: Data Analysis in an Ethnographic Study 390
PRACTICAL APPLICATION: Using Computer Databases to Facilitate Data Organization and Analysis 391
Acknowledging the Role of Researcher-as-Instrument in Qualitative Research 392
PRACTICAL APPLICATION: Planning Data Analysis for a Qualitative Study 393
CHECKLIST: Pinning Down the Data Analysis in a Qualitative Study 393
Analyzing and Interpreting Mixed-Methods Data 395
PRACTICAL APPLICATION: Using Computer Software Programs to Facilitate Mixed-Methods Data Analysis 398
Systematic Reviews of Qualitative and Mixed-Methods Studies 399
PRACTICAL APPLICATION: Observing How Experienced Researchers Have Conducted Qualitative and Mixed-Methods Research 400
CHECKLIST: Evaluating a Qualitative or Mixed-Methods Study 401
A Sample Dissertation 402
Application Exercises 407
For Further Reading 407

PART V Research Reports

Chapter 13
Planning and Preparing a Final Research Report 410
Getting Started 410
Surfing the Internet for Writing Assistance 411
Learn by Looking 411
Essential Elements of a Research Report 411
Explanation of the Research Problem or Question 412
Description of Study Design and Data-Collection Methods 412
Presentation of the Data and Their Analyses 412
Interpretations of the Data 414
Identification of Possible Weaknesses of the Study 415

Summary and Connections to a Broader Context 416
Maintaining Your Academic Integrity 416
Front Matter and End Matter 417
Preliminary Pages 417
Footnotes and Endnotes 418
Reference List 418
Appendix Content 421
Organizing a Research Report 421
Writing—and Finishing!—a Report 423
PRACTICAL APPLICATION: Writing Your Final Report 424
GUIDELINES: Writing a Clear, Coherent Report 424
PRACTICAL APPLICATION: Developing a Writing Schedule 426
GUIDELINES: Pinning Down and Adhering to a Workable Schedule 426
PRACTICAL APPLICATION: Critiquing a Final Research Report 427
CHECKLIST: Criteria for Critiquing a Research Report 428
Beyond the Unpublished Research Report: Presenting and Publishing 430
Conference Presentations 430
PRACTICAL APPLICATION: Presenting Your Research at a Professional Conference 430
GUIDELINES: Presenting an Effective Paper or Poster 431
Journal Articles 432
Sharing Authorship 433
Responding to Reviewers' Critiques 433
A Closing Thought 434
Application Exercises 434
For Further Reading 435

APPENDICES

Appendix A Using a Spreadsheet: Microsoft Excel 436

Using Excel to Keep Track of Literature Resources 436
Using Excel to Record and Recode Data 438
Reorganizing Data in Excel 441
Using Excel to Perform Simple Statistical Analyses 442

Appendix B Using SPSS 444

Creating a Data Set 444
Computing Basic Descriptive Statistics 446
Computing Inferential Statistics 447

Glossary 450

References 460

Index 467

Chapter 1 The Nature and Tools of Research

In virtually every subject area, our collective knowledge about the world is incomplete: Certain questions remain unanswered, and certain problems remain unsolved. Systematic research provides many powerful tools—not only physical tools but also mental and social tools—that can help us discover possible answers and identify possible solutions.

Learning Outcomes

1.1 Distinguish between (a) common uses of the term *research* that reflect misconceptions about what research involves and (b) the true nature of research in academic settings.

1.2 Describe the iterative, cyclical nature of research, including the steps that a genuine research project involves.

1.3 Distinguish among positivism, postpositivism, constructivism, and pragmatism/realism as philosophical underpinnings of a research project.

1.4 Identify examples of how six general research tools can play significant roles in a research project: (a) the library and its resources, (b) computer technology, (c) measurement, (d) statistics, (e) language, and (f) the human mind.

1.5 Describe steps you might take to explore research in your field.

In everyday speech, the word *research* is often used loosely to refer to a variety of activities. In some situations the word connotes simply 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 uses of the term can create considerable confusion for university students, who must learn to use it in a narrower, more precise sense.

Yet when used in its true sense—as a systematic process that leads to new knowledge and understandings—the word *research* can suggest a mystical activity that is somehow removed from everyday life. Many people imagine researchers to be aloof individuals who seclude themselves in laboratories, scholarly libraries, or the ivory towers of large universities. 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 general tools that make it possible.

WHAT RESEARCH IS NOT

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 how much 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 searching through various files or websites to discover what the realtor previously did not know. Rummaging—whether through records in one's own office, at a library, or on the Internet—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—such as collecting information, organizing it in a certain way for presentation to others, supporting statements with documentation, and referencing statements properly—these activities do not add up to true research. 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.

MyLab Education Self-Check 1.1

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.[1] 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*

[1]Some people in academia use the term *research* more broadly to include deriving new equations or abstract principles from existing equations 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.

FIGURE 1.1 The Research Cycle

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, research generally involves seven distinct steps, shown in Figure 1.1. We now look at each of these steps more closely.

1. ***The researcher begins with a problem—an unanswered question—related to a topic of interest and concern.*** The impetus for all good research is a desire to acquire new information that advances our collective understandings of physical, biological, social, or psychological phenomena. At a minimum, good researchers are *curious* researchers: They genuinely want to learn more about a particular topic. Many of them are also motivated to identify possible solutions to local, regional, or global problems—solutions that might either directly or indirectly enhance the well-being of humankind or of the physical, biological, and social environments in which we live.

As you think about your topic of interest, consider these questions: What is such-and-such a situation like? Why does such-and-such a phenomenon occur? Might such-and-such an intervention alter the current state of affairs? What does it all mean? With questions like these, research begins.

2. ***The researcher clearly and specifically articulates the goal of the research endeavor.*** A critical next step is to pin down the issue or question—which we will refer to as the **research problem**—that the researcher will address. The ultimate goal of the research must be set forth in a grammatically complete sentence that specifically and precisely identifies the question the researcher will try to answer. 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. ***The researcher often 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 possibly 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 many miles or kilometers 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?

Thus, what initially appears to be 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. ***The researcher identifies general assumptions—and possibly also specific hypotheses—that underlie the research effort.*** An **assumption** is a condition that is taken for granted, without which the research project would be pointless. For example, 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, you 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.

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 instance, a researcher might assume that

- Participants' responses in a paper-and-pencil questionnaire, face-to-face interview, or online survey are reasonably accurate indicators of their actual behaviors or opinions.
- Behaviors observed in an artificial laboratory environment can effectively reveal how people or other animal species are likely to behave in more natural, real-world settings.
- Certain assessment instruments (e.g., widely used intelligence tests, personality tests, and interest inventories) reflect relatively stable personal characteristics that are unlikely to change very much in the near future. (We examine this issue in detail in the discussion of *validity* of assessment instruments in Chapter 4.)

As you will discover in upcoming chapters, researchers can sometimes support such assumptions by citing past research findings or collecting certain kinds of data within their own research projects.

In addition to stating basic assumptions, many researchers form one or more hypotheses about what they might discover. A *hypothesis* is a logical supposition, a reasonable guess, an educated conjecture. In formal research, it might be more specifically called a **research hypothesis**, in that 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 you resolve the principal research problem. When one or more research hypotheses are proposed prior to any data collection, they are known as ***a priori* hypotheses**—a term whose Latin roots mean "from something before."

Hypotheses are certainly not unique to research. In your everyday life, if something happens, you immediately try to account for its cause by making some reasonable conjectures. For example, imagine that you come home after dark, open your 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 identify several hypotheses regarding the lamp's failure:

Hypothesis 1: A recent storm has disrupted your access to electrical power.
Hypothesis 2: The bulb has burned out.
Hypothesis 3: The lamp isn't securely plugged into the wall outlet.
Hypothesis 4: The wire from the lamp to the wall outlet is defective.
Hypothesis 5: You forgot to pay your electric bill.

Each of these hypotheses hints at a strategy for acquiring information that may resolve the nonfunctioning-lamp problem. For instance, to test Hypothesis 1, you might look outside to see whether your neighbors have lights, and to test Hypothesis 2, you might replace the current light bulb with a new one.

Hypotheses in a research project are as tentative as those for a nonfunctioning table 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.

Hypotheses—predictions—are an essential ingredient in certain kinds of research, especially experimental research (see Chapter 7). To a lesser degree, they might guide 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 studies in Chapter 8).

5. ***The researcher develops a specific plan for addressing the problem and its subproblems.*** Research is not a blind excursion into the unknown, with the hope that

the data necessary to address the research problem will magically 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 the formative stages of a research project, much can be decided: Are any existing data directly relevant to the research problem? If so, where are they, and are you likely to have access to them? If the needed data *don't* currently exist, how might you generate them? And later, after you have acquired the data you need, what will you do with 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. In Chapter 4, we discuss several general issues related to research planning. Then, beginning in Chapter 6, we describe strategies related to various research methodologies.

You should note here that 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.*

6. ***The researcher collects, organizes, and analyzes data related to the problem and its subproblems.*** After a researcher has isolated the problem, divided it into appropriate subproblems, identified assumptions (and possibly also *a priori* hypotheses), and chosen a suitable design and methodology, the next step is to collect whatever data might be relevant to the problem and organize and analyze those data in meaningful ways.

The data collected in research studies take one or both of two general forms. **Quantitative research** involves looking at amounts, or *quantities,* of one or more variables of interest. A quantitative researcher 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 isn't 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).

The two kinds of data—quantitative and qualitative—often require distinctly different research methods and data analysis strategies. Accordingly, three of the book's subsequent chapters focus predominantly on quantitative techniques (see Chapters 6, 7, and 11), and two others focus almost exclusively on qualitative techniques (see Chapters 8 and 12). Nevertheless, we urge you *not* to think of the quantitative–qualitative distinction as a mutually exclusive, *it-has-to-be-one-thing-or-the-other* dichotomy. Many researchers collect both quantitative and qualitative data in a single research project—an approach sometimes known as **mixed-methods research** (see Chapter 9). And in **action research**, one or more researchers—who are often practitioners in a particular helping profession (e.g., education, counseling, social work, medicine)—might use both quantitative and qualitative methods in an effort to improve current practices and desired outcomes (see Chapter 10). Good researchers tend to

be *eclectic* researchers who draw from diverse methodologies and data sources in order to best address their research problems and questions (e.g., see Gorard, 2010; Lather, 2006; Onwuegbuzie & Leech, 2005).

7. ***The researcher interprets the meaning of the data as they relate to the problem and its subproblems.*** Quantitative and qualitative data are, in and of themselves, *only* data—nothing more. The significance of the data depends on how the researcher extracts *meaning* from them. In research, uninterpreted data 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 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's the difference between the old perspective and the new one? It's 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 assumptions, hypotheses, and logical reasoning processes.

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 interpretations of the data.

We must emphasize two important points related to the seven-step process just described. First, *the process is* ***iterative:*** A researcher sometimes needs to move back and forth between two or more steps along the way. For example, while developing a specific plan for a project (Step 5), a researcher might realize that a genuine resolution of the research problem requires addressing a subproblem not previously identified (Step 3). And while interpreting the collected data (Step 7), a researcher may decide that additional data are needed to fully resolve the problem (Step 6).

Second, *the process is* ***cyclical***. The final step in the process depicted in Figure 1.1—interpretation of the data—is not *really* the final step at all. Only rarely is a research project a one-shot effort that completely resolves a problem; more often, it is likely to unearth new questions related to the issue at hand. And if specific hypotheses have been put forth, either *a priori* or after data have been collected and analyzed, those hypotheses are rarely proved or disproved beyond a shadow of a doubt. Instead, they 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 might turn to other hypotheses as being more likely explanations of the phenomenon in question. In any of these situations, 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 "*Voila!* I've completely answered the question about which I'm concerned." Instead, research typically

involves a cycle—or more accurately, a *helix* (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 undertaking that is static, self-contained, an end in itself. Here we see another difference between true research and the nonexamples of research presented earlier 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.

MyLab Education Self-Check 1.2

MyLab Education Application Exercise 1.1: Identifying Hypotheses and Assumptions

PHILOSOPHICAL ASSUMPTIONS UNDERLYING RESEARCH METHODOLOGIES

Let's return to Step 4 in the research process: *The researcher identifies assumptions—and possibly also hypotheses—that underlie the research effort.* The assumptions underlying a research project are sometimes so seemingly self-evident that a researcher may think it unnecessary to mention them. In fact, the researcher may not even be consciously aware of them. For example, two general assumptions underlie many research studies:

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

But are such assumptions justified? Is the world a lawful place, with some things definitely causing or influencing others? Or are definitive laws and cause-and-effect relationships nothing more than figments of our fertile human imaginations?

As we consider such questions, it is helpful to distinguish among different philosophical orientations that point researchers in somewhat different directions in their quests to make sense of our physical, biological, social, and psychological worlds.[2] Historically, a good deal of research in the natural sciences has been driven by a perspective known as **positivism**. Positivists believe that, with appropriate measurement tools, scientists can objectively uncover absolute, undeniable *truths* about cause-and-effect relationships within the physical world and human experience.

In the social sciences, many researchers are—and most others *should* be—less self-assured and more tentative about their assumptions. Some social scientists take a perspective known as **postpositivism**, believing that true objectivity in seeking absolute truths can be an elusive goal. Although researchers might strive for objectivity in their collection and interpretation of data, they inevitably bring certain *biases* to their investigations—perhaps biases regarding the best ways to measure certain variables or the most logical inferences to draw from patterns within the data. From a postpositivist perspective, progress toward genuine understandings of physical, social, and psychological phenomena tends to be gradual and probabilistic. For example, recall the earlier discussion of hypotheses being either *supported* or *not supported* by data. Postpositivists don't say, "I've just proven such-and-such." Rather, they're more likely to say, "This increases the probability that such-and-such is true."

[2]Some writers use terms such as *worldviews, epistemologies,* or *paradigms* instead of the term *philosophical orientations.*

Other researchers have abandoned any idea that absolute truths are somewhere "out there" in the world, waiting to be discovered. In this perspective, known as **constructivism**, the "realities" researchers identify are nothing more than human *creations* that can be helpful in finding subjective meanings within the data collected.[3] Constructivists not only acknowledge that they bring certain biases to their research endeavors but also try to be as up-front as possible about these biases. The emphasis on subjectivity and bias—rather than objectivity—applies to the phenomena that constructivist researchers study as well. By and large, constructivists focus their inquiries on people's *perceptions* and *interpretations* of various phenomena, including individuals' behaviors, group processes, and cultural practices.

In yet another perspective, known as **phenomenology**, the focus is entirely on how human beings *experience* themselves and their world as they go through life. Researchers with this orientation typically ask the question, "What is it like to experience such-and-such?" For example, they might ask, "What is it like . . . to have attention-deficit disorder?" ". . . to run for political office?" ". . . to undergo chemotherapy?" ". . . to immigrate to an English-speaking country without any knowledge of English?" In our view, a phenomenological orientation is also a constructivist orientation, in that people's constructed realities are essential components of their lived experiences. However, some scholars argue that the two perspectives are distinctly different entities.

Many of the quantitative methodologies described in this book have postpositivist, probabilistic underpinnings—a fact that becomes especially evident in the discussion of statistics in Chapter 11. In contrast, some qualitative methodologies have a distinctly constructivist or phenomenological bent, with a focus on ascertaining people's beliefs, perceptions, and experiences, rather than trying to pin down absolute, objective truths that might not exist at all.

But once again we urge you *not* to think of quantitative research and qualitative research as reflecting a mutually exclusive, *either-this-or-that* dichotomy. For instance, some quantitative researchers approach a research problem from a constructivist framework, and some qualitative researchers tend to think in a postpositivist manner. Many researchers acknowledge *both* that (a) absolute truths regarding various phenomena may actually exist—even if they are exceedingly difficult to discover—and (b) human beings' self-constructed beliefs and experiences are legitimate objects of study in their own right. You might see such labels as **pragmatism** and **realism** used in reference to this orientation (e.g., see R. B. Johnson & Onwuegbuzie, 2004; Maxwell & Mittapalli, 2010).

Presumably, most researchers hope that their findings will either directly or indirectly be useful to humankind as a whole. But some researchers focus almost exclusively on how human actions can lead to beneficial outcomes. In an orientation that goes by various names, including *action research, praxis,* and *social action,* a researcher places a particular human action or intervention front and center in an investigation—perhaps studying a particular approach to physical therapy, reading instruction, protection of an endangered species, or rainforest preservation—with the ultimate goal being to enhance the well-being of our planet or some of its inhabitants (e.g., see Gergen, Josselson, & Freeman, 2015). The researcher determines whether the intervention has a desired effect and then, after some analysis and reflection, may modify the intervention to further enhance its effectiveness. Clearly, these practically oriented researchers assume that some cause-and-effect relationships do exist in our world and that, more specifically, we human beings can have a beneficial impact on our physical, biological, social, or psychological environments.

MyLab Education Self-Check 1.3

MyLab Education Application Exercise 1.2: Examining Philosophical Assumptions

[3]In some fields (e.g., in business), this perspective is often called *interpretivism.*

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 that researchers use to achieve their research goals can vary considerably depending on the discipline. A microbiologist needs a microscope and culture media; an attorney needs a library of legal decisions and statute law. By and large, we don't 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 certain information that will be analyzed and interpreted at some point in the research process. Likewise, statistics merely provide ways to analyze and summarize data, thereby allowing us to see patterns within the data more clearly.

In the following sections, we look more closely at six general tools of research:

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

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 at Ephesus housed more than 12,000 papyrus scrolls and, in later years, many parchment books as well.[4]

Until the past few decades, libraries were primarily repositories of concrete, physical representations of knowledge—clay tablets, scrolls, manuscripts, books, journals, films, and the like. 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, CDs, DVDs, online databases) to store information in more compact forms. Second, they have provided increasingly fast and efficient means of locating and

[4] Many academic scholars would instead say "seventh century BCE" and "second century CE" in this sentence, referring to the more religion-neutral terms *Before Common Era* and *Common Era.* However, we suspect that some of our readers are unfamiliar with these terms, hence our use of the more traditional ones.

accessing information on virtually any topic. And third, many of them have made catalogs of their holdings available on the Internet. 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 research tools, personal computers—whether they take the form of desktops, laptops, tablets, or smartphones—are now commonplace. In addition, computer software packages and applications have become increasingly user-friendly, such that novice researchers can easily take advantage of them. But like any tool—no matter how powerful—computer technology has its limitations. Yes, computers can certainly calculate, compare, search, retrieve, sort, and organize data more efficiently and 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 can't 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. Table 1.1 offers suggestions for how you might use computer technology as a research tool.

Measurement

Especially when conducting quantitative research, a researcher needs a systematic way of *measuring* the phenomena under investigation. Some common, everyday measurement instruments—rulers, scales, stopwatches—can occasionally be helpful for measuring easily observable variables, such as length, weight, or time. But in most cases, a researcher needs one or 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.

In quantitative research, social and psychological phenomena require measurement as well, even though they have no concrete, easily observable basis in the physical world. For example, an economist might use the Dow-Jones Industrial Average 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 schoolchildren have acquired knowledge and skills related to a particular topic. Finding or developing appropriate measurement instruments for social and psychological phenomena can sometimes be quite a challenge. Thus, we explore measurement strategies in some depth when we discuss the research planning process in Chapter 4.

Statistics

As you might guess, statistics are most helpful in quantitative research, although they occasionally come in handy in qualitative research as well. Statistics also tend to be 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.

Statistics have two principal functions: to help a researcher (a) describe quantitative data and (b) draw inferences from these data. **Descriptive statistics** help the researcher capture the general nature of the data obtained—for instance, how certain measured characteristics appear to be "on average," how much variability exists within a data set, and how closely two

TABLE 1.1 The Computer As a Research Tool

Part of the Study	*Relevant Technological Support Tools*
Planning the study	• Brainstorming assistance—software used to help generate and organize ideas related to the research problem, research strategies, or both. • Outlining assistance—software used to help structure various aspects of the study and focus work efforts. • Project management assistance—software used to schedule and coordinate varied tasks that must occur in a timely manner. • Budget assistance—spreadsheet software used to help in outlining, estimating, and monitoring the potential costs involved in the research effort.
Literature review	• Literature identification assistance—online databases used to help identify relevant research studies to be considered during the formative stages of the research endeavor. • Communication assistance—computer technology used to communicate with other researchers who are pursuing similar topics (e.g., e-mail, Skype, electronic bulletin boards, list servers). • Writing assistance—software used to facilitate the writing, editing, formatting, and citation management of the literature review.
Study implementation and data gathering	• Materials production assistance—software used to develop instructional materials, visual displays, simulations, or other stimuli to be used in experimental interventions. • Experimental control assistance—software used to physically control the effects of specific variables and to minimize the influence of potentially confounding variables. • Survey distribution assistance—databases and word-processing software used in combination to send specific communications to a targeted population. • Online data collection assistance—websites used to conduct surveys and certain other types of studies on the Internet. • Field based data collection assistance—software used to take field notes or to monitor specific types of responses given by participants in a study.
Analysis and interpretation	• Organization and transcription assistance—software used to assemble, categorize, code, integrate, and search potentially huge data sets (such as qualitative interview data or open-ended responses to survey questions). • 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. • Dissemination assistance—desktop publishing software and poster-creation software used to produce professional-looking documents and posters that can be displayed or distributed at conferences and elsewhere. • Presentation graphics assistance—presentation software used to create static and animated slides for conference presentations. • Networking assistance—blogs, social networking sites, and other Internet-based mechanisms used to communicate one's findings to a wider audience and to generate discussion for follow-up studies by others in the field.

or more characteristics are associated with one another. In contrast, **inferential statistics** help the researcher make decisions about the data. For example, they might help a researcher decide whether the differences observed between two experimental groups are large enough to be attributed to the differing experimental interventions rather than to a once-in-a-blue-moon fluke. Both of these functions of statistics ultimately involve summarizing the data in some way.

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.[5] By doing so, we find that, "on average," the students work 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 these students has worked exactly 18.5 hours a week. That figure represents absolutely no fact in the real world.

If statistics offer 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 a very limited amount of information at any single point in time.[6] 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 a 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.*

Any researcher who uses statistics must remember that calculating statistical values is not—and must not be—the final step in a research endeavor. The ultimate question in research is, *What do the data indicate?* Statistics yield *information* about data, but conscientious researchers are not satisfied until they determine the *meaning* of this information.

Although a book such as this one can't 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're driving along a country road. In a field to your left, you see an object with the following characteristics:

- Black and white in color, in a splotchy pattern
- Covered with a short, bristly substance
- Appended at one end by something similar in appearance to a paintbrush
- Appended at the other end by a lumpy thing with four smaller things coming out of its top (two soft and floppy; two hard, curved, and pointed)
- 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, 2016; also see Jaccard & Jacoby, 2010):

1. ***Words reduce the world's complexity.*** Classifying similar objects and events into categories and assigning specific words to those categories can make our experiences easier to make sense of. For instance, it's 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

[5]When the word *arithmetic* is used as an adjective, as it is here, it is pronounced with emphasis on the third syllable ("ar-ith-MET-ic").

[6]If you have some background in human memory and cognition, you may realize that we are talking about the limited capacity of *working memory* here (e.g., see Cowan, 2010; J. E. Ormrod, 2016).

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 four spindly sticks on its underside is a concrete entity. The concept *cow,* however, is more abstract: It connotes such characteristics as *female, supplier of milk,* and, to the farmer, *economic asset.* Concepts and the labels associated with them allow us to think about our experiences without necessarily having to consider all of their discrete, concrete characteristics.

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 (e.g., *farmer, milk, pasteurization*) 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 two 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 *learning outcome, classroom climate,* and *student at risk* 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 theories and research, and communicating your results to others.

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

The Benefits of Knowing Two or More Languages

It should go without saying that not all important research is reported in a researcher's native tongue. Accordingly, some 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.

Knowing two or more languages has a second benefit as well: Words in a second language may capture the *meaning* of certain phenomenon in ways that one's native tongue may not. For example, the German word *Gestalt*—which roughly means "organized whole"—has no direct equivalent in English. Thus, many English-speaking psychologists use this word when describing the nature of human perception, enabling them to communicate the idea that people often perceive organized patterns and structures in visual data that, in the objective physical world, are *not* organized. Likewise, the Zulu word *ubuntu* defies an easy

translation into English. This word—which reflects the belief that people become fully human largely through regularly caring for others and contributing to the common good—can help anthropologists and other social scientists capture a cultural worldview quite different from the more self-centered perspective so prevalent in mainstream Western culture.

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 a good deal of conventional wisdom tells us that clear thinking *precedes* clear writing, writing can in fact 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 things you do and don't 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 isn't surprising, then, that writing about a topic actually enhances the writer's understanding of the topic (e.g., Kellogg, 1994; Mueller & Oppenheimer, 2014; Shanahan, 2004).

If you wait until all your thoughts are clear before you start writing, you may never begin. Thus we recommend that you start writing parts of your research proposal or report as soon as possible. Begin with a title and a **purpose statement**—one or more sentences that clearly describe the primary goal(s) you hope to achieve by conducting 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 to . . . ," you are well on your way to planning a focused research study.

PRACTICAL APPLICATION Communicating Effectively Through Writing

In our own experiences, we authors have found that most students have a great deal to learn about what good writing entails. Yet we also know that with effort, practice, mentoring, and regular feedback, students *can* learn to write more effectively. Subsequent chapters present specific strategies for writing literature reviews (Chapter 3), research proposals (Chapter 5), and research reports (Chapter 13). Here we offer general strategies for writing in ways that can help you clearly communicate your ideas and reasoning to others. We also offer suggestions for making the best use of 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 people have written (e.g., see J. E. Ormrod, 2016).

1. ***Be specific and precise.*** 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 online resources offer suggestions for writing clear, concise sentences and combining them 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 research reports. But ultimately, everything you say should relate either directly or indirectly to your research problem. If you can't 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, Dinah Jackson, a doctoral student in educational psychology, was interested in the possible effects of *self-questioning*—asking oneself questions about the 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.*** We authors have read many student research reports that seem to wander aimlessly and unpredictably from one thought to another, without any obvious organizational structure directing the flow of ideas. Using headings and subheadings is one simple way to provide an organizational structure for your writing *and* make that structure crystal clear to others.

5. ***Use concrete examples to make abstract ideas more understandable.*** There's 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 made 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 argued that only a superficial coding scheme (e.g., counting the number of main ideas included in notes) would fail to capture the true quality of the notes. She clarified 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 many 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 help 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.*** You will probably 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 finished their task. Such students have invariably been disappointed—sometimes even outraged—when we have asked them to revise their work, usually several times. The need 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 "12th edition" indicates! Multiple revisions enable you to reflect on and critically evaluate your own writing, revise and refocus awkward passages, get feedback from peers and advisors who can point out where a manuscript has gaps or lacks clarity, and in other ways ensure that the final version 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 disciplined scholar whose thoughts and work are worth reading about. If, instead, you mispel menny of yur werds—as we our doing in this sentance—your reeders may quikly discredit you as a sloppy resercher who shuldn't be taken seriusly!

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. In addition, dictionaries and word-processing spell-check functions can obviously assist you in your spelling.

GUIDELINES Using the Tools in Word-Processing Software

Most of our readers know the basics of using word-processing software—for instance, how to "copy," "paste," and "save"; how to choose a particular font and font size; and how to format text as *italicized,* <u>underlined</u>, or **boldface**. Following are specific features and tools that you may not have routinely used in previous writing projects but that can be quite useful in writing research reports:

- ***Outlining.*** An "outlining" feature lets you create bullets and subbullets to organize your thoughts.
- ***Setting headers and footers.*** A "header" is a line or two at the top of the page that appears on every page; a "footer" appears at the bottom of each page. For example,

using the "insert date" function, you might create a header that includes the specific date on which you are writing a particular draft. And using an "insert page number" function will add appropriate numbers to the tops or bottoms of successive pages.

- ***Creating tables.*** Using a "table" feature, you can create a table with the number of rows and columns you need. You can easily adjust the widths of various columns, format the text within each table cell, add new rows or tables, and merge two or more cells into a single, larger cell. Usually, an "autoformat" option will give you many possible table formats from which to choose.
- ***Inserting graphics.*** You are likely to find a variety of options under an "Insert" pull-down menu. Some of these options enable you to insert diagrams, photographs, charts, and other visuals you have created elsewhere.
- ***Creating footnotes.*** Footnotes are easy to create using an "insert footnote" feature. Typically, you can choose the symbols to be used in designating footnotes—perhaps *1, 2, 3, . . .* , *a, b, c, . . .* , or special symbols such as * and †.
- ***Using international alphabets and characters.*** Computers and computer software sold in English-speaking countries have the English alphabet as the default alphabet, but often either your word-processing software or your "system preferences" on your computer's operating system will let you choose a different alphabet (e.g., Turkish, as in the surname Kağitçibaşi) or certain characters (e.g., in Chinese or Japanese) for particular words or sections of text.
- ***Tracking changes.*** A "track changes" feature enables you to keep a running record of specific edits you have made to a document; you can later go back and either "accept" or "reject" each change. This feature is especially useful when two or more researchers are coauthoring a report: It keeps track of who made which changes and the date on which each change was made.

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

1. ***Save and back up your document frequently.*** We authors can recall a number of 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, things you have written sometimes get lost. It's imperative that you get in the habit of regularly 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 your work-in-progress frequently, perhaps every 5 to 10 minutes. Many software programs will do this for you automatically if you give them instructions about whether and how often to do it.
- 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 relative's home, or somewhere in cyberspace. One option is to save documents on a flash drive or external hard drive. Another is to copy them to an electronic dropbox, iCloud (for Macintosh), or other Internet-based storage mechanism. One of us authors uses a flash drive to back up much of her past work (including several book manuscripts) and any in-progress work; she keeps this flash drive in her purse and takes it everywhere she goes. Also, she occasionally sends herself in-progress documents as attachments to self-addressed e-mail messages—giving her an almost-current backup version of the documents in the event that an unintended keystroke somehow wreaks havoc on what she has written.
- Save various versions of your work with distinct labels 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. And service departments at computer retailers can often retrieve documents from the hard drives of otherwise "dead" machines.

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., surnames such 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 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 in red ink, 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 also comes into play. The human mind is undoubtedly the most important tool in the researcher's toolbox. Nothing equals its powers of comprehension, integrative reasoning, and insight.

Over the past few millennia, human beings have developed a number of 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, scientific methods, theory building, and collaboration with other minds.

Critical Thinking

Before beginning a research project, good researchers typically look at research reports and theoretical discussions 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 methodologies, and unwarranted conclusions. And, of course, good 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.

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; Mercier, Boudry, Paglieri, & Trouche, 2017; 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 cultural group 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 the 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—we conclude that tulips are apt 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; 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 finite and small (premise), 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 two faulty premises: that the Earth is flat and also relatively small.

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 more 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 of questions students ask themselves (mentally) and try to answer as they listen to a lecture or read a textbook influence both 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?") From such findings, Jackson generated several key premises and drew a logical conclusion from them:

> If learning information in an associative, integrative manner is more effective than learning information in a fact-by-fact, piecemeal manner, (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 better-integrated class notes and higher-quality learning. (Conclusion)

Such reasoning led Jackson to form and test several hypotheses, including this one:

> 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)

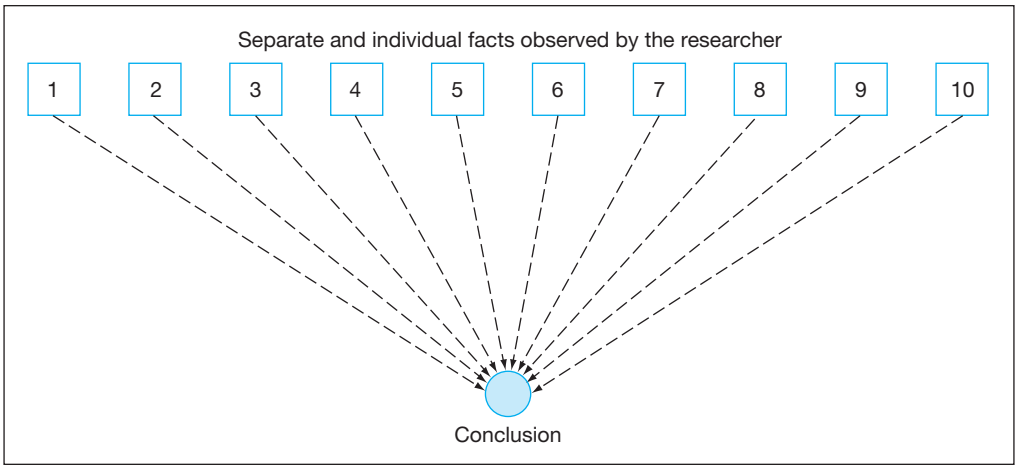
Happily for Jackson, the data she collected in her dissertation research supported this hypothesis.

Inductive Reasoning

Inductive reasoning begins not with a preestablished truth or assumption but instead with an observation. For example, 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. "Hmmm," you may have thought, "what happens if I do that again?" So you grabbed another cracker, held it out, 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 invariably observed the same result. Eventually, you 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 larger *population* from which the sample has been

FIGURE 1.2 The Inductive Process



taken. For instance, an anthropologist might draw conclusions about a certain culture after studying a certain 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 dyslexia to recommend that teachers use the instructional approach with other students with dyslexia. A sociologist might (a) conduct three surveys (one each in 1995, 2005, and 2015) asking 1,000 people to describe their beliefs about AIDS and then (b) draw conclusions about how society's attitudes toward AIDS have changed over the 20-year period.

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 electroencephalogram (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: *People who exhibit flat EEGs for 24 hours or longer will not recover.* We cannot, of course, rule out the unexplored cases, but *from the data observed,* the conclusion reached was that recovery is impossible. The EEG line from *every* case led to that *one* conclusion.

Scientific Method

During the Renaissance, people found that when they systematically collected and analyzed data, new insights and understandings might emerge. Thus was the scientific method born; the words literally mean "the method that searches for 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 term **scientific method** has referred to an approach in which a researcher (a) identifies a problem that defines the goal of one's quest; (b) posits a hypothesis that, if confirmed, resolves the problem; (c) gathers data relevant to the hypothesis; and (d) analyzes and interprets the data to see whether they support the hypothesis and resolve the question that instigated the research. In recent years, however, the term has been a controversial one because not all researchers follow the steps just listed in a rigid, lockstep manner; in fact, as noted earlier, some researchers shy away from forming any hypotheses about what they might find. Some of the controversy revolves around which article to use in

front of the term—more specifically, whether to say "*the* scientific method" or "*a* scientific method." If we are speaking generally about the importance of collecting and analyzing data systematically rather than haphazardly, then saying "*the* scientific method" makes sense. If, instead, we are speaking about a specific methodology—say, experimental research or ethnographic research (described in Chapter 7 and Chapter 8, respectively), it is probably better to say "*a* scientific method." In any event, we are talking about a somewhat flexible—although certainly also rigorous—process.

As you may already have realized, application of a scientific method usually involves both deductive logic and inductive reasoning. Researchers might develop a hypothesis either from a theory (deductive logic) or from observations of specific events (inductive reasoning). Using deductive logic, they might make predictions about the patterns they are likely to see in their data *if* a hypothesis is true, and even researchers who have not formulated hypotheses in advance must eventually draw logical conclusions from the data they obtain. Finally, researchers often use inductive reasoning to generalize about a large population from which they have drawn a small sample.

Theory Building

Psychologists are increasingly realizing that the human mind is a very *constructive* mind. People don't just passively absorb and remember a large body of unorganized facts about the world. Instead, they pull together the things they see and hear 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, 2016).

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—for instance, why the sun "goes down" at night, where babies come from, and why certain individuals behave in particular ways. People's everyday, informal theories about the world aren't always accurate. For example, imagine that an airplane drops a large metal ball as it travels forward through the air. 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 (Cook & Breedin, 1994; McCloskey, 1983).

What characterizes the theory building of a good researcher is the fact 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 being examined.* 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, 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, and relating the theory to other theories regarding overlapping phenomena (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 in the theory, and conducting additional investigations to test one or more aspects of the theory in its current state. 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. Three or more heads can be even better. Any single researcher is apt to have certain perspectives, assumptions, and theoretical biases—not to mention gaps in knowledge about the subject matter—that will limit how the researcher approaches a research project. By bringing one or more professional colleagues into a project—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.

Sometimes these colleagues enter the picture as equal partners. At other times 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—more importantly—will make a genuine contribution to the student's topic of study.

Many productive researchers keep in regular communication with others who conduct research on the same or similar topics, perhaps exchanging ideas, critiquing one another's work, and directing one another to potentially helpful resources. Such ongoing communication is also a form of collaboration—albeit a less systematic one—in that everyone can benefit from and build on what other people are thinking and finding. Increasingly, computer technology is playing a central role in this cross-communication and cross-fertilization. For example, many researchers subscribe to topic-specific **electronic discussion groups**—you may also see such terms as *list servers, online discussion forums, bulletin boards,* and *message boards*—in which any message sent to or posted on them is available and possibly sent to all subscribers. In addition, some researchers maintain professional websites 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.

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 where we don't look objectively at the data we collect. And sometimes we are so emotionally attached to particular perspectives or theories about a phenomenon that we can't abandon them when mountains of evidence indicate that we should. Figure 1.3 describes

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 failing to separate 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 $a = F \div m$ and therefore that if there is no acceleration ($a = 0$), then there is no force ($F = 0$). 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 Chicago—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 and want to make *generalizations* about a larger group.
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 has probably burned out." We unscrew the existing light bulb and replace it with a new one—and *voila!* 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 secretly hope that they don't find such evidence, of course, but they open-mindedly look for it nonetheless.
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 isn't 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're apt to think quite rationally and objectively when dealing with topics we don't 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 doesn't necessarily cause the other. An ability to distinguish between causation and correlation is critical in any research effort, as you will discover in Chapter 6.

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.

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.

MyLab Education Self-Check 1.4

MyLab Education Application Exercise 1.3: Communicating Effectively about Research

REFLECTIONS ON NOTEWORTHY 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.

Despite the lukewarm reception, Fleming 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 presentation of his 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 especially 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 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.

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

If you want to learn 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 book will give you an insight into the way great research comes into being.

The procedures used in groundbreaking research are identical to those every student follows in completing a dissertation, thesis, or other research project. Triggered by curiosity, all research begins with an observation, a question, a problem. Assumptions are made. Hypotheses might be formulated. Data are gathered. Conclusions are reached. What *you* do in a research project is the same as what many others have done before you, including 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 tasks to be meaningless busywork that have 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. 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 legitimate ways to collect and analyze data is essential for keeping up with advances in your field. The alternative—*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 academic discipline. You can do so by spending an hour or two in the *periodicals* section of your local college or university library or, alternatively, making use of your library website's online databases to download and read a number of articles at home.

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

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 gained the approval of designated peers. A **peer-reviewed** research report—you may also see the terms *juried* and *refereed*—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 **non-peer-reviewed** report (a.k.a., a *nonjuried* or *nonrefereed* one) is a report that appears in a journal or on the Internet without first being screened by one or more experts. Some non-peer-reviewed reports are excellent, but others may not be.

PRACTICAL APPLICATION Identifying Important Tools in Your Discipline

We have introduced several key research tools in the preceding pages, and we describe many more specific ones in subsequent chapters. 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 important tools in your discipline is to work closely with an expert researcher in your field.

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 prestigious 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 can't locate anyone locally, it may be possible to recruit one or more willing individuals 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?

MyLab Education Self-Check 1.5

APPLICATION EXERCISES

Similar exercises are available on Pearson MyLab Education.

1. A local school board is developing a new preschool program; the system already has half-day and full-day kindergartens. The school board members recognize that children come to kindergarten from many different backgrounds and wonder whether preschoolers who attend childcare programs that are run by school districts are better prepared for kindergarten than children who attend independent childcare programs.

   i. You are hired as a research consultant to help the board members with planning and especially with advertising the new program. What's the research problem in this situation?

   ii. Your first task as the research consultant is to identify the hypotheses and assumptions that underlie the research effort. In this case, you think the board members already have both. What's their most likely hypothesis? Is it your hypothesis, too, as the consultant?

   iii. What are some of the assumptions that underlie the effort?

   iv. You search the Internet and the library for information about school-based and independent preschools. There's a lot out there: studies, summaries, reviews, opinions. You write a lengthy report summarizing your findings and present it to the school board. Did you engage in research?

2. Answer these questions on examining philosophical assumptions.

   i. You are being invited into a school to research the school's reading instruction in the early grades. You have not been given many details about what sort of study the school is interested in. What questions would you ask school administrators and teachers in the initial meeting? How would you determine which approach to take?

   ii. During the initial meeting, staff and teachers seem to pose causal sorts of questions—such as "What is the effect of . . .?" or "What is the impact of . . .?" What

sorts of philosophical assumptions are underlying those questions? How would this inform your design of a study at the school? What challenges would you face in implementing this sort of study in the school setting? What would be important to identify at the outset of the study?

iii. The reading specialists at the school seem to be more interested in teachers' perspectives in implementing reading instruction and how students are responding to the program, rather than identifying causal relationships and outcomes. What philosophical assumptions might they be operating from? What sorts of research questions could be generated to study reading from teachers' perspectives? What might data collection look like?

FOR FURTHER READING

General Research Design

Bouma, G., & Carland, S. (2016). *The research process* (6th ed.). New York, NY: Oxford University Press.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Los Angeles, CA: Sage.

Goodwin, C. J. (2013). *Research in psychology: Methods and design* (7th ed.). New York, NY: Wiley.

Howell, K. E. (2013). *An introduction to the philosophy of methodology.* Los Angeles, 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.

Letherby, G., Scott, J., & Williams, M. (2012). *Objectivity and subjectivity in social research.* Los Angeles, CA: Sage.

McMillan, J. H., & Wergin, J. F. (2010). *Understanding and evaluating educational research* (4th ed.). Hoboken, NJ: Pearson.

Nieswiadomy, R. M. (2012). *Foundations in nursing research* (6th ed.). Hoboken, NJ: Pearson.

Niglas, K. (2010). The multidimensional model of research methodology: An integrated set of criteria. In A. Tashakkori & C. Teddlie (Eds.), *Mixed methods in social & behavioral research* (2nd ed., pp. 215–236). Los Angeles, CA: Sage.

Repko, A. F., & Szostak, R. (2017). *Interdisciplinary research: Process and theory* (3rd ed.). Los Angeles, CA: Sage.

Rosnow, R. L., & Rosenthal, R. (2013). *Beginning behavioral research: A conceptual primer* (7th ed.). Hoboken, NJ: Pearson.

Effective Writing

Beebe, L. (Ed.). (1993). *Professional writing for the human services.* Washington, DC: National Association of Social Workers Press.

Flesch, R. (1974). *The art of readable writing.* New York, NY: Harper & Row.

Glicken, M. D. (2007). *A guide to writing for human service professionals.* New York, NY: Rowman and Littlefield.

Mitchell, M. L., Jolley, J. M., & O'Shea, R. P. (2013). *Writing for psychology* (4th ed.). Belmont, CA: Wadsworth/Cengage.

Pinker, S. (2014). *The sense of style: The thinking person's guide to writing in the 21st century.* New York, NY: Viking.

Strunk, W., Jr., & White, E. B. (2009). *The elements of style* (50th anniversary ed.). New York, NY: Pearson/Longman.

Weisman, D., & Zornado, J. (2013). *Professional writing for social work practice.* New York, NY: Springer.

Williams, J. M., & Bizup, J. (2014). *Style: Lessons in clarity and grace* (11th ed.). Hoboken, NJ: Pearson.

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.

Bickle, J., Mauldin, R., & Giere, R. N. (2005). *Understanding scientific reasoning* (5th ed.). New York, NY: 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, NY: Guilford Press.

Jung, S. (1995). *The logic of discovery: An interrogative approach to scientific inquiry.* New York, NY: Peter Lang.

Poplin, M. S. (1987). Self-imposed blindness: The scientific method in education. *Remedial and Special Education, 8*(6), 31–37.

Popper, K. (2000). *The logic of scientific discovery.* New York, NY: Routledge.

Ravitch, S. M., & Riggan, M. (2016). *Reason and rigor: How conceptual frameworks guide research.* Los Angeles, CA: Sage.

Shank, G. D. (2006). *Qualitative research: A personal skills approach* (2nd ed.). Hoboken, NJ: Pearson. [See Chapter 7.]

Chapter 2 The Problem: The Heart of the Research Process

The main research problem or question is the axis around which the whole research effort revolves. It clarifies the goals of the research project and can keep the researcher from wandering in tangential, unproductive directions.

Learning Outcomes

2.1. Identify strategies for choosing and refining an overall research problem or question.
2.2. Subdivide a main research problem or question into useful subproblems (or subquestions).
2.3. Pin down a proposed research study by (a) identifying a relevant theoretical or conceptual framework; (b) if appropriate, stating one or more *a priori* hypotheses; (c) identifying the general concepts or more-specific variables to be examined; (d) defining terms; (e) stating assumptions; (f) if applicable, identifying delimitations and limitations; and (g) explaining the study's potential importance or significance.

The heart of every research project—the axis around which the entire research endeavor revolves—is the problem or question the researcher wants to address. The first step in the research process, then, is to identify this problem or question with clarity and precision.

FINDING RESEARCH PROJECTS

Issues in need of research are everywhere. Some research projects can enhance our general knowledge about our physical, biological, social, or psychological world or shed light on historical, cultural, or aesthetic phenomena. For example, an ornithologist might study the mating habits of a particular bird species, an anthropologist might examine the moral beliefs and behaviors of a particular cultural group, and a psychologist might study the nature of people's logical reasoning processes (as one of us authors did in her doctoral dissertation). Such projects, which can advance theoretical conceptualizations about a particular topic, are known as **basic research**.

Other research projects address issues that have immediate relevance to current procedures, practices, and policies. For example, an agronomist might study the effects of various fertilizers on the growth of sunflowers, a nursing educator might compare the effectiveness of different instructional techniques for training future nurses, and a political scientist might determine what kinds of campaign strategies influence one or more demographic groups' voting patterns. Such projects, which can inform human decision making about practical problems, are known as **applied research**.

Keep in mind, however, that the line between basic research and applied research is 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 an

endangered bird species, an ornithologist might lead the way in saving the 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.

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 identified as being in **pdf** form, referring to **portable document format**; these items are often electronic photocopies of articles that have appeared in academic journals and similar sources.

You might also want to look at typical research projects for doctoral dissertations. For example, your university library probably has a section that houses the completed dissertations of students who have gone before you. Alternatively, you might go to the electronic databases in your library's catalog. Among those databases you are likely to find ProQuest Dissertations & Theses Global, which includes abstracts—and in many cases, the complete texts—for millions of dissertations and theses from around the world.

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 aren't suitable for research because they lack the interpretation-of-data component; they don't 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. ***A research project should not be simply a ruse for achieving self-enlightenment.*** All of us have large gaps in our knowledge that we may want to fill. But mere self-enlightenment should not be the primary purpose of a research project (see Chapter 1). Gathering information to know more about a certain topic is entirely different from looking at a body of data to discern how it contributes to the solution of a problem or to address a currently unanswered question.

A student once submitted the following 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 project whose sole purpose is to compare two sets of data does not qualify as a suitable research endeavor.*** 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:

	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 a government website to reveal what is already known.

3. ***Simply calculating a correlation coefficient between two related sets of data is not acceptable as a problem for research.*** Why? Because a key ingredient in true research—*making sense* of the data—is missing. 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 after they have collected data and performed a simple statistical procedure, their work is done. 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 is of little use. It does, however, suggest questions that a research project might address: What is the underlying *cause* of the correlation between children's and parents' intelligence test scores? Is it genetic? Is it environmental? Does it reflect some combination of genetic heritage and environment?

4. ***A problem that results only in a yes-or-no answer is not a sufficient problem for a research study.*** Why? For the same reason that merely calculating 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's 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, might be 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 issues and not dwell on those that will make little or no contribution. 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 is "How do heredity and environment *interact* in their influences on behavior?"

PRACTICAL APPLICATION Identifying and Presenting the Research Problem or Question

How can a beginning researcher formulate an important and useful research problem or question? Here we offer guidelines both for choosing an appropriate problem or question and for pinning it down sufficiently to focus the research effort.

GUIDELINES Choosing an Appropriate Problem or Question

Choosing a good research problem or question requires genuine curiosity about unanswered questions. But it also requires enough knowledge about a topic to identify the kinds of investigations that are likely to make important contributions to one's field. Here we offer several strategies that are often helpful for novice researchers.

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. 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 2 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, mistakenly attributing them to natural "sloshing" as a result of the Earth's movement through space, rather than to the moon's gravitational pull.

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, hear, and read: Why does such-and-such happen? What makes such-and-such tick? What are people thinking when they do such-and-such?

2. ***Read the existing research literature about a topic.*** One essential strategy is to find out what things are already known and believed 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 about a topic 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 personally know or believe to be true (Neuman, 2011)

Reading relevant literature has other advantages as well. It can give you a conceptual or theoretical framework on which you can build a rationale for your study; we talk more about such frameworks later in the chapter. Reading the literature can also offer potential research designs and data-collection strategies. And it can help you interpret your results and relate them to previous research findings in your 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 also urge you 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 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 record of it.

3. ***Seek the advice of experts.*** Another simple yet highly effective strategy for identifying a research problem or question is to ask an expert: What needs to be done? What burning

questions are still out there? What previous research findings don't seem to 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 on campus and elsewhere.

Some beginning researchers—including many students—are reluctant to approach well-known scholars for fear that these scholars don't 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're familiar with their work and 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 places 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 or questions. At some 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're 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 other individuals 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 13). 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 may also 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're pursuing an issue of broad scientific or social concern—especially one that is currently a hot topic in your field.

7. ***Be realistic about what you can accomplish.*** Although it's important to address a problem or question that legitimately needs addressing, it's equally important that your research project be a *manageable* one. For example, how much time will it take you to collect the necessary data? Will you need to travel great distances to acquire the data? Will you need expensive equipment? Will the project require knowledge and skills far beyond those you currently have? Asking yourself and then answering such questions can help you keep your project within reasonable, accomplishable bounds.

GUIDELINES Writing a Purpose Statement

Remember, the heart of any research project is your problem or question, along with the goal or goals you intend to pursue in your efforts to address the problem or question. 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 address the problem.

Thus, after identifying a research problem or question, you must articulate it in such a way that *it is carefully phrased and represents the one or more goals of the total research effort.* Following are several general guidelines to help you do that in the overall **purpose statement** that introduces your project to others.

1. ***Describe your problem or question clearly and completely.*** When introducing your research problem or question to others—for instance, when you present it in the introductory purpose statement in your research proposal—you should state it so clearly that anyone else can understand the issue you want to investigate. However, you can state your problem or question 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, we present four meaningless half-statements—verbal fragments that only hint at a purpose. 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 or question.

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 midwestern metropolitan school system?

Retirement plans of adults becomes:

How do adults' retirement plans 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 statements of these problems and questions, the areas studied are carefully limited so that the studies are 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 your purpose statement implies.*** Novice researchers sometimes identify a research problem or question without thinking about its implications. Consider the following purpose statement in a research proposal submitted by a student whom we will call 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 somewhere around 40,000 public and private secondary schools. These schools, north to south, extend from Alaska to Florida; east to west, from Maine to Hawaii. How does John intend to contact each of these schools? By personal visit? At best, he might be able to visit two or three schools per day, so if he worked 365 days a year—in which case many school officials would have to agree to meet with him on a weekend or holiday—he would need more than 40 years to collect his data. And even if John had an exceptionally long life—not to mention exceptional persistence—the financial outlay for his project would be exorbitant.

"But," John explains, "I plan to gather the data by sending a questionnaire." Fine! Each letter, with an enclosed questionnaire and a return postage-paid envelope, might cost two dollars to mail. At best, he could expect only a 50% return rate on the first mailing, so one or more follow-up mailings would be required for nonreturnees. And we would need to figure in the cost of envelopes, stationery, printing, and data analysis.

A faster and less expensive option, of course, would be to conduct the survey by e-mail, possibly with a hotlink to an interactive website. In that case, John would need to track down the name and chief administrator of every one of those 40,000 schools. How long might it take him to do that? And how many of his e-mail messages might be automatically relegated to an administrator's spam box and thus never opened and read?

Obviously, John didn't intend to survey every secondary school in the United States, yet that's what he wrote that he would do.

3. ***Say precisely what you mean.*** You can't 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 purpose 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:

> In this study, I propose to survey the science programs *in selected secondary schools throughout the United States* for the purpose of . . . .

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 purpose statement 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 imprecise purpose statement can lead others to have reservations about the quality of the overall research project. If a researcher cannot be meticulous and precise in describing the problem or question to be addressed, others might wonder 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.

Another potential difficulty is this one: A researcher *talks about the problem or question* but never actually *states what it is.* Recognizing that the problem or question needs to be considered within a certain context or background—which it usually does—the researcher launches into a generalized discussion, continually obscuring the problem or question, never clearly articulating it. Take, for example, the following introductory paragraph for a research proposal:

> 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 the 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 have suggested that you look at examples of dissertations that students have completed at your university and elsewhere. Look at the abstracts for a few of those dissertations and notice the directness with which the goals of the research projects are set forth. The purpose of a study should be presented in the first sentence or two: "The purpose of this study was to . . ." No mistaking it, no background buildup necessary—just a straightforward plunge into the task at hand. All research problems and purpose statements should be described with the same clarity.

4. ***State the problem or question 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 write that they intend to *prove* that certain statements are 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 might 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.

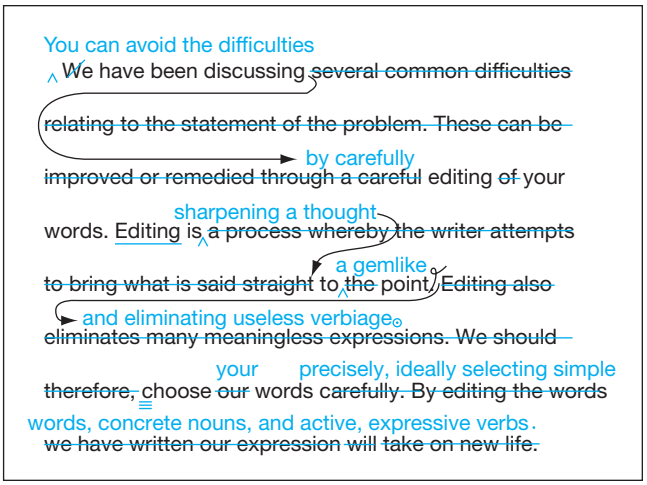
Good researchers try to keep open minds about what they might find. Perhaps they will find the result they hope to find, perhaps not. Any hypothesis should be stated as exactly that—a *hypothesis*—rather than as a foregone conclusion. As you will see later in the chapter, hypotheses play important roles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

Let's rewrite the preceding purpose statement, 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.

FIGURE 2.1 Editing to Clarify Your Writing: An Example



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* mean 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. For example, we replaced an unnecessarily wordy phrase—"a process whereby the writer attempts to bring what is said straight to the point"—with just seven words: "sharpening a thought to a gemlike point." And we replaced the last two sentences with a single sentence that pinned down what good word choice might involve: "simple words, concrete nouns, and active, expressive verbs."

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 present a research statement that is clear, precise, and accurate.

CHECKLIST

Evaluating a Purpose Statement

_____ 1. Write a purpose statement for a research study that reveals the problem or question to be addressed.

_____

_____

_____

_____

_____

_____ 2. Review your purpose statement and ask yourself three questions:
- Are the problem/question and purpose presented in one or more complete, grammatical sentences?
- 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 purpose statement as necessary.

_____

_____

_____

_____

_____

_____ 4. Look at your edited statement and reflect on three additional questions:
- Does the answer to your problem or question have the potential for providing important and useful information?
- Will the result be more than a simple exercise in gathering information, answering a yes-or-no question, or making a simple comparison?
- Is the project focused enough to be accomplished with a reasonable expenditure of time, money, and effort?

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

_____ 6. Show some of your colleagues or fellow 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, rewrite your purpose statement once again:

_____

_____

_____

_____

_____

MyLab Education Self-Check 2.1

MyLab Education Application Exercise 2.1: Stating the Research Problem

DIVIDING THE MAIN RESEARCH PROBLEM OR QUESTION INTO SUBPROBLEMS

Many research problems and questions are too large or complex to be solved without subdividing them. A good strategy, then, is to divide and conquer. Almost any general problem or question can be broken down into smaller units, or **subproblems**—sometimes in the form of more specific questions—that are easier to address and resolve. Furthermore, by viewing the main problem or question through its subproblems or more specific questions, a researcher can often get a better idea of how to approach the entire research endeavor. To keep our discussion simple, we will use the term *subproblems* to encompass both declarative subproblem statements and specific questions that a researcher intends to address.

Subproblems Versus Pseudo-Subproblems

The researcher must distinguish subproblems that are integral parts 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 a researcher must make before being able to effectively address the general research problem and its subproblems. Consider the following as examples:

- Would a quantitative approach, a qualitative approach, or a mixed-methods approach be most suitable for addressing the problem?
- If a sample must be chosen from a larger population, what is the best way to choose the sample, and how large should it be?
- What specific methods should be used to collect data?
- What procedures should be used to analyze the data?

For each pseudo-subproblem, you must decide whether (a) a little common sense and some creative thinking might help in solving it, or (b) you lack the knowledge to address the difficulty. In the latter case, you have several options:

1. Turn to the index of this text to see whether your pseudo-subproblem is discussed.
2. Peruse the "For Further Reading" lists at the end of each chapter in this book to see whether they include sources that might help you, and consult general research methods books in your discipline.
3. Search your university library's catalog and online databases to find potentially helpful books and journal articles. If your library doesn't own what you need, you can typically obtain it through interlibrary loan.
4. Seek the suggestions and advice of more experienced researchers in your field. Recall a point previously 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 researchable issue in its own right.*** A subproblem should constitute a logical subdivision that can be investigated as a distinct project within the larger research endeavor. The solutions to the subproblems, taken together, can then be combined to address and possibly resolve the main problem or question.

It is essential that each subproblem be stated clearly and succinctly. As mentioned earlier, a subproblem is often 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 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 (or question) to make sure that (a) they don't extend beyond the main problem and (b) they address all significant aspects of the main problem.

4. ***Subproblems should be small in number.*** If the main problem or question is carefully stated and properly limited to a feasible research effort, the researcher will find that it usually contains two to six subproblems. Sometimes a researcher will come up with as many

as 10, 15, or 20 subproblems. If you find yourself in this situation, you should study the individual subproblems to see whether (a) some are actually procedural issues (i.e., pseudo-subproblems), (b) some might reasonably be combined into larger subproblems, or (c) the main problem or question is more complex than you originally believed. If the last of these is true, you may want to reconsider whether the solution to your overall problem or question is realistically achievable given the time and resources you have.

Identifying Subproblems

To identify subproblems, you must begin with the main problem or question itself. Write it down, and then carefully scrutinize it to identify more specific problems that should be isolated for in-depth study. All of the subproblems should, in total, comprise your overall problem or question.

You can use either paper and pencil or mind-mapping 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 your main problem or question on paper and then box off the subproblem areas. More specifically, you might follow these steps:

1. Write your problem or question on a clean sheet of paper, leaving considerable space between the lines.
2. Critically examine what you've written in Step 1 to identify specific topics that require in-depth treatment in order for your overall problem or question to be resolved. Draw a box around each topic.
3. Make sure that the text within each box includes one or more words that indicate the need for data interpretation (e.g., *analyze, discover, compare*). Underline these words.
4. Arrange the entire problem or question—which now has its subproblems in boxes—in a graphic that shows the structure of the whole research design.

We use a problem in musicology to illustrate this technique. More specifically, we revisit the motets-of-Palestrina problem, which in this case is described within the context of a purpose statement:

> 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 lifespan 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 <u>to analyze</u> the motets of Palestrina written between 1575 and 1580 to discover their distinctive contrapuntal characteristics, <u>to analyze</u> the same characteristics in the motets of William Byrd written between 1592 and 1597, and to determine what <u>a comparison of these two analyses</u> 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 then *comparing* the two analyses. The three underlined 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 can see exactly what the overall research 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 are parallel in structure: The

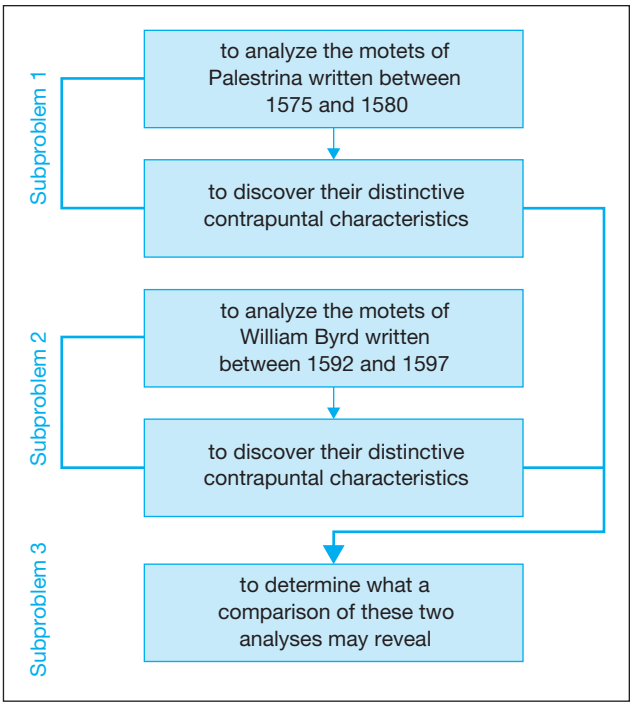


FIGURE 2.2 A Structural Representation of the Palestrina–Byrd Problem

analytical aspect of the subproblem is stated in one box, and the purpose of the analysis is stated in the box right below it. 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—should ultimately resolve the main research problem.

Using Mind-Mapping (Brainstorming) Software

Some computer software programs can facilitate the process of breaking problems into subproblems; you might see these referred to as either *mind-mapping* or *brainstorming* software. Three commercially available, downloadable programs are Inspiration, iMindMap, and XMind. Alternatively, you can find a variety of free software programs (*freeware*) on the Internet; examples are Coggle (coggle.it), FreeMind (freemind.sourceforge.net), and Popplet (popplet.com). Such programs allow you to brainstorm research ideas and construct graphic networks of interrelated concepts, terms, and principles. For example, you can start by putting your main problem, idea, or concept inside a box or oval in a prominent place on your computer screen. As you brainstorm other, related ideas, you put those on the screen as well, and you can draw lines or arrows to show connections among your various ideas. In some programs, you can also label the lines or arrows, perhaps using such phrases as "which includes," "can have an effect on," or "is a subproblem of." And virtually all of these programs enable you to print out your diagram or export it to a word-processing document or other computer file.

As an illustration, we take a remark that, at one time or another, many middle school and high school students make about some of their peers: "No one likes the popular kids." On the surface, this remark is self-contradictory: By definition, the word *popular* means "well liked." What dynamic might underlie so many adolescents' dislike of so-called "popular" classmates? Herein lies a genuine research problem, which we might pose in the form of a question:

Why do many adolescents claim to dislike allegedly "popular" students?

Using Inspiration® 9, we can put this question in an oval, as shown at the top of Figure 2.3. To fully address the problem, we probably want to determine (a) what kinds of students are

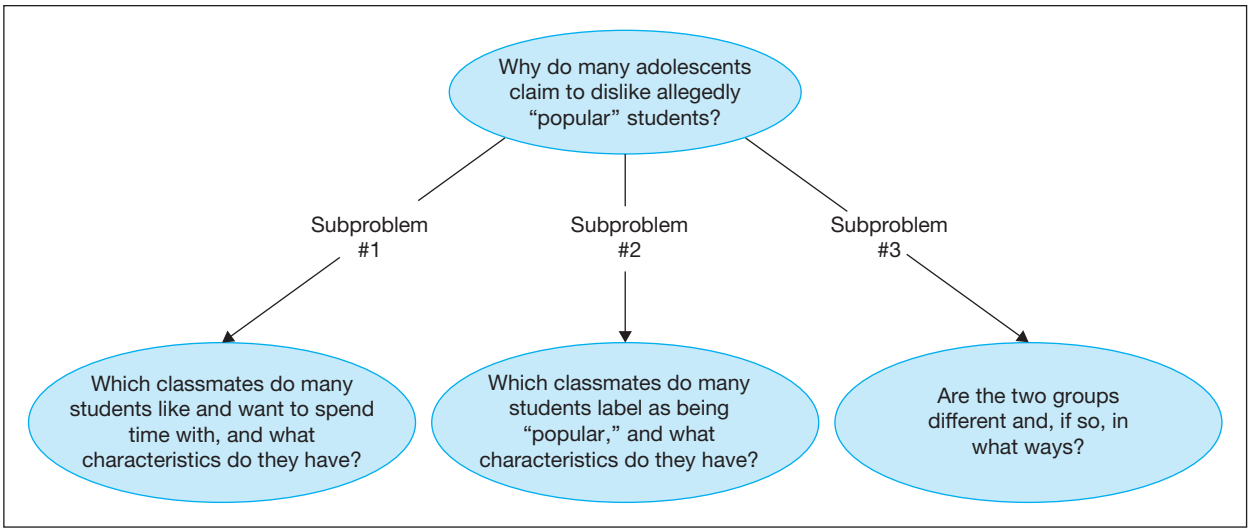


FIGURE 2.3 Translating the Common Adolescent Assertion that "No One Likes the Popular Kids" into a Main Research Question and Three Subproblems

truly popular ones, in the sense that peers like them and enjoy being in their company; (b) what kinds of students peers refer to as "popular" ones; and (c) what differences, if any, exist between the two groups. We can state these three issues as subproblems, also shown in Figure 2.3.

In fact, some researchers have already addressed and answered the four questions presented in Figure 2.3. More specifically, genuinely popular students are seen as being kind, trustworthy, and socially skillful. Some of these students are also labeled as being popular, but other allegedly "popular" students are perceived as being stuck-up, exclusive, and in some cases, mean and aggressive (e.g., see Mayeux, Houser, & Dyches, 2011).

MyLab Education Self-Check 2.2

MyLab Education Application Exercise 2.2: Identifying Research Subproblems

EVERY RESEARCH PROBLEM NEEDS FURTHER DELINEATION

Up to this point, we have been discussing only the research problem (or question), a purpose statement (which usually encompasses the problem), and the problem's subparts. The purpose statement identifies the one or more goals of the research effort, and the subproblems suggest ways of approaching those goals in a manageable, systematic way. But goals alone are not enough. To fully understand a proposed research project, we need other information as well. Both the researcher and the people reading the research proposal should ultimately have a clear understanding of every detail of the process.

At the beginning of any research endeavor, the researcher should minimize possible misunderstandings by

- Identifying a general theoretical or conceptual framework
- If appropriate, stating one or more *a priori* hypotheses
- Identifying the general concepts and possibly also specific variables that are the focus of the investigation
- Defining terms
- Stating underlying assumptions
- If applicable, identifying delimitations and limitations

Such things comprise the *setting of the problem.* We look at each of them in more detail in the following sections. We also include a section on "Importance of the Study" because this topic typically appears in proposals, dissertations, journal articles, and other lengthy research reports.

Identifying a Relevant Theoretical or Conceptual Framework

Many researchers tie their research studies to a particular **theoretical framework**—that is, to a preexisting or newly proposed set of concepts and principles that, in combination, might shed light on cause-and-effect and certain other relationships within the phenomenon under investigation. Such a framework may provide a basis for one or more *a priori* hypotheses about what the data might reveal and to some degree guides the research design and data analyses. In some cases, a researcher's main research problem is to find support for or to disconfirm a particular theory. Alternatively, the research problem might involve comparing the relative efficacy of two or more different theories in explaining new data regarding the issue at hand.

Whereas a theoretical framework implicitly includes possible explanations of cause-and-effect relationships, a **conceptual framework** simply identifies key concepts and inter-concept connections that, taken as a whole, provide a helpful "lens" through which the researcher looks at certain processes within the phenomenon under investigation and can impose some sort of *meaning* on the phenomenon. Here we authors ask you to recall our seemingly silly discussion of the concept *cow* in Chapter 1. As illustrated in that discussion, the many concepts central to various topics and academic disciplines—and possibly also one or more new concepts that a researcher creates—can help the researcher and other individuals both (a) think more clearly and abstractly about a phenomenon and (b) draw inferences about how those concepts are likely to manifest themselves in new situations.

Oftentimes researchers create diagrams or other graphics that present their conceptual frameworks in visual form. For example, Figure 2.4 shows a general conceptual framework that educational psychologist Hadley Solomon has used in her research regarding possible reasons that some teachers prematurely leave the teaching profession. As shown at the top of the graphic, Dr. Solomon proposes that *emotion* plays a key role in teachers' decisions to retire early or move on to other career paths. She further proposes that motivational variables (*goals, beliefs,* and *values*) are influential, and she ties these concepts to three theoretical orientations in the motivation literature (listed in small font at the lower left side). Teachers'

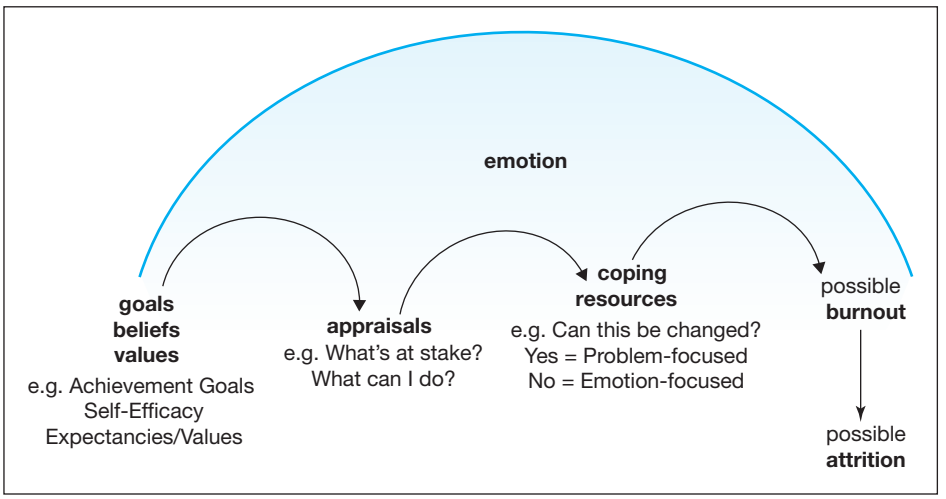


FIGURE 2.4 A Conceptual Framework that Might Be Used in an Investigation of Possible Reasons Underlying Teacher Burnout and Departure from the Teaching Profession

Source: Created by and used with permission from Hadley Solomon.

goals, beliefs, and values might influence their *appraisals* of how effective they can be in their work with students—more specifically, how important their work is and how much control they have over what they can accomplish. Their appraisals, in turn, can influence how well and in what ways they think they can *cope* with the obstacles they face in their efforts to enhance students' achievement and general well-being; if teachers determine that the obstacles are insurmountable, they may experience *burnout* and possibly leave the teaching profession (*attrition*).

We urge you not to agonize over the distinction we're making between theoretical and conceptual frameworks here. Other scholars whose work we have read have distinguished between them differently than we do here. And still others seem to use the two terms interchangeably. Obviously, the terms overlap in meaning, with virtually any theoretical framework having a conceptual framework at its core. And as illustrated in Figure 2.4, a general conceptual framework might include one or more theoretical frameworks.

Stating Hypotheses

As noted in Chapter 1, hypotheses are intelligent, reasonable guesses about how a research problem or question might be resolved. Our focus here is on *a priori* hypotheses—those that a researcher poses in advance, usually in conjunction with the overall research problem or question and its subproblems. When several hypotheses are posed in a research proposal, there is often a one-to-one correspondence between subproblems and their corresponding hypotheses, in which case there are as many hypotheses as there are subproblems.

Hypotheses can guide the researcher toward choosing particular types of research designs, collecting particular kinds of data, and analyzing those data in particular ways. The data may, in turn, support or not support each hypothesis. We must emphasize once again that researchers can rarely, if ever, *prove* their hypotheses. Ultimately, *a priori* hypotheses are nothing more than *tentative propositions set forth to assist in guiding the investigation of a research problem or question.*

A researcher who deliberately sets out to "prove" a hypothesis doesn't have the objective, impartial open-mindedness so important for good research. The researcher might bias the procedure by looking only for 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.

A priori hypotheses are essential to most experimental research (see Chapter 7), and they are sometimes posed in other kinds of quantitative research as well. In contrast, many researchers who conduct strictly qualitative studies intentionally *do not* speculate in advance about what they will find, in large part as a way of keeping open minds about where their investigations will take them and what patterns they will find in their data.

Distinguishing Between Research Hypotheses and Null Hypotheses in Quantitative Research

Up to this point, we have been talking about *research* hypotheses—those educated guesses that researchers hope their data might support. But because researchers can never really prove a hypothesis, they 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 can't actually prove this 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 students in Program A—they can reject the no-difference hypothesis and thus 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**. Most null hypotheses are *not* appropriate as *a priori* hypotheses. Instead, they 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 General Concepts and Possibly Also Specific Variables That Are the Focus of the Investigation

At a minimum, a researcher must identify certain general concepts that underlie a proposed research study. The conceptual framework presented in Figure 2.4 includes many concepts: *emotion, goals, beliefs, values, appraisals,* and so on. And the research question and subquestions presented in Figure 2.1 imply two distinct concepts: *genuinely popular students* (as reflected in peers' positive perceptions of them) and *allegedly "popular" students.*

Many researchers identify, in advance, certain specific variables that will be systematically studied. We have occasionally used the term *variable* in earlier discussions in this chapter and in Chapter 1, but we haven't yet explained what we've meant by the term. We do so now: A **variable** is any characteristic, experience, behavior, or outcome in a research investigation that, by its very nature, 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 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.

The terms *concept* and *variable* overlap, to be sure, and thus we authors urge you not to worry too much about how they might be somewhat different entities. But in our experience, quantitative researchers are more likely to use the term *variable,* in large part because they often pin down in advance exactly what characteristics, experiences, behaviors, and/or outcomes they intend to examine and in what precise ways they plan to assess those things. Explicit identification of variables at the beginning of a study is most common in experimental studies (see Chapter 7) and certain kinds of descriptive studies (see Chapter 6). In contrast, many qualitative researchers prefer to let important concepts or variables "emerge" as data are collected (e.g., see the discussion of grounded theory studies in Chapter 8).

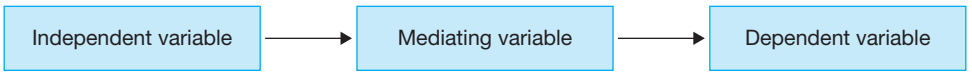
Whenever a research project involves an investigation of a possible cause-and-effect relationship—as is typically true in experimental studies—at least two variables must be specified up front. A variable that the researcher studies as a possible cause of something else—in many cases, this is one that the researcher directly manipulates—is called an **independent variable**. A variable that is potentially caused or influenced by the independent variable—that "something else" just mentioned—is called a **dependent variable** because its status *depends* to some degree on the status of 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:

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 get home, you put one carton in your refrigerator freezer but absentmindedly leave the other one on the kitchen counter. You then go out 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 temperature and is therefore the dependent variable.

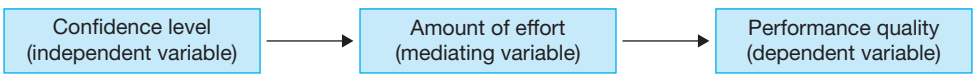
Now let's consider an example in medical research. Imagine that you want to compare the relative effectiveness of two different drugs that are used to treat high blood pressure. You take a sample of 60 men who have high blood pressure and randomly assign each man to one of two groups: The men in one group take one drug, and the men in the second 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 research question or *a priori* hypothesis 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:

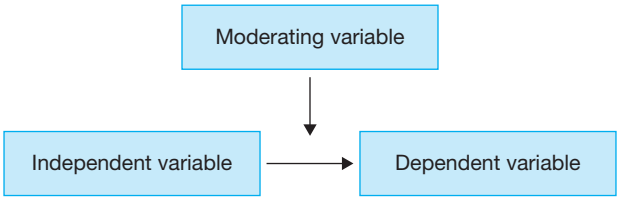


For example, consider the common finding that people who are confident in their ability to perform a particular new task do, on average, actually perform the task better than less-confident people, even if the two groups of people had the same ability levels prior to engaging in the task. Looking at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

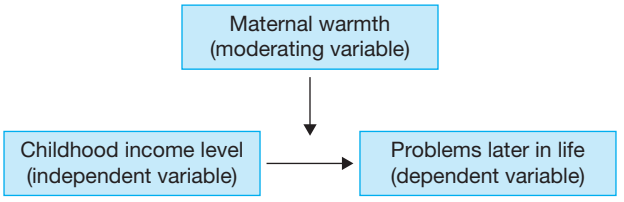
But *why* does this relationship exist? One likely mediating variable is that highly confident people exert more effort in performing the new task than do people with less confidence (e.g., Bandura, 1997; Schunk & Pajares, 2005). The mediating variable, then, is *amount of effort,* as follows:



Still another variable of potential interest is a **moderating variable**—a variable that, although not intervening between the independent and dependent variables, influences the nature and strength of their cause-and-effect relationship, as shown in the next diagram:



As an illustration, 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 this:



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 self-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 either increases or decreases (i.e., has an effect on) 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.

Identifying independent and dependent variables is often quite helpful in choosing both (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.

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, you can gain practice in distinguishing among independent, dependent, mediating, and moderating variables.

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

Following are eight purpose statements for hypothetical research projects. 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. We ask you to 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 the 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.)

Defining Terms

What, precisely, do the terms in the overall research problem or question, purpose statement, and subproblems actually mean? For example, if we say that the purpose of the research is to analyze the contrapuntal characteristics of motets, what are we talking about? What are *contrapuntal characteristics?* Or if we say that a study will investigate the relationship between people's self-confidence levels and the quality of their performance on a task, we need to pin down what we mean by both *self-confidence* and *performance quality.* Without knowing explicitly what specific terms mean—or, more specifically, what the *researcher* means by them—we cannot properly evaluate the researcher's proposed study.

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 within the contexts of the problem or question and any subproblems. Other people who read

the researcher's initial proposal or final 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.

The researcher must be careful to avoid *circular definitions,* in which the terms to be defined are used in the definitions themselves. For example, if a researcher were to define *self-confidence* as "degree of confidence one has in one's own abilities," readers would still be in the dark about what *confidence* actually means within the context of that particular study.

When talking about a variable that has no cut-and-dried, easy-to-pinpoint manifestation in the physical world—and especially when conducting quantitative research—it is often helpful to include an **operational definition**. That is, the researcher defines the variable in terms of how it will be identified and assessed in the research study. For instance, a researcher might define *self-confidence* as a high score on a self-report questionnaire that has items such as "I can usually achieve what I set out to do" and "I think of myself as a smart person." Likewise, a researcher might define *intelligence* as a score on a certain intelligence test or define a student's *popularity* as the number of classmates who identify the student as being a desirable social partner. As another example, let's return to the first scenario in the preceding 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's simply the length of a rat's lifespan in days or some other unit of time. Somewhere in the research proposal, however, the researcher will need to be more specific about how physical exercise and health will be assessed. 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 Assumptions

We have previously discussed assumptions in Chapter 1. Assumptions are so basic that without them, the research problem itself could not exist. For example, suppose a researcher is trying to determine, by means of a pretest and a posttest, whether one approach to in-service technical training in the workplace is superior to another approach. A basic assumption in such a situation is that the pretest and posttest measure knowledge of the subject matter in question.[1] The researcher must also assume that the instructor(s) in the study can effectively teach the desired new skills and that the employees are capable of learning the subject matter. Without these assumptions, the 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 research problem or question should be openly and unreservedly set forth. If others know the assumptions a researcher is making, 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 my problem or question? Ideally, your answer should bring your assumptions into clear view.

Identifying Delimitations and Limitations

A purpose statement describes what the researcher intends to do. But in some cases, it's also important to know what the researcher does *not* intend to do. Whatever things the researcher is not going to do in a research project are known as *delimitations.*

Research problems and questions typically emerge out of 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 boundaries of the problem under investigation.

[1]Alternatively, 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 assessment instruments in Chapter 4.

For example, in the Palestrina–Byrd problem, it's 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 had a determinative influence on Byrd's compositions. But given how the problem has previously been expressed in the purpose statement, the researcher doesn't need to be concerned with *influences* on the motets of the two composers. The researcher should be primarily interested in the *characteristics* of the motets, including their musical style, musical individualism, and contrapuntal likenesses and differences. What the researcher does *not* need to do is to worry about collecting data extraneous to this goal, no matter how enticing or interesting such an exploratory safari might be.

Good researchers also acknowledge that their research projects have certain weaknesses, or *limitations,* that might cast shadows of doubt on results and conclusions. No research project can be perfect, and an honest researcher won't pretend that it is. For example, when studying a certain aspect or quality of human behavior, a researcher might consider such questions as the following:

- Will my sample consist only of people of a certain age range, geographic location, or cultural background? If so, how generalizable are my results likely to be to other populations?
- In what environment will I be conducting the study—in a laboratory, in a classroom, in a workplace setting, on the Internet, or elsewhere? How might this environmental context affect the results I obtain?
- How will I be assessing specific variables in my study? How accurate are my assessments likely to be?
- What personal biases might I be bringing to the study? Are they likely to influence the data I collect or my interpretations of the data?
- What "shortcuts" will I be taking in order to make my study logistically feasible? Might these shortcuts weaken the strength of any conclusions I might draw?

Weaknesses related to these and other issues must be clearly stated in a discussion of limitations, both in the research proposal and in the final research report.

Importance of the Study

In most dissertations and other research reports, researchers set forth their reasons for undertaking the study. Such a discussion may be especially important in a research proposal. Some proposed studies seem to go far beyond any relationship to the practical world. Of such research efforts readers might silently ask, "Of what *use* is it? Will it significantly enhance our understanding of a puzzling phenomenon? How might it contribute to existing theoretical and empirical knowledge about the topic? And what *practical value* does the study have? Will it make an appreciable difference in the health of the planet or in the well-being of one or more species living on it?" Such questions need to be answered.

MyLab Education Self-Check 2.3

MyLab Education Application Exercise 2.3: Selecting Variables to Study

WRITING THE FIRST CHAPTER OR SECTION OF A RESEARCH PROPOSAL

In any research proposal or final research report, the first order of business is to present the general research problem or question, typically within its larger context. For example, as a doctoral student at the University of Maryland, Baltimore County, Christy Leung conducted a mixed-methods study concerning the experiences of Chinese women who had immigrated

to the United States. Following is what she wrote on the first page of her final dissertation, but her strategy would have been similar in her dissertation proposal:

> America has long been recognized as a nation of immigrants . . . many immigrants believe that having freedom and equal opportunity for success and prosperity is possible. Immigrants come to the U.S. with a belief that through hard work, motivation, and persistence, they will be able to earn a better living and provide a better life for their children (Clark, 2003). Many groups, including the Chinese, have chosen to leave their home country because of this belief. The Chinese people have a long history of migration to and settlement in the U.S. to pursue the American dream. Chinese immigrants were once predominantly men who migrated as contract labor workers (e.g., Yung, Chang, & Lai, 2006). However, a series of political incidents and subsequent legislations led to a different wave of Chinese immigration to the U.S. after World War II (Yung et al., 2006; Zhoa, 2002). Changes in the pattern of international migration are important for understanding the adaptation and well-being of immigrants (Massey, Arange, Hugo, Kouaouci, Pellegrino, & Taylor, 1993). (Leung, 2012, p. 1)

In the three paragraphs that followed, Leung expanded on the diverse characteristics and motives of Chinese immigrants, described some of the unique challenges that women were apt to face in moving to the United States, and identified certain key concepts related to the issues her study was addressing (e.g., *socialization, availability of social supports,* and *acculturation*). At that point, Leung had provided sufficient information for readers to understand her purpose and goals in conducting the study:

> (T)he overall goal of this research project was to examine Chinese immigrant mothers' reasons for migration, experiences of migrating to the U.S., . . . acculturation strategies, adjustment, and parenting. . . . (Leung, 2012, p. 3)

After presenting the main research problem or question—often within the context of a purpose statement—a research proposal should identify more specific subproblems to be addressed, along with any *a priori* hypotheses the researcher may have posed related to these subproblems. Somewhere in the introductory section or chapter, a theoretical or conceptual framework should be identified, key concepts or variables should be identified and defined, basic assumptions should be elucidated, and, if appropriate, delimitations and limitations should be put forth. 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 dissertation or other lengthy research report, such topics often comprise the first chapter or section. The document then generally continues with an in-depth discussion of investigations that others have done, usually titled "Review of the Related Literature" or something of that nature. We discuss this review in the next chapter.

PRACTICAL APPLICATION Writing the First Section of a Proposal

In a checklist earlier in this chapter, you developed a purpose statement that presumably encompassed a general research problem or question. 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 completing the following exercise.

1. ***State the subproblems.*** On a blank sheet of paper or new computer document, write the purpose statement you developed earlier. Now inspect your statement carefully and do these things:

a. Within the statement, box off or highlight those areas that need in-depth treatment in order for the problem to be fully explored. Consecutively number these distinct 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 declarative sentences or in questions. Make sure each subproblem includes a word that reflects data interpretation.

2. ***Identify or create a theoretical or conceptual framework that can guide your study.*** Will you be drawing from an existing theory as you plan and conduct your investigation? If so, briefly describe the theory and its relevance to your research problem and to one or more of its subproblems. If, instead, you are unaware of a useful theoretical perspective, identify major concepts that will provide a foundation for your study. Possibly you might also create a graphic that illustrates the theoretical or conceptual framework on which you will base your study, with lines that show logical interconnections (e.g., cause-and-effect relationships, category and subcategories) among the concepts depicted.

3. ***State any* a priori *hypotheses you might have.*** Are you expecting to find certain kinds of results related to one or more of your subproblems? If so, write your research hypotheses, along with a brief rationale for each one. Your rationales should be either theoretically or logically defensible. The sections on deductive logic, inductive reasoning, and theory building in Chapter 1 can help you complete this step.

4. ***Identify and define key concepts and (possibly) specific variables that will be the focus of your study.*** Specify the particular characteristics, conditions, and/or behaviors that are either stated or implied in your research problem or question or in one or more of your subproblems. Give a short but precise explanation of what each concept or variable means *in your particular study*—for instance, how you intend to study it and in some way determine its existence and, if applicable, its specific values.

5. ***Write your assumptions.*** Make 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 or other entities you will be studying, the relevance (or nonrelevance) of the environmental context in which you will be conducting your study, and your assessment techniques.

6. ***Identify and describe delimitations of your study—that is, specific boundaries beyond which you will not go.*** Review the earlier discussion of delimitations. Now write several topics and questions related to your research problem or question that your research project will *not* address.

7. ***Identify and describe possible limitations of your study.*** Identify potential weaknesses of your study related to your proposed sample, data-collection environment, assessment techniques, and personal biases, as well as any planned "shortcut" strategies that may adversely affect the quality of your results and the credibility of your conclusions.

8. ***Explain the importance and significance of your study.*** In a short paragraph or two, explain why your study is worthy of being conducted. 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. For now, however, keeping it in a separate section with its own label can help you remember that *explaining* your study's potential contribution to your field of inquiry is important in its own right.

9. ***Type your proposal.*** 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?
- If so, have you ascertained that the project will advance the frontiers of human knowledge in an important way?

_____

_____

- If not, have you asked one or more experts 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 are the potential strengths of your proposed project?

_____

_____

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

_____

_____

_____ 3. What procedures will you follow going forward?
- Do you have a tentative plan for reviewing 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 for how you will 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 of your 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've given you many suggestions for identifying an appropriate problem or question to drive your research. We can't overemphasize this fact: If you don't precisely pin down both (a) your research problem or question and (b) your goals for conducting your study, 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 or Question

Earlier in the chapter we offered guidelines for identifying and stating an appropriate research problem or question. We now give you a few general suggestions for fine-tuning the problem or question you've identified.

1. ***Conduct a thorough literature review.*** You have presumably already looked at some of the literature related to your overall research problem or question. A next critical step is to make sure you know enough about your topic that you can ask yourself important methodological and procedural questions and then make solid decisions about how you might address them in your research endeavor. You may find that you need to revise your research plan significantly once you have delved more deeply into the scholarly literature related to your topic.

2. ***Think concretely about how you will conduct the subsequent steps you must take in your project.*** 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 potential bottlenecks and pitfalls that might cause problems later on.

3. ***Discuss your research problem or question with others.*** Beginning researchers frequently need to revise their research problems, questions, and goals in order to clarify them and make them more manageable. One good way to do this is to show your purpose statement to other people. If they don't understand what you intend to do, further explanation and clarity are needed. One can learn a great deal from trying to explain something to someone else.

As you proceed with your project, continue to ask others for feedback. Don't be overly discouraged by one or two individuals who are nonconstructively critical or dismissive. Many great discoveries have been made by people who were repeatedly told that they couldn't do what they set out to do.

4. ***Remember that your project will take time—lots of time.*** All too often, we authors 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 or question, conducting the necessary literature review, collecting and interpreting 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're 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's worth however much of your time and effort it takes to do the job well.

5. ***Also 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.

APPLICATION EXERCISES

Similar exercises are available on Pearson MyLab Education.

1. The following are incomplete statements that hint at a research problem. For each incomplete statement, write a complete sentence that clearly describes a researchable problem that could effectively address the topic.
   - **i.** Full inclusion of children with disabilities in general education classrooms.
   - **ii.** Moral reasoning of undergraduate psychology students.
   - **iii.** Writing longhand versus using a word-processing program.
   - **iv.** Students' experiences in after-school programs.
2. Identify the subproblems embedded within each of the following research problems.
   - **i.** The purpose of this study is to determine if students can learn a memory strategy and then apply the memory strategy to learn vocabulary terms in chemistry classes. Two specific memory strategies will be investigated: the pegword method and the keyword method.
   - **ii.** Does the use of a peer-coaching model as a tool for enhancing workplace productivity among newly hired accounting majors facilitate the exchange of ideas and foster greater confidence in decision making on the job?
3. A *variable* is any quality or characteristic of a research investigation that has two or more possible values. For each of the following research statements, identify the variables.
   - **i.** Research statement: The purpose of the current study is to test for sex differences in the self-esteem of fourth-grade children. What is the independent variable, and what is the dependent variable?
   - **ii.** Research statement: For decades, researchers have questioned the effects of pictures on learning from text. Some suggest that pictures distract the reader, whereas others indicate that pictures assist the reader in comprehending the text. This study compares ninth-grade readers' comprehension of science text with and without pictures. What is the independent variable, and what is the dependent variable?
   - **iii.** Research statement: Research has consistently indicated that the transition to college life is difficult. The purpose of this study is to describe the typical first-year experience for women at a small liberal arts college. What is the variable under study here? Why is there no independent variable?

FOR FURTHER READING

Booth, W. C., Colomb, G. G., Williams, J. M., Bizup, J., & Fitzgerald, W. T. (2016). *The craft of research* (4th ed.). Chicago, IL: University of Chicago Press.

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.

Gay, L. R., Mills, G. E., & Airasian, P. (2012). *Educational research: Competencies for analysis and application* (10th ed.). Columbus, OH: Pearson. [See Chapter 2.]

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis.* New York, NY: Guilford Press.

Hesse-Biber, S. N. (2010). *Mixed methods research: Merging theory with practice.* New York, NY: Guilford Press. [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. (2010). *Research in education: Evidence-based inquiry* (7th ed.). Columbus, OH: Pearson. [See Chapter 3.]

Medawar, P. B. (1979). *Advice to a young scientist.* New York, NY: Harper & Row.

Neuman, W. L. (2011). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Hoboken, NJ: Pearson.

Plano Clark, V. L., & Badiee, M. (2010). Research questions in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Mixed methods in social & behavioral research* (2nd ed., pp. 275–304). Los Angeles, CA: Sage.

Roberts, C. M. (2010). *The dissertation journey: A practical and comprehensive guide to planning, writing, and defending your dissertation* (2nd ed.). Los Angeles, CA: Corwin.

Schram, T. H. (2006). *Conceptualizing and proposing qualitative research* (2nd ed.). Columbus, OH: Pearson. [See Chapter 5.]

Wimmer, R. D., & Dominick, J. R. (2014). *Mass media research: An introduction* (10th ed.). Boston, MA: Wadsworth/Cengage.

ANSWERS TO THE CONCEPTUAL ANALYSIS EXERCISE "Identifying Independent, Dependent, Mediating, and Moderating Variables":

1. The phrase "effects of . . . on" tells 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. Test 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 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 doesn't 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 apparently unmeasured mediating variables), which in turn increase body mass index (dependent variable).

This page intentionally left blank

Chapter 3 Review of the Related Literature

Many who have already conducted research on your topic have laid foundational elements—not only previous research findings but also insightful concepts and theories—on which you might productively build. As groundbreaking physicist and mathematician Isaac Newton wrote in 1676, "If I have seen further, it is by standing on the shoulders of giants."

Learning Outcomes

3.1. Describe several purposes that a literature review can serve during the planning of a research project.

3.2. Explain how you might effectively use five general resources to locate related literature: (a) the library catalog, (b) online databases, (c) reference librarians, (d) Internet search engines, and (e) other researchers' reference lists.

3.3. Describe concrete strategies you can use to evaluate, organize, and synthesize literature related to a research problem.

As noted in Chapter 2, reading the literature related to your topic of interest can help you formulate a specific research problem. It can also help you tie your problem—and, later, your findings as well—to a larger body of research and theoretical understandings about your topic. In this chapter we discuss the importance of the literature review and give you suggestions for reviewing the related literature thoroughly but efficiently.

UNDERSTANDING THE ROLE OF THE LITERATURE REVIEW

Research proposals and reports typically have a section—often an entire chapter in a thesis or dissertation—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, although 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 many benefits:

1. It can help you ascertain whether other researchers have already addressed and solved your research problem 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 alert you to controversial issues and gaps in understanding that have not yet been resolved—issues and gaps you might address in your own work.

5. It can show you how others have handled methodological and design issues in studies similar to your own.
6. It can reveal sources of data you may not have known existed.
7. It can introduce you to assessment tools that other researchers have developed and effectively used.
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, especially if you find that others have invested considerable time, effort, and resources in studying it.

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

In most instances, researchers begin their review of the literature early in the game, and they draw on existing theories and previous research studies to help them pin down their research problem or question. Extensive literature reviews *up front* are especially common in quantitative research, where they might help researchers formulate specific *a priori* hypotheses in conjunction with their problem or its subproblems. In some forms of qualitative research, however, researchers worry that too much knowledge about current perspectives and previous findings might unduly bias their own data collection and interpretation; hence, they postpone a thorough literature review until relatively late in the research process (e.g., see the discussion of grounded theory research in Chapter 8).

MyLab Education Self-Check 3.1

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, and conference presentations, and perhaps also on Internet websites. Obviously, you can't simply wander aimlessly through the library stacks or the 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 research problem or question and its subproblems. For example, imagine that you want to investigate possible reasons why people commit serious crimes against individuals whom they see as somehow "different" from and "inferior" to themselves. Obvious keywords for this topic are *aggression, hate crimes, racism,* and *prejudice.* These are very general concepts, 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, websites, 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 hate-crimes problem, these might include such terms as *stereotyping, racial pride, in-group bias,* and *antisocial behavior*—and may also help you focus your research problem a bit more.

Armed with your keywords—which you will undoubtedly continue to revise—you can proceed in several directions. In the following sections, we describe five good starting points: the library catalog, online databases, reference librarians, the Internet, and other researchers' citations and reference lists. In a subsequent section about style manuals, we talk about the kinds of information you will need to collect about the sources you use.

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, when you entered 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 bookshelves (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 might 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 is 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 asking 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 on the topic; some of these have the word "Handbook" in their titles (e.g., *Handbook of Peer Interactions, Relationships, and Groups; Handbook of Cultural Developmental Science*). And don't overlook general textbooks in your discipline. A good university-level textbook can give you a broad overview of a topic, including important concepts, theoretical perspectives, a sampling of relevant research, and critical references. However—and this is important—as a general rule, you should *not* cite basic textbooks to support the points you are trying to make in your literature review.

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're 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 you to find periodicals about any single topic close together, but you must know the call number(s) relevant to your discipline and topic. Other university libraries organize and shelve paper periodicals alphabetically by title; this approach enables you to find any particular periodical without having to consult the library catalog, but you 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 microform 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're quite easy to use once you've had a little practice. Don't be afraid to ask someone behind the periodicals or microforms desk to show you how to use them.

One general rule of thumb is to 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 or question. You should ignore this rule, of course, if you're specifically interested in how perspectives about your topic have changed over the years.

A second rule of thumb is to focus on publications that are likely to have credibility with experts in the field. For example, credible books often come from publishing houses and university-affiliated publishers that specialize in scholarly works (e.g., Sage, Routledge, Oxford University Press). And as previously noted in Chapter 1, reputable journals are typically *peer-reviewed,* in that notable scholars have carefully scrutinized article manuscripts before they ever appear in published form; a quick look at the names and affiliations of a journal's editors and editorial board can give you a good sense of the rigor with which articles have been screened. We urge you *not* to be seduced by bestselling paperbacks on trendy topics, as their authors and contents haven't necessarily been vetted by experts.

If you have access to the Internet from your home computer—as most people do these days—then you also have access to countless online library catalogs around the world. Typically, the home page for your own institution will have a quick link to the library and its catalog. And an Internet search can quickly give you links to many other university and public library catalogs.

A Few Words About Call Numbers The *call numbers* referred to earlier are the unique identification codes that books, journals, 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 10 basic areas of knowledge and subsequent subareas, each divided decimally. The Dewey decimal system is the principal classification system used in many public libraries.
- ***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. You can find descriptions of more specific categories and subcategories on many Internet websites.

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 aren't always classified exactly as we ourselves might have classified them.

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.

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

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 online catalog. In many cases, when one goes to a library shelf to get a particular book or journal, the most useful information 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. For example, to identify the needs and shopping patterns of different populations, marketing researchers often draw on sociologists' and geographers' concepts and data-collection techniques, and psychologists can learn a great deal about human thought processes by using the positron emission tomography (PET), magnetic resonance imaging (MRI), and magnetoencephalography (MEG) technologies of neurophysiologists. Thus, you're 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 Online Databases

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. For this and other reasons, virtually all college and university libraries provide access to many online databases that enable searches of thousands of journals and such other

TABLE 3.2 Examples of Online Databases

Database	*Subject Area(s) Covered*
Academic Search Premier	Education, humanities, multicultural issues, sciences, social sciences
America: History and Life	History of the United States and Canada
AnthroSource	Anthropology
Applied Science and Technology Source	Applied sciences and technology (e.g., computing, engineering, resource management, telecommunications, transportation)
Art and Architecture Source	Broad range of art topics (e.g., advertising, architecture, art history, folk art, graphic arts, video)
Biological Abstracts	Biology, medicine, agriculture
Business Source Premier	Business, economics
Directory of Open Access Journals	All disciplines
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	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) Index	Public and social policy, social sciences
ProQuest Dissertations and Theses Global	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
SPORTDiscus	Recreation, physical fitness, exercise physiology, coaching, sports medicine
Web of Science	Humanities, sciences, social sciences
WorldCat	All disciplines

sources as books, chapters in edited books, dissertations, government documents, technical reports, and newspapers. Table 3.2 lists examples.

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 can conduct either an "easy search" (the default mode) or an "advanced search." If you click on "advanced search," you can type one or more words or phrases in boxes near 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 an author, title, keyword, word or phrase in the abstract, or some other entity.
2. In pull-down menus to the left of the second and any subsequent boxes at the top of the screen, you can tell the computer to do one of the following:
   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).[1]
3. Options in the lower portion of the computer screen allow you to limit your search results still further, perhaps by specifying a particular type of resource (e.g., peer-reviewed), population, age group, or range of publication dates. One of the options is to search in **open-access journals**—scholarly journals that anyone can read on the Internet and download without cost. Many (although not all) of these journals are peer-reviewed and reflect researchers' interest in making their findings available to a wider audience than would otherwise be possible.
4. Once you have limited your search to some degree (at a minimum by completing Step 1), you click on the "Search" button 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 ("did not match any documents"). If references appear, you can click on them to get more information—usually an abstract and, in some cases, the entire text. If your search has been unsuccessful, you probably need to "edit your search" by eliminating 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 the "Search" button once again.
6. Each time you identify a potentially useful source, you can use one or more tools to keep track of it, perhaps adding it to an electronic folder, printing it, or e-mailing it to yourself. You might also import the source to a bibliographic database software program on your computer; we will describe such software later in the chapter.

As is true for some of the articles in PsycINFO, many databases provide entire documents. For example, ProQuest Historical Newspapers: *The New York Times* allows you to search—and then also read—news articles, editorials, letters to the editor, birth announcements, obituaries, advertisements, and virtually any other entry in any issue of the *Times* 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, arts, humanities, business, and law.

One especially helpful database during a literature search is the Web of Science, which can tell you which 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 isn't 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 an "old" one, and thus you would want to find more up-to-date articles on the same topic. In the Web of Science database, its "Citation Network" feature enables you to search

[1] These three options for including or excluding certain words/phrases in your search are known as *Boolean operators*.

the reference lists of all sources in its database and find more recently published works that cite the article by Sowell and her colleagues. If you were to use the Web of Science for this specific purpose (as one of us authors did), you would find that the article by Sowell and colleagues has been cited by hundreds of other researchers and so obviously has been an influential one in neuroscience.

An 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 certain online databases. Many professional associations give current members access to electronic copies of articles published in the associations' journals. In addition, 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 identified are likely to provide complete articles and other documents you can download and print (e.g., look for sites with a "pdf" label), whereas others provide abstracts and links to companies that charge a service fee for the complete documents. Within Google Scholar, one option is to link your search to your own university library's resources; as this edition of the book goes to press, you can do this by clicking on the "My Library" option at the top of the home page. However, we urge you *not* to use Google Scholar as your *only* mechanism for identifying sources relevant to your research problem, as it reportedly can miss some key works in your field.

Another, more specialized database—one especially helpful for researchers interested in medicine and related topics—is PubMed, developed and updated by the U.S. National Library of Medicine (nlm.nih.gov). And for documents produced by various federal agencies in the United States, you can use the Federal Digital System, or FDsys, developed and maintained by the U.S. Government Printing Office (gpo.gov).

Also, check out Google Books (books.google.com), which provides excerpts from many books and in some cases the entire texts of books—especially those whose copyrights have expired. If you're interested in browsing a recently published book to determine whether its contents might be useful to you, you can sometimes access its table of contents and one or more excerpts by using the "Look inside" feature on Amazon's website (amazon.com) or the "Read sample" feature on the Barnes & Noble website (barnesandnoble.com).

Yet another invaluable database is WorldCat (worldcat.org), 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, audio 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. Please don't hesitate to consult with a reference librarian about databases that might 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 or standing 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 resources, online databases, or any of the library's other resources. Larger university libraries may even have librarians who specialize in your academic discipline.

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 can't possibly keep up with them all. As we have already said twice in this chapter, 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, and 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, PubMed, and Google Books. With each passing year, the Internet becomes an increasingly valuable source of information for researchers. As our readers almost certainly already know, an Internet search begins with a **search engine** at a website such as Google.com, Bing.com, or DuckDuckGo.com. 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 to be used in your search. (For example, to limit your search only to children who have autism, you should type "+children" and "+autism." Otherwise, you might 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 many different types of websites. For instance, it may lead you to government websites that can provide helpful documents and information, including those for the U.S. Census Bureau (census.gov), U.S. Department of Education (ed.gov), U.S. Department of Labor (dol.gov), National Aeronautics and Space Administration (nasa.gov), and U.S. Geological Survey (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 (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 peer-reviewed:* 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—we repeat,* ***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

unpublished research reports posted on the Internet haven't 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 can't verify as the work of credible scholars.

Using Citations and Reference Lists of Scholars 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 resource—in our minds, an *essential* one—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 suggest that you 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.*

Identifying Your Sources: The Importance of Adhering to an Appropriate Style Manual

Whenever you find a potentially useful source of information, you should record enough information about it to ensure that (a) you can track it down later if you need to get more information from it and (b) future readers of your research proposal or final research report can find it and read it for themselves. Thus you will need to record the following:

- Author(s)
- Title (in the case of a journal article, both the title of the article and the title of journal)
- Date of publication (the year and, if relevant to the source, the month and day)
- Page numbers (for both journal articles and chapters in edited books)
- Publishing company and its bricks-and-mortar location (for books)

If you use a source you have found on the Internet, you should also record where or how you found it. One common practice is to record the address (**Uniform Resource Locator**, or **URL**) at which you found the resource and, if that address is likely to change over time, the date on which you retrieved it. Alternatively, many online documents posted since the year 2000 have a **Digital Object Identifier**, or **DOI**—a unique, permanent number that enables others to find a document again even if its precise location on the Internet has changed in the meantime. DOIs are especially helpful when research reports and other scholarly works are available only in electronic form (for more information, go to doi.org).

Exactly how you will identify each source in your research proposal or final report depends on your academic discipline. Different disciplines prescribe different styles for both in-text citations and final reference lists. For example, many social scientists use American Psychological Association (APA) style; we authors use APA style in identifying our sources in this book. In contrast, historians typically use Chicago style, and many scholars in the humanities use Modern Language Association (MLA) style. Table 3.3 lists four commonly used styles, along with sources of information about each one. You should consult with your academic advisor about the appropriate style to use.

TABLE 3.3 Commonly Used Styles in Research Reports

Style	*Manual*	*Online Assistance*
APA Style: American Psychological Association	American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author. *Also see:* American Psychological Association. (2010). *Concise rules of APA style* (6th ed.). Washington, DC: Author. *Also see:* American Psychological Association. (2012). *APA style guide to electronic references* (6th ed.). Washington, DC: Author.	blog.apastyle.org owl.english.purdue.edu
Chicago Style: University of Chicago	*Chicago manual of style* (17th ed.). (2017). Chicago, IL: University of Chicago Press. *Also see:* Turabian, K. L. (2013). *A manual for writers of research papers, theses, and dissertations: Chicago style for students & researchers* (8th ed.). Chicago, IL: University of Chicago Press.	chicagomanualofstyle.org
CSE Style: Council of Science Editors	Council of Science Editors. (2014). *Scientific style and format: The CSE manual for authors, editors, and publishers* (8th ed.). Reston, VA: Author.	Many university libraries and writing centers provide online assistance; search the Internet for "CSE style"
MLA Style: Modern Language Association	Modern Language Association. (2016). *MLA handbook* (8th ed.). New York, NY: Author.	style.mla.org owl.english.purdue.edu

Considering all of the resources we have described in this chapter, you might be thinking that you will be spending the next 10 years conducting your literature review! Don't worry. In the Practical Application sections that follow, we describe (a) how to plan an organized and efficient literature search and (b) how to distinguish between research reports that are and are not worth taking seriously.

PRACTICAL APPLICATION Planning a Literature Search

In Chapter 2 you learned how to identify a research problem or question. You also learned 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 its subproblems provide a way to focus your attention as you read the literature. One concrete and effective approach, using either paper and pencil or mind-mapping (brainstorming) software, involves the following steps:

1. Write the problem in its entirety on the page or computer screen.
2. Write each subproblem in its entirety as well.
3. Brainstorm topics you need to learn more about before you can adequately address each subproblem. These topics become your "agenda" as you read the literature.
4. Go to the library catalog, its online databases, and the Internet to seek out resources related to your agenda.

As an illustration, we return to the research question "Why do many adolescents claim to dislike allegedly 'popular' students?" and its three subproblems, presented once again in ovals at the top of Figure 3.1. Addressing Subproblems 1 and 2 requires knowledge about relevant personal characteristics in adolescence, such as likeability, popularity, and social skills. It also requires knowledge about possible data-collection strategies; questionnaires, interviews, and direct observations of student behaviors are three possibilities. Meanwhile, depending on

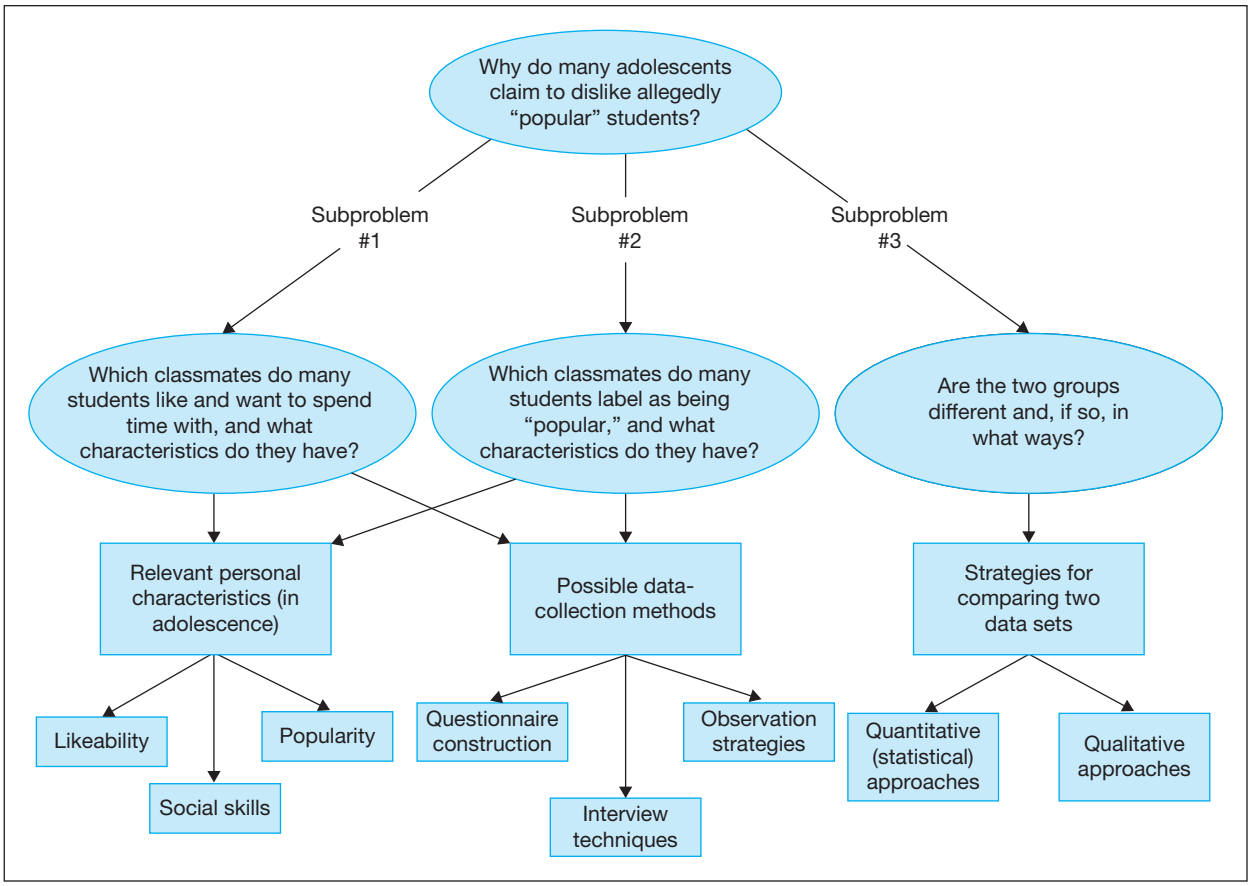


FIGURE 3.1 Using Mind-Mapping Software to Plan a Literature Review

the one or more forms that data-collection strategies take, addressing Subproblem 3 requires knowledge about comparing quantitative data (especially through statistical analyses), qualitative data, or both. Using Inspiration® 9, we present these topics within rectangles in Figure 3.1. A researcher would need to learn more about each of them in the literature review.

The sequence of steps just described can help you keep your literature review within the realm of your research problem or question. 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 consider 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 several times*, especially as your analysis of the literature points you in potentially fruitful new directions. Following are suggestions for maximizing your efficiency at the library.

1. ***Before you go to the library, acquire appropriate software to create a database for the resources you are going to gather.*** If you will be conducting a very limited search of literature relevant to your research problem, you might simply use spreadsheet software

to keep track of potentially useful sources (see Appendix A). But if you're conducting a relatively thorough one—as would be true for a thesis or dissertation—we strongly recommend that you use software specifically designed for creating bibliographic databases. Some software programs are available commercially; examples are Biblioscape, EndNote, and RefWorks. Several no-cost software programs (i.e., freeware) are available as well; as this book goes to press, examples are BiblioExpress, Mendeley, ReadCube, and Zotero.

Bibliographic software programs typically have designated places, or *fields*, for all the information you would need in order to create a reference list—such as author, title, date, journal title and page numbers (for articles), and publishing company and location (for books). 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 or MLA style). Furthermore, the programs have additional fields for call numbers, keywords, any notes you take while reading an item, and (often) any graphics you want to include. And some of them let you incorporate full texts of certain sources, especially journal articles, conference papers, and other documents available in pdf form. If you decide to use a bibliographic software program, you might want to watch one or more of its online video tutorials to learn how to use its many features.

By putting the information you collect in an electronic database, you can later 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 of your sources listed by publication date, the program will, with appropriate instructions, rearrange them in chronological order.

2. ***Identify the materials you want to read (books, articles, etc.), and determine whether your library has them.*** You will probably identify many of the sources you need by consulting the library catalog and perusing indexes and abstracts in online databases. As you make a list of your desired sources, keep the following suggestions in mind:

- ***Keep track of the specific searches you conduct.*** For example, make lists of which indexes and other databases you consult, as well as which keywords you use and in what order. Keeping such records minimizes the likelihood that you will duplicate your efforts.
- ***Whenever possible, let computers make your lists for you.*** For instance, if you're using one of your library's online databases, you will probably be able to print out the sources you identify or, if you prefer, e-mail your list (and in some cases actual journal articles) to yourself. And if you're using a bibliographic software program to organize your literature, you may be able to *import* all of the bibliographic information for a particular source from a library database directly into your own database.
- ***Check the library holdings for the books and journals you identify.*** More specifically, make note of whether the library owns the sources you need and, if so, where they are located and whether they are currently on loan to another user. In the case of journals, you should also check to see whether they are on paper, online, or in microform (e.g., microfilm, microfiche). If the library does *not* have something you need, keep the reference; we will talk about alternative strategies for obtaining such materials shortly.

3. ***Develop an organized plan of attack for finding the sources you have identified.*** Arrange any paper and microform sources you need to obtain according to where they are located in the library. For instance, you can organize books by call number. You can organize journal articles first by paper versus microform, then by the specific journals in which the articles appear. If your university has two or more separate libraries, you will also want to organize your sources by the specific buildings in which they're located.

4. ***Track down your sources.*** After you have organized your sources, you're ready to go find them and look them over. Keep a record of whether each item was (a) found and used, (b) found but determined not to be helpful, or (c) not found. As you peruse the stacks, don't forget to browse neighboring shelves.

5. ***When you find a useful resource, record both (a) essential bibliographic details and (b) potentially helpful details about its contents.*** After skimming a source and determining that it might be useful to you, make sure you have all the information you will need when you include the source in your reference list (author[s], title, date of publication, etc.). In some cases, of course, a source will have so much information that you will simply want to make a photocopy of relevant pages[2] or, in the case of a book, check it out and take it home. If you don't take the actual source home with you, you will want to take sufficient notes about its contents to enable you to re-create its ideas when you are writing your literature review; in other words, you will want to *annotate* each source. For example, it might be helpful to record information about the following:

- Relevant theoretical perspectives
- Data-collection methods
- Details about the samples used (e.g., age range, gender, geographic location)
- Findings
- Conclusions

And you should definitely make note of possible shortcomings of a source. For example, did researchers use such a restricted sample of participants that the findings might not be typical for a larger population? Were data-collection methods so subjective that they virtually guaranteed that the researchers would "find" what they wanted to find? Were conclusions based on faulty reasoning (e.g., inferring a causal relationship from strictly correlational data)? A good literature review not only describes but also *evaluates* what others have previously done and written.

As you make notes, you may find it helpful to keep track of the pages on which you've found certain ideas; furthermore, page numbers are mandatory if you're quoting a source word for word. Also, put quotation marks around word-for-word transcriptions of phrases or sentences in a source. Later, the quotation marks can remind you that these are *someone else's words*—not yours—and must be identified as such in your literature review (more on this point in an upcoming discussion of plagiarism).

"Make haste slowly" is a sound rule for any researcher. Be careful not to write half-complete notes that—when you read them later—are either unintelligible or so lacking in essential information that they're practically useless. If you're 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 the sources themselves for such details.

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, your university catalog system will probably let you put a *hold* on the book so that you're the next one in line to check it out when it comes in. If your 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 save it for you for a few days at the circulation desk.

[2]In the United States, federal copyright law allows one copy for personal use.

- ***Recall the book.*** Many university libraries allow professors to check out books for an entire semester or academic year, but the professors 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—maybe for several months or longer—you can ask for the book to be *recalled* so that you can have access to 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 an **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 for older, out-of-print books.*** Some books identified through Google Books are no longer protected by copyright and thus might be available in online form 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 fee to see the entire book.
- ***Order books from 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 (amazon.com) and Barnes & Noble (barnesandnoble.com).
- ***Use an online document delivery service.*** Some of the online databases listed in Table 3.2 (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 directly to your research problem, you can order complete copies through ProQuest Dissertations & Theses Global (proquest.com). Some of these services may be available to you free of charge through your university library; others may require a fee.

As you conduct your literature review, you will undoubtedly need to repeat this cycle of steps several times. With each go-around, however, you will become more and more of an expert on the topic you're 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, 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

Evaluating a Research Article

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

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

_____ 3. 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 author's research problem or question?

_____ 4. If new 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?

_____ 5. Did the author explain procedures clearly 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?

_____ 6. Do you agree with the author's interpretations and conclusions? Why or why not?

_____ 7. Is the article logically organized and easy to follow? What could have been done to improve its organization and readability?

_____ 8. 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?

KNOWING WHEN TO QUIT

Certainly, you shouldn't read only one or two articles and think that you're done. Get used to looking for and reading new research reports; for a researcher, this is a lifelong endeavor. There are always, *always* new things to learn about a topic.

At some point, however, you must be practical and bring your preliminary literature review to a close. 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, you will eventually begin to realize that you've become quite familiar with certain arguments, methodologies, and findings. Perhaps you will see the same key researchers 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 reasonably knowledgeable about critical parts of the related literature.

Notice our use of the adjective *preliminary* to modify "literature review" in the second paragraph of this section. As you begin to write your review of the literature, you may find certain gaps in your knowledge that need filling. And later on, after you've collected your data, you may find intriguing results within them that additional explorations of related literature might help you sensibly interpret. Thus, you should plan on spending some additional time in your university library or its online equivalent as your project proceeds.

MyLab Education Self-Check 3.2

MyLab Education Application Exercise 3.1: Beginning a Literature Review

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 1998, Jones found that such-and-such. . . . Also, Smith (2004) discovered that such-and-such. . . .
> A few years later, Black (2012) 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, a researcher doesn't merely report the related literature. The researcher 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. In many cases, the subproblems within your main problem or question can 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:

- Identify common themes that run throughout the literature.
- 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.

When you write a literature review that does such things, you have contributed something new to the knowledge in the field even *before* you've 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 13.)

In courtrooms, witnesses swear to "tell the truth, the whole truth, and nothing but the truth." This principle applies to literature reviews as well. Good researchers don't intentionally bias their reviews by omitting or misrepresenting prior research findings that undermine arguments they're trying to make. We remind you of a point previously made in Chapter 2: *Absolute honesty and integrity are assumed in every statement a scholar makes.*

PRACTICAL APPLICATION Writing the Literature Review

Soon after you have read, evaluated, organized, and synthesized the literature relevant to your research problem, you should begin writing the section or chapter that describes the literature you've 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. A section that presents related literature is a discussion of the research studies and other scholarly writings that bear directly on your own research effort.

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

2. ***Develop a plan for the overall organizational structure of your review.*** Writing a good review of the related literature requires advance planning. Before beginning to write your literature review, create an outline of the topics you intend to address and the points you intend to make. A careful consideration of your problem or question and its subproblems should suggest relevant areas for discussion and the order in which they should be addressed.

Begin your discussion of the 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 on the topic. Such studies give an overall historical perspective and provide a context for your own efforts.

3. ***Continually emphasize relatedness to your research problem.*** Keep your readers constantly aware of how the literature you are discussing has relevance to your own project. 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 other people's research and writing; when written in this manner, no attempt is made to demonstrate the relatedness of the literature to the problem being researched. If you can't identify a relationship, you would do well to consider whether you should include the source at all.

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 in the middle of a section. For example, 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 changes in someone else's words a form of plagiarism. For example, early in Chapter 2 we say:

> Some research projects can enhance our general knowledge about our physical, biological, psychological, or social world or 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 can enhance our general knowledge about our physical, biological, psychological, or social world or shed light on historical, cultural, or aesthetic phenomena. Such research can advance theoretical conceptualizations about a topic.

All you would have done here is to replace "Some research projects" with "Basic research" at the beginning of your first sentence and make a few minor 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 your 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. Once again, absolute honesty and integrity are essential as you write.

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, in this case 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 authors do, citations in parentheses within the body of the text—must depend on the particular style manual you're using, which, as mentioned earlier, depends on your particular discipline.

7. ***Minimize your use of direct quotations from other people's writings.*** 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're trying to make about the literature in general; such is the case when we authors present excerpts from students' dissertations 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.

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. We strongly recommend that you *use quotations only when you have a very good reason*—for example, when the specific words that an author uses are as important as the ideas that the author presents. Consistently using other people's words, even when you give those people 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.

Current 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.[3]

As important as what other authors 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 continue to ask throughout the research process—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 simple heading "Summary," you can condense your review into a synopsis of how the existing literature on your topic contributes to an understanding of the specific problem or question you're trying to address.

9. ***Remember that your first draft will almost certainly NOT be your last draft.*** Here we are simply repeating a point made in Chapter 1—a point that applies to a literature review as well as to any other part of a research report. First drafts almost inevitably leave a lot to be desired, in part because (as 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 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

[3]Many publishers now use their websites to post 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—as is especially likely to be the case if you include excerpts from one of *their* publications in a publication of your own—you can often submit a permission request online.

blanks in bold red font (in a computer document) or with sticky notes (on a paper copy), and then make another visit to the library (either to the actual building or to its online resources) in order to fill in those blanks.

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 might 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've 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.

MyLab Education Self-Check 3.3

MyLab Education Application Exercise 3.2: Citing and Paraphrasing Previous Research

MyLab Education Application Exercise 3.3: Organizing Your Literature Review

A SAMPLE LITERATURE REVIEW

At this point, it should be helpful for our readers to look at an excerpt from what is, in our view, a well-written literature review for a doctoral dissertation (Hess, 2016, pp. 22–43). The author of the review, Chelsie Hess, wanted to examine possible relationships between teachers' social and academic support of students and the students' subsequent academic achievement in English/language arts and mathematics. To address her goal, she was able to use portions of an existing data set collected in the Measures of Effective Teaching (MET) project, a large-scale research study funded by the Bill and Melinda Gates Foundation (e.g., Bill & Melinda Gates Foundation, 2010).

Two qualities of the proposal are particularly worth noting. First, the author does not describe the results of relevant studies she is citing in a piecemeal, one-at-a-time fashion; instead, she regularly combines studies with similar findings into a single citation list. Second, the general ideas she presents seem to flow logically from one to another; there's no obvious place where a reader might stop and think, "I'm confused. Did she just switch to a new topic?"

The excerpt we have chosen appears on the left-hand side in the following text. Our commentary appears on the right.

DISSERTATION ANALYSIS 1

Theoretical Grounding for Interactional Quality

The powerful role of teachers in students' academic learning has begun to be established in the literature. However, researchers have not been successful in explaining what exactly it is about a teacher that determines whether students will

Comments

Notice how the first two sentences provide an overview of the discussion that will follow in this section.

be successful. A teacher's experience, educational attainment, and salary are not consistently predictive of students' academic outcomes, and when there have been significant results the effects have been small in magnitude (Croninger et al., 2007; Hanushek et al., 1999; Hanushek & Rivkin, 2006; Jepsen & Rivkin, 2009; Kane et al., 2008; Murnane & Steele, 2007). It is the thesis of this study that a primary influence on students' achievement is having a teacher who effectively supports students' social and academic development through sympathetic, organized, affirming, and academically effective interactions in the classroom. In order to improve the quality of the teacher-student relationship, researchers must first document classroom interactions and their effects (Pianta et al., 2012).

All three "et al." citations in the third sentence have previously been cited with all authors; hence the use of "et al." to cite them once again is appropriate.

In the statement of a "thesis" in this paragraph, the author is essentially presenting a hypothesis that she will be testing in her study. Later, in her Methodology chapter, she translates this thesis into a specific a priori *hypothesis.*

A range of theoretical models converge on the expectation that ongoing sensitive and affectionate interactions with caregivers are essential to children's well-being. In the field of child development, for example, a child's security and willingness to explore the environment are seen to emerge out of first close relationships with one or more familiar caregivers. Adults who are sensitive and create an emotionally supportive, predictable, consistent, and safe environment encourage children to be self-reliant explorers of their environment (Ainsworth, 1979; Bowlby, 1969), and these same concepts have been transferred to and validated in the school environment (Birch & Ladd, 1998; Hafen et al., 2014; Hamre & Pianta, 2001).

With the first sentence of this paragraph, the author transitions into a discussion of concepts and theories that might explain the empirically documented relationship between the quality of teacher–student interactions and student achievement levels.

Of course, there are differences in the roles and effects that adults play at home and at school. At home, a parent takes on numerous functions, for example, tending to the child's physical needs and socializing him or her to take on responsibilities. Teachers play many roles as well and take on the unique duty of imparting academic knowledge and skills. A student's ability to learn is influenced by who is teaching, what is being taught, and the cultural and physical context where the learning is occurring. How teachers implement instruction and build connections with their students are especially influential factors in learning. The importance of a positive relationship between an adult and a child is undisputed, yet the effects of supportive interactions extend beyond social-emotional development (Crosnoe & Benner, 2015).

Hamre and Pianta and colleagues introduced the teaching through interactions (TTI) framework of effective teaching as a lens through which to study classroom structures (e.g., how the school day is organized) and processes within the classroom (e.g., teacher-student interactions). These authors and their colleagues have rigorously tested and elaborated the framework in over 4,000 early childhood and elementary classrooms across the United States (Hamre et al., 2013) and more recently in secondary settings (Allen, Pianta, Gregory, Mikami, & Lun, 2011; Malmberg & Hagger, 2009). The conceptual framework is unique in that it includes three distinct domains (i.e., emotional support, classroom organization, and instructional support), and recognizes the behavioral, cognitive, emotional, and motivational components of teacher-student interactions (Pianta et al., 2012).

In this paragraph, the author introduces a conceptual framework *that will guide much of her study.*

The source "Hamre et al., 2013" has more than five authors; thus, the "et al." is appropriate even in its first citation.

Over the past two decades, Hamre and Pianta have dedicated their efforts to identify and understand the complex social systems of the classroom, along with the

added complexity of teacher-student interactions. The framework has identified three broad domains in an attempt to capture the dynamic of interactions, which includes everything from a teacher's warmth and sensitivity in the classroom to the regular use of scaffolding for increasingly deep academic understandings.

Emotional Support

Pianta's early work revolved around the influences of teacher-child relationships and the emotional support given by early childhood teachers in children's later success in school (Hamre & Pianta, 2001; Pianta, 1994, 1999; Pianta & Nimetz, 1991). Thus the first domain included in the Hamre et al. (2013) TTI framework emphasizes the emotional climate of the classroom and the teacher's emotional expressions, positive affect, sensitivity, and regard for student perspectives. The importance of an adult's expression of emotional support for children has long been recognized and is rooted in early attachment theory (Ainsworth, 1979; Bowlby, 1969). Teachers who are warm and sensitive tend to be more attuned and responsive to students' social, emotional, and academic needs (Hamre & Pianta, 2005; Pianta et al., 2008).

In this subsection within the more general section "Theoretical Grounding for Interactional Quality," the author discusses the first of the three domains delineated in the teaching through interactions (TTI) framework previously introduced. Through the use of the subheading "Emotional Support" and subsequent subheadings, the author can more clearly communicate her overall organizational scheme for the more general section.

Consistent with Pianta's research, when teachers are more attuned and responsive, students are likely to report a greater enjoyment of school and learning and a positive sense of peer community (Gest, Madill, Zadzora, Miller, & Rodkin, 2014). If students feel emotionally connected and supported, then it should come as no surprise that these students, on average, have more positive academic attitudes, are more engaged, and have higher achievement scores (Crosnoe et al., 2004; Deci & Ryan, 2000).

Notice the phrase "on average" in the last sentence of this paragraph. By using it, the author effectively communicates that some students are exceptions to this general rule.

Not only do students thrive in classrooms when teachers are sensitive to their feelings, they also flourish in classrooms where students are encouraged to speak their minds and converse with one another. Regard for student perspectives is included in the TTI framework and has been well documented in educational and motivational research. Students are most motivated to learn when adults support their need to feel competent and autonomous at school (Deci & Ryan, 2000). Students benefit most when teachers actively scaffold the learning experience with a balance of control, autonomy, and mastery in the classroom. For example, student learning is inhibited when there is a mismatch between a student's need for autonomy and the teacher's need to exercise control (Cornelius & Herrenkohl, 2004; Eccles, Wigfield, & Schiefele, 1998). Along with the need for meaningful choices, students are motivated to learn when they feel valued as an individual.

In using the term "need for autonomy," the author is drawing from another popular theoretical perspective: that students are more motivated to succeed when they believe they have some independent choice and control over their actions and activities.

(The author continues with subsections discussing the other two domains outlined in the TTI framework: classroom organization and instructional support.)

Note: From *Associations Between Teacher Interactional Quality and Student Achievement: A Classroom-Level Analysis of Randomized and Non-Randomized Teacher Assignments in the Measures of Effective Teaching Project* (pp. 31–34) by Chelsie Hess, 2016, doctoral dissertation, University of Northern Colorado, Greeley (soon to be available through the online database ProQuest Dissertations & Theses: Full text). Reprinted with permission.

APPLICATION EXERCISES

Similar exercises are available on Pearson MyLab Education.

1. Answer the following questions on beginning a Literature Review.

   i. Consider this tentative research hypothesis: Failing seventh-grade English students enrolled in a peer-tutoring program have higher exam grades than failing seventh-grade English students not enrolled in a peer-tutoring program. Explain the very first step a researcher should take when beginning a review of the literature.

   ii. Use an online database to find three books or journal articles in which you might find information or research related to this hypothesis. Which database should you choose to use first, and why?

   iii. List the complete reference for each book or journal article, and provide a brief summary or abstract of the contents of the book and/or journal article. Use the correct reference format for your discipline.

2. Knowing how to cite and paraphrase properly is critical for any writing activity in which you rely on previously published material. Read through the paragraphs given below, and then answer the questions. The passage is taken from a draft of a review of the literature written by a researcher planning a study on kindergarten readiness.

   With some 500,000 children in California entering kindergarten each year, the concept of "school readiness" increasingly receives public attention and funding. Following the National Education Goals Panel, policymakers operationalize school readiness as including a child's physical well-being and motor development, social/emotional development, approaches to learning, language development, cognition, and general knowledge, but research on correlates of socioemotional success in schools lags behind the other areas.

   The view from national survey data is optimistic: Approximately 85% of kindergarteners are judged by their parents to make friends easily, successfully join groups, and show empathy toward peers, whereas fewer than 25% regularly fight or argue with others. Critically, however, the data reflect clear differences among ethnic groups, with white, non-Hispanic children judged by parents and teachers as most socially skilled. As the percentage of minority children in California grows, researchers must address factors in early childhood that encourage the social outcomes expected in American kindergartens and schools.

   i. The researcher has relied on information from other sources, but the citations have been removed for this exercise. Identify at least two places where citations are required, and explain why.

   ii. The phrase "school readiness" appears in quotations. What, if anything, is required as a citation?

   iii. How might you paraphrase this sentence without plagiarizing?

   "Approximately 85% of kindergarteners are judged by their parents to make friends easily, successfully join groups, and show empathy toward peers, whereas fewer than 25% regularly fight or argue with others."

3. Answer the following questions on organizing your Literature Review.

   i. You have finished reviewing relevant articles for your literature review. How would you begin writing the literature review? How would you make decisions regarding how best to organize literature for review? Which would not be a recommended approach for writing a literature review?

   ii. Draft an advance organizer that lays out areas you will review. Also draft an outline for the review.

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. (2014). *Conducting research literature reviews: From the Internet to paper* (4th ed.). Los Angeles, CA: Sage.

Galvan, J. L. (2012). *Writing literature reviews: A guide for students of the social and behavioral sciences* (5th ed.). Los Angeles, CA: Pyrczak.

Hardy, J., & Dittman, H. (2007). *Learn Library of Congress classification* (2nd North American ed.). Friendswood, TX: Total Recall Publications.

Heyvaert, M., Hannes, K., & Onghena, P. (2016). *Using mixed methods research synthesis for literature reviews.* Los Angeles, CA: Sage.

Jesson, J. K. (2011). *Doing your literature review: Traditional and systematic techniques.* London, England: Sage.

Machi, L. A., & McEvoy, B. T. (2016). *The literature review: Six steps to success* (3rd ed.). Los Angeles, CA: Corwin.

Martelo, M. L. (2011). Use of bibliographic systems and concept maps: Innovative tools to complete a literature review. *Research in the Schools, 18*(1), 62–70.

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.

Chapter 4 Planning Your Research Project

Before constructing a house, a builder acquires or develops a detailed set of plans—how to frame the walls and roof, where to put doors and windows of various sizes, where to put pipes and electrical wiring, what kinds of materials to use, and so on. These plans enable the builder to erect a strong, well-designed structure. Researchers should pay similar attention to detail in planning a research project.

Learning Outcomes

4.1. Distinguish between primary data and secondary data, and describe a variety of forms that data for a research project might take.

4.2. Compare quantitative versus qualitative research methodologies in terms of their typical purposes, processes, data-collection strategies, data analyses, and final reports.

4.3. Describe various strategies you might use to enhance both the credibility and the generalizability of your findings.

4.4. Differentiate between substantial and intangible phenomena and among nominal, ordinal, interval, and ratio scales of measurement.

4.5. Describe various types of validity and reliability that, as applicable, are important for assessment strategies used in a research study, as well as several ways in which you might determine or enhance the validity and/or reliability of an assessment strategy.

4.6 Discuss ethical issues related to protection from harm, voluntary and informed participation, right to privacy, and honesty with colleagues. Also, explain the roles of institutional review boards and professional codes of ethics in minimizing potential ethical problems in a research study.

When we talk about a general strategy for addressing a research problem or question, 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 time and expense are 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 your data will take—always with the central goal of addressing your research problem in mind—at the very beginning of your project.

PLANNING A GENERAL APPROACH

In planning a design, a researcher in quest of new knowledge and understandings should not be shackled by discipline-specific methodological restraints. The course of a research project will frequently lead the researcher into new and unfamiliar territories that have historically been associated with other content areas. A sociologist trying to resolve a problem in sociology may come face to face with problems that are psychological or economic. And on the way to finding a solution for a problem in criminology, a student in criminal justice might venture into the realms of mental illness and behavioral genetics. Any good researcher must be *eclectic*, willing to draw on whatever sources seem to offer productive methods or data for resolving the research problem.

Instead of limiting their thinking to departmentalized knowledge, researchers might better 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 problems and questions related to children, senior citizens, families, communities, cultural groups, organizations, employees, mental and physiological processes, learning, motivation, social interactions, instructional interventions, crime, rehabilitation, medical treatments, nutrition, language, and religion.
- ***Things.*** In this category are problems and questions related to animal and vegetable life, viruses and bacteria, inanimate objects (rocks, soil, buildings, machines), matter (molecules, atoms, subatomic matter), stars, and galaxies.
- ***Records.*** In this category are problems and questions related to newspapers, personal journals, letters, Internet websites, registers, speeches, minutes, legal documents, mission statements, census reports, archeological remains, sketches, paintings, and music.
- ***Thoughts and ideas.*** In this category are problems and questions related to concepts, theories, perceptions, beliefs, ideologies, semantics, poetry, and political cartoons.
- ***Dynamics and energy.*** In this category are problems and questions related 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, and black holes.

We don't 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. Each item in the lists is a potential **unit of analysis**—the "who" or "what" that is the focus of study.

Research Planning Versus Research Methodology

Don't confuse overall research planning with research methodology. Whereas the general approach to *planning* a research study may be similar across academic disciplines, the specific *methodology* is apt to depend somewhat on the discipline, the research problem or question at hand, and the kinds of data that need to be collected and analyzed. You can't 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 can't study chromosomes with a questionnaire, and you can't study people's opinions with a microscope.

Thus, in planning a research design, it's extremely important for the researcher not only to choose a viable research problem or question but also to consider the kinds of data

that an appropriate investigation will require, as well as reasonable means of collecting and interpreting those data. Many beginning researchers become so entranced with the glamour of their overall problem or question that they fail to consider practical issues related to data availability, collection, and interpretation. For example, comparing the brain-wave patterns of children who are gifted versus those of average ability may be a laudable research project, 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 machine at your disposal?
- If so, do you have the technical skills to use it?
- Are you sufficiently knowledgeable to interpret the electroencephalographic data 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's probably better that you abandon this project in favor of one for which you have the appropriate knowledge, skills, and resources. Your research should be *practical* research, built on precise and realistic *planning* and executed within the framework of a clearly conceived and feasible *design.*

THE NATURE AND ROLE OF DATA IN RESEARCH

By definition, the term *research* as we are using it in this book involves collecting and interpreting relevant data. 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.

Researchers must always remember that data are not absolute reality or truth—if, in fact, any single "realities" and "truths" can ever be determined. (Recall the discussions of *postpositivism* and *constructivism* in Chapter 1.) Rather, data are merely *manifestations* of various physical, biological, social, or psychological phenomena that we want to make better sense of. For example, we often see what other people do—the statements they make, 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 really know.

Data Are Transient and Ever Changing

Data are rarely permanent, unchanging entities. Instead, most of them are transient—they may be accurate for only a very short time. Consider, for 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 to the next block, the data they have collected are already out of date. Some people in the previous block who expressed a particular opinion may have subsequently engaged in a neighborhood discussion, watched a television documentary, or looked at a website that changed their opinion. Some people may have moved away, and others may have moved in; some may have died, and others may have been born. Tomorrow, next week, next year—what we thought we had "discovered" may have changed quite a bit.

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 isn't necessarily true the next. Even the most carefully collected data can have an elusive quality about them; at a later time, they may have no counterpart in reality whatsoever. Data are volatile: They evaporate quickly.

Primary Data Versus Secondary Data

For now, let's take a *positivist* perspective and assume that out there—somewhere—is a certain Absolute Truth waiting to be discovered. A researcher's only perceptions of this 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.

Imagine, for a moment, that you live in a dungeon, where you can never see the sun—the Truth. Instead, you see a beam of sunlight on the dungeon floor. This light might give you an idea of what the sun is like. The direct beam of sunlight is *primary data.* Although the shaft isn't the sun itself, it has come directly from the sun.[1]

But now imagine that, rather than seeing a direct beam of light, you see only a diffused pattern of shimmering light on the floor. The sunlight (primary data) has fallen onto a shiny surface and then been reflected—distorted by imperfections of the shiny surface—onto the floor. The pattern is in some ways similar but in other ways dissimilar to the original shaft of light. This pattern of reflected light is *secondary data.*

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

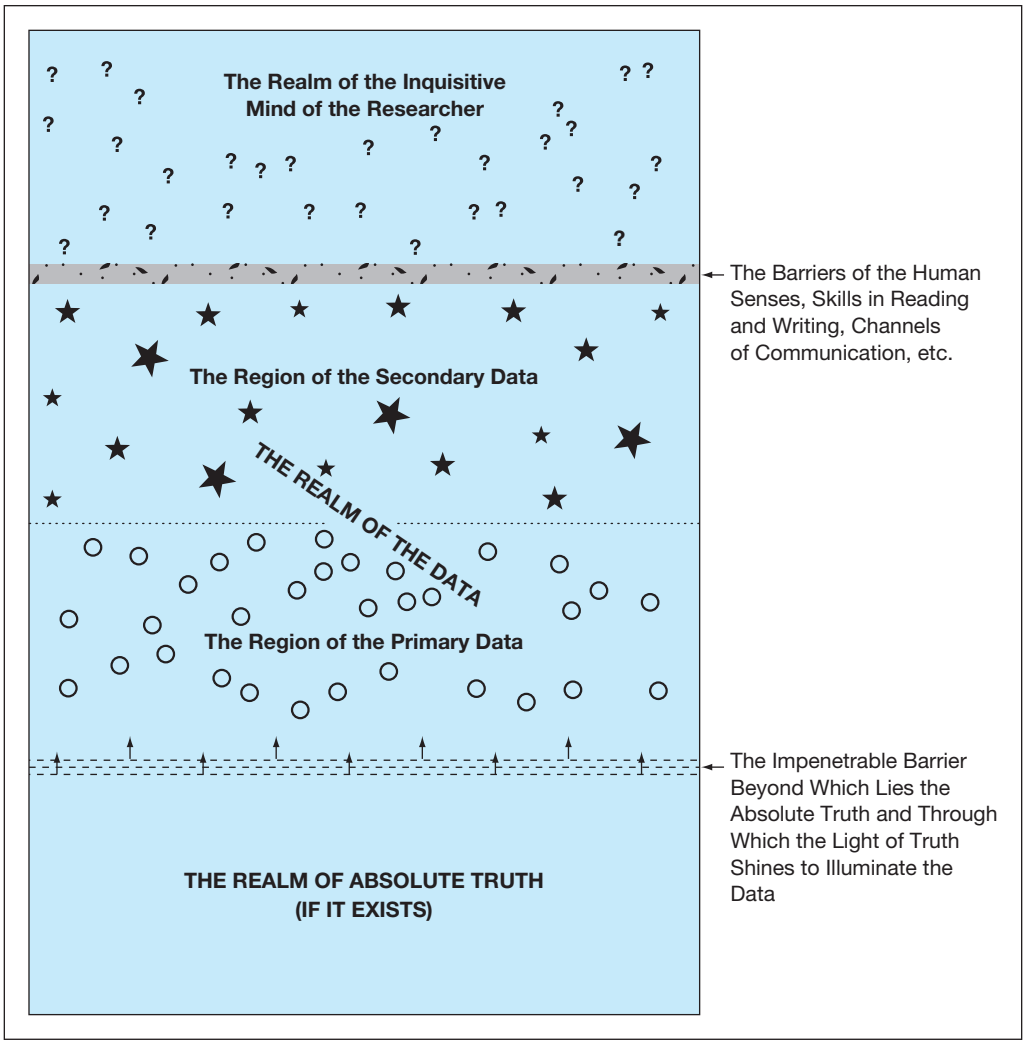
Now along comes a reporter who interviews both you and the driver and then writes an account of the accident for the local newspaper. When your sister reads the account the following morning, she gets, as it were, the reflected-sunlight-on-the-floor version of the event. The newspaper article provides secondary data. The data are inevitably 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 others actually observed.

Figure 4.1 represents what we have been saying about data and their relation to any possible Truth that might exist. 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 composed of many things, including the limitations of the human senses, the weaknesses of assessment strategies, the inadequacy of language to precisely communicate people's thoughts, and the inability of two human beings to witness the same event and report it in exactly the same way.

Researchers must never forget the overall idea underlying Figure 4.1. Keeping it in mind can prevent them from making exaggerated claims or drawing unwarranted conclusions. No researcher can ever glimpse Absolute Truth—if such a thing exists at all—and researchers can perceive data that reflect that Truth only through imperfect senses, imperfect

[1]For 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.*

FIGURE 4.1 The Relation Between Data and Truth



assessment tools, and imprecise channels of communication. Such 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. . . .*

Planning for Data Collection

Basic to any research project are several fundamental questions about the data. To avoid serious trouble later on, the researcher must answer them 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 a ridiculously simple one, but in fact, a specific, definitive answer to it is fundamental to any research effort. To adequately address the research problem or question, what data are mandatory? What is their nature? Are they historical documents? Interview excerpts? Questionnaire responses? Observations? Measurements taken before and after an experimental intervention? Specifically, what data do you need, and what are their characteristics?

2. ***Where are the data located?*** In our experiences as college professors, we have seen many fascinating problems that students would like to address in their 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 *exactly where*? If you're planning a study of documents, where are the documents you need? Are they available on the Internet or, if not, at exactly which library and in what collection will you find them? What agency, society, or other organization has the files you must examine? Specify their location—either by website or by town, street address, and postal code.

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. ***What limits will be placed on the nature of acceptable data?*** Not all gathered data will necessarily be acceptable for use in a research project. Sometimes 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. The restrictions identified are sometimes called the **criteria for the admissibility of data**.

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. 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?

Now imagine a sociologist who plans to conduct a survey to determine people's attitudes and beliefs about a controversial issue in a particular area of the country. The sociologist constructs a 10-item survey that will be administered and collected at various shopping malls, county fairs, and other public places over a 4-week period. Some people will respond to all 10 items, but others may respond to only a subset of the items. Should the sociologist include data from surveys that are only partially completed, with some items left unanswered? And what about responses such as "I'm not going to waste my time on your stupid questions!"—responses indicating that a person wasn't interested in cooperating?

The agronomist and the sociologist should be specific about such things—ideally, in sufficient detail that, if appropriate, another researcher might reasonably replicate their studies.

5. ***How will the data be interpreted?*** This is perhaps the most important question of all. The four former hurdles have been overcome. You have the data in hand. But you must also spell out precisely what you intend to do with them to address your 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 provide a solution, and, if so, might they reasonably lend themselves to credible interpretations? If your answer to either of these questions is *no,* you must rethink the nature of your problem. If, instead, both answers are *yes,* an important next step is to consider an appropriate methodology.

MyLab Education Self-Check 4.1

LINKING DATA AND RESEARCH METHODOLOGY

Imagine that a man from a remote village travels to the big city, from which he subsequently takes his first ride on a commercial airline to reach a second destination. No one else in his village has ever ridden in an airplane, so after he returns home, his friends ask him about his trip. One friend asks, "How fast did you move?" "How far did you go?" and "How high did you fly?" A second one 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, but because they ask different kinds of questions, they obtain different kinds of information. Although neither of them gets the "wrong" story, neither does each one get the whole story.

In research, too, different research problems and questions require different kinds of data, which, in turn, often require different research methodologies. To some extent, then, *the desired data dictate the method.* As an example, consider historical data, those pieces of information gleaned from written records of past events. You can't extract much meaning from historical documents by conducting a laboratory experiment. An experiment simply isn't 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." Many highways can take us toward a better understanding of the unknown. They may traverse different terrains, but they all converge on the same general destination: the enhancement of human knowledge and understandings.

Comparing Quantitative and Qualitative Methodologies

On the surface, quantitative and qualitative approaches involve similar processes—for instance, they both entail identifying a research problem or question, reviewing related literature, and collecting and analyzing data. But by definition, they're suitable for different types of data: Quantitative studies involve numerical data, whereas qualitative studies primarily make use of nonnumerical data (e.g., verbal information, cultural artifacts). And to some degree, quantitative and qualitative research designs are appropriate for addressing different kinds of problems and questions.

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 for 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 others 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 are teachers' and students' views on the effectiveness of case-based instruction?" 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 learn 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 4.1 Typical Characteristics of Quantitative Versus 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 • Preplanned 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?	• Numerical 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 • Primarily deductive reasoning	• Coding for themes and categories • Acknowledgment that analysis is subjective and potentially biased • Primarily inductive reasoning
How are the findings communicated?	• Numbers • Statistics, aggregated data • Graphs • Formal voice, scientific style	• Words • Narratives, individual quotes • Visual illustrations, nonnumerical graphics • Personal voice, literary style (in some disciplines)

Table 4.1 presents typical 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 tend to seek explanations and predictions that, in most cases, will generalize to other persons and places. The intent is often to identify relationships among two or more variables and then, based on the results, to confirm or modify existing theories or practices.

Qualitative researchers tend to seek better understandings 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.

Process In part because quantitative studies have historically been the mainstream approach to research, carefully structured guidelines exist for conducting them. Concepts, variables, hypotheses, and assessment techniques tend to be defined before the study begins and to remain constant throughout. Quantitative researchers choose methods that allow them to objectively *measure* the variable(s) of interest. In order to minimize the chances of collecting biased data, they also try to remain relatively detached from the phenomena being investigated and from the participants in their studies.

A qualitative study is often more holistic and *emergent,* with the specific focus, design, data-collection techniques (e.g., observations, interviews), and interpretations developing and possibly changing along the way. Researchers try to enter a situation with open minds, prepared to immerse themselves in its complexity and, in the case of research with human beings, to personally interact with participants. Specific categories emerge from the data, leading to information, patterns, and/or theories that help explain the phenomenon under study.

Data Collection Quantitative researchers typically identify only a few variables to study and then collect data specifically related to those variables. Methods of assessing each variable are identified, developed, and standardized, with considerable attention given to the validity and reliability of specific instruments and other assessment strategies (more about such qualities later in the chapter). Data are often collected from a large sample that is presumed to represent a particular population so that generalizations can be made about the population.

Qualitative researchers operate under the assumption that reality isn't 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. Both verbal data (interview responses, documents, field notes) and nonverbal data (drawings, photographs, videotapes, artifacts) may be collected. When the focus of study is some aspect of human existence, qualitative researchers tend to select a relatively small number of participants who might best shed light on the phenomenon under investigation.

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 analyses, conducting predetermined statistical procedures and using relatively 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 analyses are more subjective in nature: They scrutinize their bodies 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 might formulate a preliminary theory through inductive reasoning (e.g., by observing a few situations), engage in the theory-building process described in Chapter 1, and then try to support the 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 often move into a more deductive mode to verify or modify it with additional data.

Reporting Findings Quantitative researchers typically reduce their data to summarizing statistics (e.g., means, medians, correlation coefficients). In most cases, *average* performances are of greater interest than the performances of specific individuals (you will see exceptions in the single-case designs described in Chapter 7). Graphs might be used to visually display certain trends or comparisons. Results are usually presented in a report that uses a formal, scientific style with impersonal language.

Qualitative researchers often construct interpretive narratives from their data and try to capture the complexity of a particular phenomenon. They may use photographs or diagrams to illustrate particular characteristics or interactions. Especially in certain disciplines (e.g., anthropology), qualitative researchers may use a more personal, literary or narrative style than quantitative researchers do in their final reports, 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

Given that quantitative and qualitative methodologies are useful in addressing somewhat different kinds of research problems and questions, we can gain better understandings of our physical, biological, social, and psychological worlds when we have both methodologies at our

disposal. Fortunately, the two approaches aren't necessarily mutually exclusive; many researchers successfully combine them in a *mixed-methods design.* For example, it isn't 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 9.

PRACTICAL APPLICATION Choosing a General Research Approach

As you choose your own general approach to a research project—whether to use a quantitative approach, a qualitative approach, or a combination of the two—you should base your decision on the specific problem or question you want to address and on the skills you have as a researcher, *not* on what tasks you want 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

Qualitative studies have become increasingly popular in recent years, even in some disciplines that have historically placed heavy emphasis on quantitative approaches. Yet we authors 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 course of many months. In the following paragraphs, we offer important considerations for novice researchers 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 diverging but possibly equally valid realities relevant to your problem, then qualitative research might be the better route to take.

2. ***Consider the nature of your research question.*** Qualitative designs can be quite helpful for addressing exploratory or interpretive research questions. But they're apt to be of little use in testing specific hypotheses about cause-and-effect relationships.

3. ***Consider the extensiveness of the related literature.*** If the literature base is weak, underdeveloped, or altogether missing, a qualitative design can give you the freedom and flexibility you need to explore a specific phenomenon and identify important facets, concepts, and processes related to this phenomenon.

4. ***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're skimming the surface of a phenomenon and wish to do so using a large number of participants, a quantitative study will be more efficient.

5. ***Consider the amount of time you have available for conducting the study.*** Qualitative studies typically involve an extensive amount of time both on and off any relevant research sites. If your time is limited, you may not be able to complete a qualitative study satisfactorily.

6. ***If applicable, 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. Furthermore, gaining initial entry into one or more research sites (e.g., social meeting places, people's homes) may take considerable advance planning and several preliminary contacts.

7. ***Consider the extent to which you feel comfortable working without much structure.*** Qualitative researchers tend to work with fewer specific, predetermined procedures than quantitative researchers do; their work can be exploratory in many respects. Thus, they must think creatively about how best to address various aspects of a research problem or question, and they need a high tolerance for ambiguity.

8. ***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 the like, which might not be 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 researchers who have little affection for mathematics—can often be a more straightforward task.

9. ***Consider your writing skills.*** As noted earlier, qualitative researchers must have excellent writing skills. Communicating findings is the final step in all research projects; the success of your research will ultimately 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 4.2 lists some common research methodologies and the types of problems for which each is appropriate. In later chapters of the book, we look more closely at most of these methodologies.

TABLE 4.2 Common Research Methodologies

Methodology	*General Characteristics and Purposes*
Action research	A type of applied research that focuses on an existing problem, with the goal of improving current practices and desired outcomes. (See Chapter 10.)
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 8.)
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 8.)
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. (See Chapter 6.)
Design-based research	A multistep, iterative study in which certain instructional strategies or technologies are implemented, evaluated, and modified to determine possible factors influencing learning or performance. (See Chapter 10.)
Developmental research	An observational-descriptive type of research that either compares people in different age groups (a *cross-sectional study*) or follows a particular group over a lengthy period of time (a *longitudinal study*). Such studies are particularly appropriate for looking at developmental trends. (See Chapter 6.)
Ethnography	A type of qualitative inquiry that involves an in-depth study of an intact cultural group in a natural setting. (See Chapter 8.)
Experimental research	A study in which participants are randomly assigned to groups that undergo various researcher-imposed treatments or interventions, followed by systematic assessments of the treatments' effects. (See Chapter 7.)

(continues)

TABLE 4.2 Common Research Methodologies (*continued*)

Methodology	*General Characteristics and Purposes*
Ex post facto research	An approach in which one looks at conditions that have already occurred and then collects data to determine whether a relationship exists between these conditions and subsequent characteristics or behaviors. (See Chapter 7.)
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 8.)
Historical research	An effort to reconstruct or interpret historical events through the gathering and interpretation of relevant historical documents and/or oral histories. (For example, see Brundage, 2013; Yow, 2015.)
Narrative inquiry	A type of qualitative research that focuses on the stories and other recollections of individuals who have had experiences related to a phenomenon under investigation. (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 6.)
Phenomenological research	A qualitative method that attempts to understand participants' perspectives and views of physical or social realities. (See Chapter 8.)
Quasi-experimental research	A method similar to experimental research but without random assignment to groups. (See Chapter 7.)
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 6.)

MyLab Education Self-Check 4.2

MyLab Education Application Exercise 4.1: Selecting a Method that Suits Your Purpose

CHOOSING METHODS THAT ENHANCE THE VALUE OF YOUR RESEARCH PROJECT

Regardless of whether you take a quantitative or qualitative approach, your project should be worthy of your time and effort. More specifically, you must use one or more methods that, in combination, will enable you to obtain meaningful results and draw defensible conclusions. We urge you to look at this issue from two angles: (a) how much others will evaluate your results and conclusions as being *credible* and (b) if appropriate, how much others will believe them to be applicable—*generalizable*—to other circumstances and contexts.

Enhancing the Credibility of Your Findings

A research study has **credibility** to the extent that other people (a) agree that its design and methods are appropriate for the research problem or question, (b) judge its results as being reasonably accurate and trustworthy, and (c) find the researcher's interpretations of the data to be plausible. Quantitative researchers often use the term **internal validity** when evaluating the credibility of their own and others' research findings; we examine this concept more closely in the discussion of experimental research in Chapter 7. Qualitative researchers use a variety of terms: not only *credibility* but also *trustworthiness, confirmability,* and *interpretive rigor.* If you can't obtain results and draw conclusions that other scholars find credible, you're almost certainly wasting your time and effort on what is, for all intents and purposes, a trivial enterprise.

Following are several strategies you might use to enhance the credibility of a study:

- ***Triangulation.*** In **triangulation**, multiple sources of data are collected with the hope that they will all converge to support a particular assertion, hypothesis, theory, or conclusion. This approach is often used 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 address a single research problem.

- ***A controlled laboratory study.*** Some quantitative researchers conduct their studies in laboratory settings where they can carefully control environmental conditions, measurements, and any experimental treatments. Laboratory studies are especially valuable when a research problem involves precisely pinning down cause-and-effect relationships.
- ***Extensive time in the field.*** Researchers who investigate phenomena in their natural, real-world contexts may spend several months—perhaps even a year or more—collecting data, forming tentative hypotheses and developing assertions, and looking for further evidence that either supports or disconfirms those hypotheses and assertions. For example, such might be the case for a biologist who is studying the mating habits of a certain bird species, a social scientist who is trying to shed light on the belief system of a little-understood social or political group, or an archaeologist who is examining or reexamining a particular pre-Columbian site in South America.
- ***Unobtrusive measures.*** On some occasions, people behave in a nontypical manner when they know that they're being observed and studied—a phenomenon known as **reactivity**. One strategy for getting around this problem is to use an **unobtrusive measure**—that is, to gather data about people's behaviors when those people don't know their actions are being recorded. We offer two real-life examples to illustrate. In one study, a university library assessed 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). You should note, however, that ethical issues can sometimes arise when researchers observe people without permission from those individuals; we discuss ethics later in this chapter.
- ***Thick description.*** A researcher who uses **thick description** describes a situation in sufficiently rich, "thick" detail that readers can draw their own conclusions from the data presented. This approach is most commonly used in qualitative research—for example, when a researcher has studied a particular social, cultural, or religious group for a lengthy period of time. In contrast, talking about every data point collected in a quantitative study is usually "TMI"—way, *way* too much information.
- ***Respondent validation.*** In **respondent validation**, a researcher takes conclusions back to the participants in a study and asks quite simply, Do you agree with my conclusions? Do they make sense based on your own experiences? Respondent validation is sometimes used in qualitative research, especially when a researcher wants to gain insights into the nature of people's beliefs, attitudes, or subjective experiences.
- ***Analysis of contradictory data, such as discrepant perspectives or quantitative outliers.*** In this strategy, a researcher actively looks for data that are inconsistent with existing hypotheses or assertions—what qualitative researchers might call *discrepant voices* and quantitative researchers often call *outliers*—and then continually revises hypotheses or assertions until all collected data have been accounted for. This strategy, too, is more commonly used in qualitative research.
- ***Follow-up studies specifically designed to eliminate alternative explanations for findings.*** Quantitative and qualitative researchers alike may realize that their results could reasonably be interpreted in two or more disparate ways, rendering their own conclusions potentially suspect. In such instances, they might conduct one or more follow-up investigations specifically designed to discredit the legitimacy of other explanations. For example, a quantitative researcher might collect data in other conditions or another setting or, instead, conduct a second experiment that has specifically been designed to rule out competing interpretations. Such strategies bring us to our next topic: generalizability.

Enhancing the Generalizability of Your Findings

By **generalizability**, we mean the extent to which results obtained and conclusions drawn can be applied to—or in some other way useful for understanding—other people, situations, or contexts. When referring to this aspect of a study's value, quantitative researchers often use the term **external validity**. Once again, qualitative researchers differ in their terminology, depending on the precise meaning they have in mind; for instance, you might see such terms as *transferability, particularizability, replicability,* or *confirmability.*

Three commonly used strategies for enhancing the generalizability or applicability of a research study are the following:

- ***A naturalistic, real-world setting.*** Earlier we mentioned that researchers sometimes use laboratory experiments to help them control environmental conditions, observations, and experimental treatments. Laboratory studies have a downside, however: They provide an artificial setting that might 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 yield results that have broader applicability to other real-world contexts.[2] Keep in mind, however, that we're really talking about a continuum here: Research studies can have varying degrees of artificiality versus real-world authenticity.[3]
- ***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. (Note that drawing conclusions from only a small sample of category members is *inductive reasoning* at work.) 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 the same kinds of granite found in other locations. The principle might also 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. Many qualitative researchers don't intend to make broad generalizations; hence, they might specifically choose participants based on distinct, unique characteristics they want to learn more about. However, if the goal is to make claims about a wide population, researchers ideally want participants in a research study to be a *representative sample* of that population. In Chapter 6 we describe a number of possible sampling strategies, which vary in the extent to which they enable a researcher to identify and enlist a truly representative sample.
- ***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. Taken together, these studies provide evidence that the conclusion has applicability across diverse situations.

MyLab Education Self-Check 4.3

MyLab Education Application Exercise 4.2: Enhancing Credibility of a Research Study

[2]The artificial nature of laboratory research has been a concern in psychology for many years. In most cases, however, quantitative studies conducted in a laboratory and those conducted in real-world settings lead to similar conclusions about human nature, especially when lab-based studies reveal large differences among treatment groups (e.g., see Anderson, Lindsay, & Bushman, 1999; G. Mitchell, 2012).

[3]When a research study is set in a naturalistic, close-to-reality environment, it is often described as having *ecological validity.*

CHOOSING APPROPRIATE ASSESSMENT STRATEGIES

Many methodologies require one or more **assessment** strategies; that is, researchers must in some way capture and evaluate the nature of important characteristics, behaviors, or other variables under investigation. Appropriate assessment procedures provide a solid basis on which any good research project rests. Just as a building with a weak foundation is unlikely to be safe for habitation, so, too, will a research effort employing questionable assessment techniques provide little of value in addressing the problem or question at hand.

Very often, assessment also involves **measurement**; that is, it imposes a range of *numbers* on an entity being assessed. Although numbers are most commonly associated with quantitative methodologies, some degree of measurement is almost inevitable in qualitative research as well. At a minimum, qualitative researchers are apt to *count* things—perhaps the members of certain groups or the frequencies of certain events. And during data analyses, some of them numerically *code* their observations to reflect various categories into which different observations fall.

For some research problems, you can use one or more existing assessment instruments—perhaps an oscilloscope to measure patterns of sound, a published personality test to assess a person's tendency to be either shy or outgoing, or a rating scale that a previous researcher has developed to assess marital satisfaction. For other research problems, you may have to develop your *own* assessment procedures—perhaps a survey to assess people's opinions about welfare reform, a paper-and-pencil test to assess what students have learned from a particular instructional unit, or a checklist to evaluate the quality of a new product.

Distinguishing Between Substantial and Intangible Phenomena

Researchers often have relatively straightforward strategies for assessing **substantial phenomena**—things that have physical substance, an obvious basis in the physical world. For example, an astronomer might measure the patterns and luminosity of light in the night sky; a neurologist might measure the intensity and location of activity in the brain; a chemist might measure the mass of a compound both before and after transforming it in some way. All of these are reasonable approaches to assessing substantial phenomena. Some devices designed to assess substantial phenomena, such as high-powered telescopes and magnetic resonance imaging (MRI) machines, are highly specialized and used only in particular disciplines. Others, such as clocks, balance scales, and blood pressure monitors, are applicable to many fields of inquiry.

Yet many researchers strive to assess and possibly measure **intangible phenomena**—things that exist only as concepts, ideas, opinions, feelings, or other non-substance-based entities. For example, a researcher might want to assess the economic "health" of business, the degree to which students have "learned," or the extent to which people "value" physical exercise. The researcher might seek to assess these intangibles, not with telescopes or MRI machines, but with the Dow Jones Industrial Average, achievement tests, questionnaires, or interviews.[4]

Regardless of the nature of the phenomena being assessed, assessment strategies typically involve comparisons with certain quantitative or qualitative standards. Standards

[4]In previous editions of this book, we have used the term *insubstantial* rather than *intangible* for this category; however, we have found that some of our readers misinterpret *insubstantial* to mean "unimportant" or "imaginary." You may sometimes see our substantial–intangible distinction referred to as *manifest variables* (which can be directly observed and measured) versus *latent variables* (which lie below the surface and can be assessed only indirectly through their effects on another, observable entity; e.g., see Bartholomew, 2004).

for many quantitative assessments are fairly simple and clear-cut. For instance, *time* can be assessed based on information from a stopwatch or calendar. Variables such as *width* and *weight* can be based on comparison with a metric scale (e.g., centimeters and milligrams) or imperial scale (e.g., inches and pounds). Students' numerical scores on an academic achievement test might be compared with scores of other students across the nation.

Qualitative standards are apt to be more sophisticated. 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 such original sources, it's potentially possible to assess the thoughts and ideas of others and to approximate their similarity to or deviance from those sources.

Assessing Intangible Phenomena: An Example

Assessing intangible phenomena—those that have no obvious, concrete basis in the physical world—can sometimes involve considerable creativity. For example, imagine that we want to assess—and in some respects to measure—the interpersonal dynamics within a small group of people. Let's take a group of nine people who work together in the human resources department of a large corporation. They attend a recognition dinner at an exclusive hotel and 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 4.2.

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 assess it, let alone to measure it.*

One possible approach to assessing the group's interpersonal dynamics is to give each group member a slip of paper on which to write three sets of names, one set each for (a) one or more individuals in 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 4.3. 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

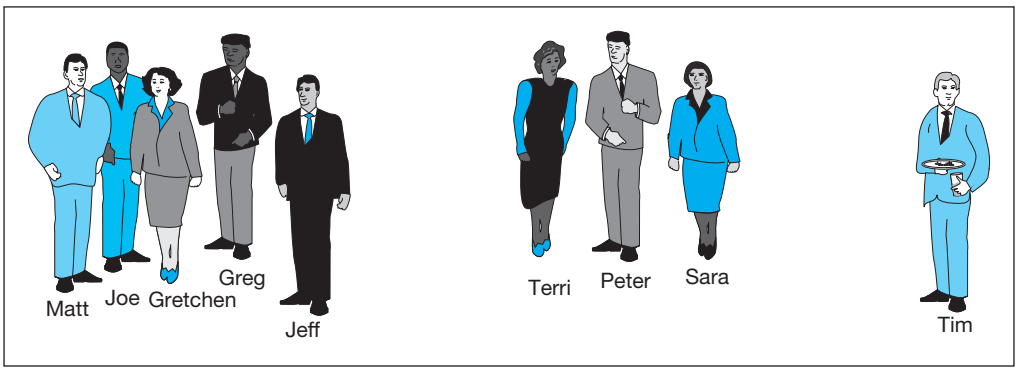
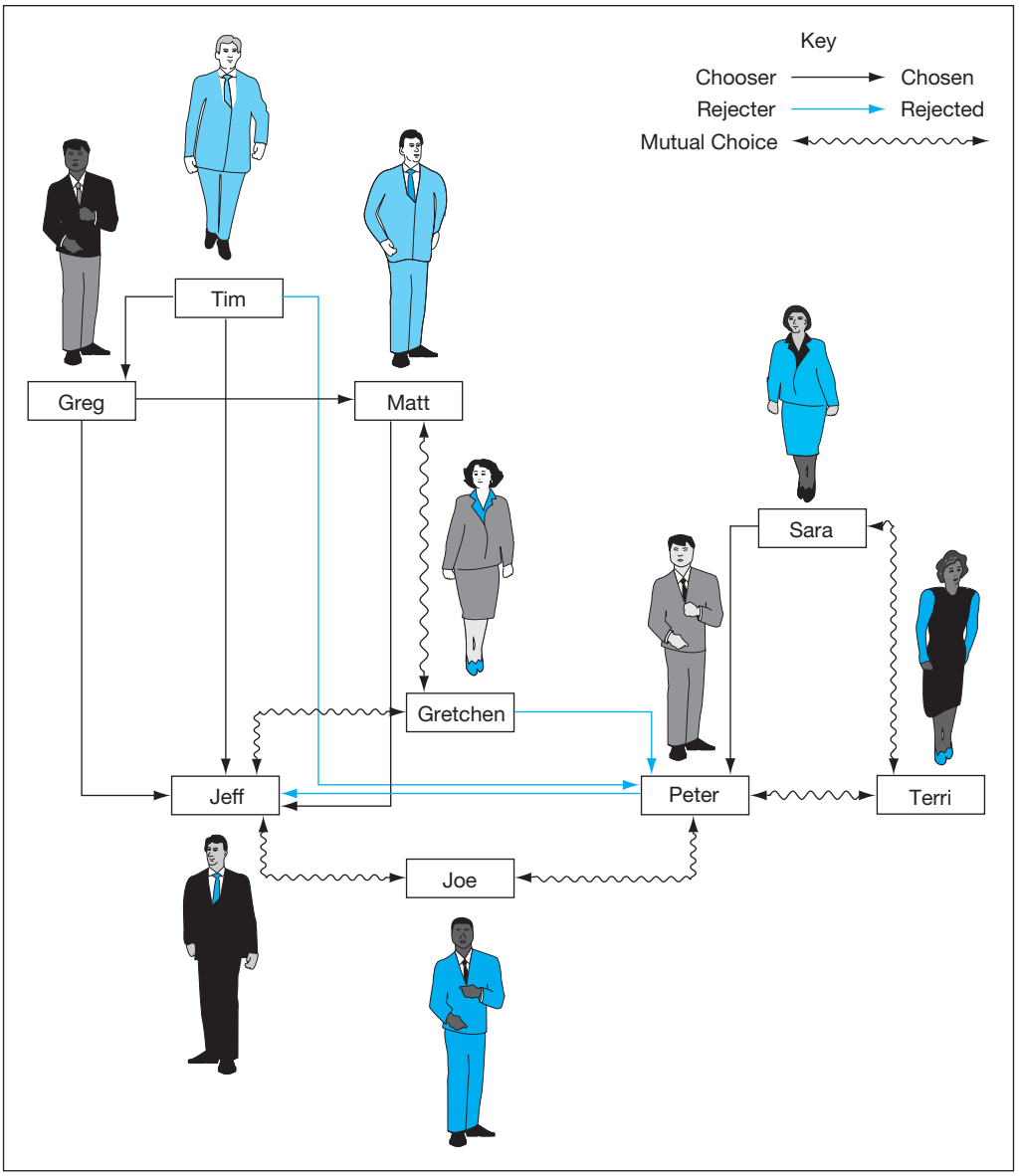


FIGURE 4.2 Conversation Groups in a Hypothetical Human Resources Department

FIGURE 4.3
Sociogram of Interpersonal Dynamics



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 4.3. 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, clusters of facts suggest the following conclusions:

- Jeff seems to be 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 4.3). The sociogram also reveals Jeff's popularity with his colleagues.
- Probably some factions and interpersonal tensions exist within the 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.

TABLE 4.3 Data from Figure 4.3 Presented as a Sociometric Matrix

		How Each Person Was Rated by the Others								
		Gretchen	Joe	Greg	Sara	Peter	Jeff	Tim	Matt	Terri
How Each Person Rated the Others	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

- 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.
- Tim is apparently the social 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've illustrated what it means to interpret data by assessing an intangible phenomenon and analyzing the resulting data. Notice that we didn't just observe the behaviors of nine individuals at a social event; we also looked below the surface to identify possible hidden social forces at play. Our example is a simple one, to be sure. Assessment 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

Whenever we assign a number to a data point, we are, in a sense, "measuring" it. All forms of measurement fall into one of four categories, or **scales**: nominal, ordinal, interval, and ratio (Stevens, 1946). The scale of measurement will ultimately dictate the statistical procedures (if any) that can be used in processing the data.

Nominal Scales

The word *nominal* comes from the Latin *nomen,* meaning "name." In a **nominal scale**, the numbers do nothing more than name, or label, each data point; they say absolutely nothing about the *quantity* of something. For example, if we're studying the behaviors of human beings or certain other animal species, we might label all males as "1" and all females as "2." If we're coding data from a survey about people's preferred flavors of ice cream, we might assign a "1" to vanilla, "2" to chocolate, "3" to strawberry, "4" to chocolate chip, and so on. And if we're comparing the belief systems of various political parties in the United States, we might use 1s for Democrats, 2s for Republicans, 3s for people affiliated with other parties, and 0s for nonaffiliated individuals.

Only a few statistics are appropriate for analyzing nominal data. We can determine the *mode* as an indicator of the most frequently occurring category within our data set; for example, we might determine that the greatest number of people in our sample prefer vanilla ice cream over all other flavors. We can calculate the *percentages* of various subgroups within our total sample; for example, we might determine that of the people in our sample, 39% are Democrats, 42% are Republicans, 5% are affiliated with another party, and 14% are unaffiliated. We can also use a *chi-square test* to compare the relative frequencies of people in various categories; for example, we might discover that more males than females are Republicans but that more females than males are Democrats. (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**, we can think in terms of the symbols > (greater than) and < (less than). Thus, we can compare various pieces of data in terms of one having greater quantity than another. In essence, an ordinal scale allows us to *rank-order* data—hence the term *ordinal.*

As an example, we can roughly assess members of a workforce by grades of proficiency: unskilled, semiskilled, or skilled. Likewise, we can indicate people's level of education by classifying them as being unschooled or having completed an elementary, high school, college, or graduate education. Many high schools rank their graduates in terms of overall achievement levels: one student (the valedictorian) is ranked *first* in the graduating class, another (the salutatorian) is ranked *second,* then *third, fourth, fifth,* and so forth.

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. And we can determine the extent of the relationship between two characteristics by means of Spearman's *rank-order correlation.*

Interval Scales

An **interval scale** 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 doesn't indicate a total absence of heat.[5] For instance, when Gabriel Fahrenheit was developing his Fahrenheit scale, he made his zero point the lowest temperature obtainable with a mixture of salt and ice, and his 100 degrees was what he determined to be human beings' average body temperature. These were purely arbitrary decisions. They placed the freezing point of water at 32° and the boiling point at 212° above zero.

In education, many researchers treat certain scores on large-scale standardized achievement tests as being on interval scales; in fact, some scores on achievement tests are specifically calculated to reflect consistent intervals of quantity.[6] For example, consider three students—Tom, Dick, and Harry—who get scores of 40, 50, and 60, respectively, on a national science achievement test. We could say not only (a) that Tom has learned the least and Harry has learned the most about science but also (b) that the *difference* between Tom's and Dick's achievement levels is the same as the *difference* between Dick's and Harry's. In other words, the intervals between 40 and 50 and between 50 and 60 are equivalent. But now let's take a fourth student, Fred, who gets a score of 0 on the test. Does this mean that Fred knows *absolutely nothing* about

[5]Some readers may know the Celsius scale by its previous name: the *centigrade* scale. The scale is currently named after the Swedish scientist Anders Celsius.

[6]The achievement test scores we describe here are *standard scores,* which are based on the assumption that students' absolute achievement levels are *normally distributed.* We discuss both the normal distribution and standard scores in Chapter 11.

science? Probably not; unless he has lived under a rock since Day 1, he has learned at least a few things about the world around him. In fact, even if his home *has* been the underside of a rock, he probably has some knowledge about his rocky environment.

Interval scales of measurement allow statistical analyses that aren't 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.

Many people who conduct surveys use rating scales to measure certain intangible characteristics, and they often assume that the results such scales yield are interval data. But are they *really* interval data? In some cases they might be, but in other situations they might not. Let's look at an example. 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:

Place an X on the scale below at the point where you would rate the availability of your professor for conferences.

0	10	20	30	40	50	60	70	80	90	100
Never available		Seldom available			Available by appointment only		Generally available			Always available

Notice that the scale includes points ranging from 0 to 100. At five points along the scale are descriptive labels that can help students determine how they should rate their professor's availability. The numbers themselves reflect equal intervals, but the specific ratings that students assign may not. For instance, is the difference between "never available" and "seldom available" equivalent to the difference between "available by appointment only" and "generally available"? Not necessarily: Some students may think of the word *seldom* as being almost as bad as the word *never,* or they might think of "generally available" as being quite a bit better than "available by appointment only." If this is true, then the rating scale is really yielding ordinal rather than interval data.

Ratio Scales

Two commonly used measurement instruments—a thermometer and a tape measure—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 can*not* say that 80°F is twice as warm as 40°F. Why? Because this scale doesn't 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 tape measure, 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's no more desk in either direction beyond those limits. A measurement of "zero" means that there's 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 assessing 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 wouldn't 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 tape

TABLE 4.4 Contrasting the Four Types of Measurement Scales

	Scale	*General Characteristics*	*Examples*	*Statistical Possibilities*
Non-Interval Scales	Nominal scale	Scale that "measures" only in terms of names or designations of discrete units or categories	• Sex/gender • Preferred flavors • Political affiliation	Enables one to determine the mode, percentage values, or chi-square
	Ordinal scale	Scale that measures in terms of such values as "more" or "less," "larger" or "smaller," but without specifying the size of the intervals	• Proficiency level • Education level • Some rating scales	Enables one also to determine the median, percentile rank, and rank correlation
Interval Scales	Interval scale	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	• Fahrenheit and Celsius temperature scales • Some scores on standardized achievement tests • Some rating scales	Enables one also to determine the mean, standard deviation, and product moment correlation; allows one to conduct most inferential statistical analyses
	Ratio scale	Scale that measures in terms of equal intervals and an absolute zero point	• Age • Blood pressure • Annual salary	Enables one also to determine the geometric mean and make proportional comparisons; allows one to conduct virtually any inferential statistical analysis

measure can do that: A meter is a *multiple* (by 100) of a 1-centimeter distance; an inch is one-twelfth (a *fractional part*) of a foot. The ratios are 100:1 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.

Table 4.4 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 assessment methods you've used will support the statistical operations you're contemplating.

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 aren't 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 1 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 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 10 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?
4. A government researcher 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?
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 psychologist is developing an instrument designed to assess 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?

MyLab Education Self-Check 4.4

Validity and Reliability in Assessment

Especially if you're conducting a quantitative study, chances are good that you will use one or more structured instruments or other structured assessment strategies—perhaps tests, questionnaires, checklists, prespecified lists of interview questions, or observer rating scales—that enable you to collect data somewhat systematically. To ensure that your study's findings have credibility in the eyes of others, each of your strategies must have both *validity* and *reliability* for its intended purpose. The validity and reliability of your assessment strategies will influence the extent to which you can (a) legitimately learn something about the phenomenon you are studying, (b) obtain statistical significance in any quantitative data analyses you conduct, and (c) draw meaningful conclusions from your data.

Validity of Assessment Strategies

Earlier in the chapter, we introduced the term *internal validity,* which many quantitative researchers use to describe the extent to which a research study yields credible findings. When talking about specific assessment strategies, however, the term *validity* has a somewhat different, albeit related, meaning. The **validity** of an assessment strategy is the extent to which the strategy yields accurate assessments of the characteristic or phenomenon in question. Certainly no one would question the premise that a tape measure 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 the mercury 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 socioeconomic class? And how well does a sociogram capture the interpersonal dynamics in a group of nine people? Especially when a researcher is assessing *intangible* phenomena—phenomena without a clear-cut basis in the physical world—the researcher's assessment strategies might 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 and consider its validity for this purpose. Some of the labels are quite fuzzy and hard to pin down. 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, instead, 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? There are 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 an assessment 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, an individual 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. But *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 or culturally 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. (This phenomenon is known as the *social desirability effect*—one possible source of *bias* in data collection.We discuss bias in considerable detail in Chapter 6.)

The validity of an assessment strategy can take several different forms, each of which is important in different situations. Here we use four terms traditionally used by quantitative researchers, but in our view, they apply equally well to qualitative research:

- **Face validity** is the extent to which, on the surface, a strategy *appears to yield* an accurate assessment of 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 a strategy is truly assessing what the researcher wants to assess.
- **Content validity** is the extent to which an assessment instrument or procedure adequately reflects the full breadth of the characteristic being assessed. Content validity is often a consideration when a researcher wants to assess people's *achievement* in some area—typically using either a paper-and-pencil instrument or a predetermined set of

hands-on performance tasks. When designed to assess some form of achievement, an assessment tool has high content validity if its items, questions, or tasks 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 strategy correlate with assessments of another, presumably related characteristic (this other characteristic is the *criterion*). For example, a personality test designed to assess a person's shyness or outgoingness has criterion validity if it yields scores similar to other assessments of a person's general sociability. A strategy 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. If the criterion is assessed at a subsequent point in time, criterion validity is also known as *predictive validity*.[7]
- **Construct validity** is the extent to which an assessment strategy yields credible results regarding a characteristic that cannot be directly observed but is assumed to exist based on patterns in people's behaviors or creations (such a characteristic is a *construct*). Motivation, creativity, racial prejudice, happiness—all of these are constructs, in that none of them can be *directly* observed and assessed. When researchers observe behaviors, ask questions, present tasks, or evaluate humanmade products as a way of assessing an underlying construct, they should obtain some kind of evidence that their approach does, in fact, assess the construct in question.

Sometimes there is universal agreement that a particular assessment tool or strategy yields a valid assessment of a characteristic; such is the case for balance scales, barometers, and oscilloscopes. But whenever there is *not* such widespread agreement, researchers should provide evidence that the strategies they are using have validity for their purposes.

The validity of any assessment strategy can vary considerably depending on the purpose for which it is being used. In other words, *the validity of a strategy 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 estimate 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 relationships.

Determining and Potentially Enhancing the Validity of an Assessment Strategy

In one way or another, you must convince other scholars that any assessment strategies you use have validity for your purposes. Sometimes previous researchers have already done this for you, in which case you can simply cite their work in your literature review or methods section. But in other situations you may have to make your own case that your assessment strategies have merit. Here we offer several approaches that researchers sometimes use to justify their use of particular assessment strategies:

- ***When creating a structured paper-and-pencil or performance-based assessment tool, take steps to be sure that it adequately represents the breadth and depth of the characteristic being assessed.*** Especially when *content validity* is a primary concern, the assessment tool should reflect the full scope and complexity of the domain or other characteristic being assessed. For example, to create a viable achievement test for a particular content domain, educational researchers often begin by constructing a two-dimensional grid, or **table of specifications**, that lists the specific

[7]Our description of criterion validity here might initially strike you as being similar to the term *triangulation.* However, triangulation involves pulling together two or more different sets of data to determine whether they lead to similar conclusions. In contrast, the term *criterion validity* is used in reference to a single assessment instrument or procedure.

topics and behaviors (e.g., factual knowledge, problem-solving skills, critical thinking) that reflect achievement in the domain. In each cell of the grid, the researcher indicates the relative importance of each topic–behavior combination. The researcher then develops a series of tasks or test items that reflects the various topics and behaviors in appropriate proportions.

- ***Take a multitrait–multimethod approach.*** In a **multitrait–multimethod approach**, two or more different characteristics are each assessed using two or more different techniques (e.g., see Campbell & Fiske, 1959; Campbell & Russo, 2001). Varying assessments of a single characteristic should be highly correlated; meanwhile, a single way of assessing two or more *different* characteristics should *not* be highly correlated. For example, in educational research, a quantitative researcher might measure the constructs *academic motivation* and *social motivation* using both self-report questionnaires and teacher observation checklists. Statistical analyses should reveal that the two measures of academic 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.

  The multitrait–multimethod approach is most commonly associated with quantitative methodologies. Yet a multitrait–multimethod approach is often used in qualitative research as well, in which case a researcher might simply describe it as *triangulation.*
- ***Strive for consistency with a particular conceptual framework.*** As previously noted in Chapter 2, a *conceptual framework* is a proposed network of interconnected concepts that might underlie the basic nature of a phenomenon being studied. Tying an assessment strategy to a particular conceptual framework can be particularly helpful when the strategy is designed to obtain qualitative data that don't easily lend themselves to sophisticated statistical analyses.
- ***Conduct one or more pilot tests of your assessment strategy.*** Unless you are using a widely used and validated paper-and-pencil instrument, it's often a good idea to try out your strategy in advance with people, animals, artifacts, or whatever other entities you will be assessing. In doing so, you can carefully scrutinize your strategy for obvious or possible weaknesses and then modify it as needed.
- ***Enlist an expert panel of judges.*** Even experienced researchers sometimes ask two or more experts in a particular field to scrutinize an assessment instrument or other assessment strategy and give informed opinions about its validity for assessing the quantity or quality of a characteristic under investigation. Experts can also offer recommendations for improving a strategy—perhaps by suggesting additional criteria to include in a checklist or recommending additional questions or alternative wordings to use in an interview or questionnaire.

Although none of the approaches just described guarantees the validity of an assessment strategy, each one increases the likelihood of such validity, and in combination, they can be even more powerful.

Reliability of Assessment Strategies

Imagine that you're 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 aren't 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 degree to which an assessment strategy consistently yields very similar results when the entity being assessed hasn't changed. As we've just seen in our waist-measuring situation, even strategies that assess concrete physical characteristics aren't necessarily completely reliable. As another example, think of a measuring cup that a baker might use while making cakes. When measuring a half cup of flour, the baker won't always measure *exactly* the same amount of flour each time.

Strategies designed to assess social and psychological characteristics (intangible 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 assessing 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, on two separate occasions, we were to ask the recognition-dinner attendees portrayed in Figure 4.2 to identify the colleagues they like best and least, they wouldn't necessarily give us the same responses on each occasion, even if the interpersonal dynamics within the group have remained constant.

As is true for validity, reliability can take different forms that are more or less relevant to different research projects. Following are four forms that are frequently of concern:

- **Interrater reliability** (also known as *interrater agreement*) 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 assessment instrument yields the same results for the same people over a short time interval.
- **Equivalent-forms reliability** is the extent to which two different versions of the same assessment instrument (e.g., "Form A" and "Form B" of an achievement test) yield similar results.
- **Internal consistency reliability** is the extent to which all of the items or other tasks within a single assessment instrument yield similar results.

You must understand two important points about reliability. First, *an assessment strategy can be valid only to the extent that it is also reliable.* In other words, we can assess something accurately only when we can also assess it with some degree of consistency. Hence, by increasing the reliability of a measurement instrument, we might also increase its validity.

Yet even if we enhance the reliability of an assessment strategy, we don't necessarily increase its 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're apt to get similar measures of an individual's head circumference on different occasions—but absolutely no validity. As noted earlier, head size is *not* a good indication of intelligence level.

Mathematically Determining the Reliability of an Assessment Strategy Any determination of reliability requires the use of numbers. For example, in a research project that requires the collection of data by means of direct observations, two different research assistants might independently assign every observation to a particular category on a checklist (reflecting a nominal scale) or evaluate various characteristics or behaviors with regard to a particular rating scale (reflecting an ordinal or interval scale). The researcher would then calculate the percentage of agreement between the two assistants. A 98% agreement rate would indicate high reliability; a 50% rate would be quite worrisome.

In quantitative research with human participants (or other suitable entities), a more complex statistical analysis is typically conducted. A quantitative approach 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 assessment instrument to individuals at two different points in time—perhaps a day, a week, or a month apart (test–retest reliability)[8]
c. Giving each individual two parallel versions of the same assessment instrument (equivalent forms reliability)
d. Administering only one assessment 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 assessment results tend to be similar for each individual (see Chapter 11 for a discussion of correlation coefficients)

Enhancing the Reliability of an Assessment Strategy We offer three possible strategies for enhancing the reliability—and thus indirectly also enhancing the validity—of an assessment strategy:

- ***Identify specific, concrete criteria to be used in categorizing or rating a characteristic, behavior, or other variable.*** This strategy is critical whenever subjective judgments are required. For example, a researcher who wants to document the frequency of aggressive behaviors among high school students would need to be quite specific about what an aggressive behavior *is.* Some behaviors might be categorized as *physical* aggression and more specifically as pushing, hitting, poking, or in some other way having the potential to cause bodily harm. Other behaviors might instead be categorized as *psychological* aggression—that is, causing mental anguish or reducing self-esteem, perhaps through ridicule, ostracism, demeaning rumors, or cyberbullying. Ideally, descriptions of each category and subcategory would include concrete, illustrative examples.
- ***When an assessment involves a structured paper-and-pencil or performance-based instrument, use the instrument in a consistent manner for all participants or other entities under investigation.*** Quantitative researchers call this strategy **standardization**. More specifically, standardization involves using similar content, format, administration procedures, and evaluation criteria during the assessment. There's an important exception here: When participants have disabilities, assessment procedures may need to be tailored to accommodate individual strengths and limitations. For example, participants who are visually impaired may not be able to respond to a written questionnaire, participants who are deaf may have trouble with interviews, and participants with brain injuries or other physical disabilities may require considerable time to show what they can do. To completely standardize procedures for such individuals could seriously jeopardize the validity of the results obtained.
- ***Explicitly train raters to apply criteria in a similar manner.*** This strategy, too, is especially important when subjective judgments are required. Any single rater might have biases that slant ratings or categorizations in certain directions; for instance, in a study of aggression in high school students, a research assistant might dismiss an instance of psychological aggression as being nothing more than "playful teasing" and thus not worthy of recording. Having two or more raters independently use an assessment strategy and then training them sufficiently that they are consistent in their assessments enhances both the validity and reliability of the strategy. Revision of the strategy itself to *enhance* raters' consistency may also be in order.

[8]In recent years, technology has enabled relatively easy collection of *multiple* measures of a single characteristic over a short period of time—an approach that can potentially increase both the test–retest reliability and the validity of the assessment. The section on *experience-sampling methods* in Chapter 6 provides more details.

Thinking of Validity and Reliability as Reflections of Error in Your Assessments

Very few assessment strategies are perfect, in the sense that they capture the quality or quantity of a variable with complete accuracy. In other words, the results that most strategies yield will have some degree of *error* in them. In many cases—and especially when we're studying intangible phenomena—a strategy may allow us to assess a characteristic only indirectly and thus may be subject to a variety of biasing factors (e.g., people's responses on a rating scale might be influenced by their interpretations, prejudices, reading skills, memory lapses, etc.). Such sources of error negatively impact the *validity* of the assessment strategy. Yet, as pointed out earlier, even when we're assessing substantial phenomena, we may get slightly different results from one time to the next simply because our assessment tool is imprecise (e.g., the waist or head size we measure may depend on how snugly we pull the tape measure). Such sources of error impact the *reliability* of the assessment. Generally speaking, validity errors reflect biases in the assessment strategy itself and are relatively constant sources of error. In contrast, reliability errors reflect *use* of the strategy and are apt to vary unpredictably from one occasion to the next.

Especially for quantitative research, we cannot overemphasize the importance of determining and maximizing the validity and reliability of any assessment instruments and other assessment strategies. Without reasonably valid and reliable indicators of quantitatively assessed characteristics and phenomena, researchers cannot possibly obtain informative and useful data for addressing and solving their research problems or questions.

As you plan your research project, you should clearly identify the nature of the assessment strategies you will use and carefully examine them with respect to their potential validity and reliability. Furthermore, in your research proposal and final research report, you should describe any instrument or other assessment strategy in explicit, concrete terms. For example, if you're 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're assessing some aspect of human thought or behavior, you should describe the questions asked or tasks administered, the overall length of the assessment, and the method of evaluating responses.

CONCEPTUAL ANALYSIS EXERCISE Identifying Problems with Validity and Reliability in Assessment Strategies

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

- Face validity
- Content validity
- Criterion validity
- Construct validity
- Interrater reliability
- Test–retest reliability
- Equivalent-forms reliability
- 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 5 points higher than students who answered either of the other two 40-question sets.
3. In order to determine what kinds of situations prompt cooperative behavior 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 "mutual collaboration" or, instead, reflect "dependence" of one gorilla on another.
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 ratings don't seem to have any correlation with the blood-test results.
5. In a 2-week period during the semester, a researcher gains entry into several college classrooms in order to administer a short questionnaire regarding college students' beliefs about climate change. The instrument 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 questionnaires; as a result, the researcher later discovers that a few students were in two of the classes surveyed and thus responded to the questionnaire twice. Curiously, however, these students have sometimes given different responses to particular statements on the two different occasions, and hence their two overall scores are also different.
6. In order to get a sense of how harmonious most long-term marriages are, a researcher administers a questionnaire to 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 harmonious marriage (e.g., "We usually agree about how to spend our money," "We split household chores in a manner that suits us both"), 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. Thus, if one were to look only at responses to the first 30 items, one would conclude that most long-term couples live in relative peace and harmony. But if one were to look only at responses to the last 30 items, one would think that married couples quarrel a lot. *Note:* We recommend that questionnaires *not* be slanted in a one-way direction (e.g., with all items characterizing a harmonious marriage and none characterizing a *dis*harmonious one); see the "Constructing a Questionnaire" guidelines in Chapter 6.
7. College students in an introductory 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 who have chosen 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 task is being used as 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."

8. A researcher constructs and uses a questionnaire intended to assess the extent to which older adults display open-minded attitudes about and tolerance toward a particular religious group. However, several experts in the researcher's field of study suggest that the questionnaire assesses not how open-minded and tolerant people actually *are* regarding that religious group but, instead, what people would like to *believe* or *convey* about their own open-mindedness and tolerance.

MyLab Education Self-Check 4.5

ETHICAL ISSUES IN RESEARCH

In 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—during the *planning* stage—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 institutional 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 nonhuman 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 especially vulnerable populations. 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 to engage in some kinds of activities. Special care must be taken with participants who can't easily advocate for their own needs and desires—such as children, elderly individuals, and people with significant physical or mental disabilities.

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 disguised form of coercion (Scott-Jones, 2000).[9]

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 that 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—such as birth certificates, newspaper articles, Internet websites, and public "tweets" on Twitter—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.*

Most 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 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 might 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

[9]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 example, reading and writing a review of a research article (Scott-Jones, 2000).

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 6-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 6 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 2 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-xxx-xxxx) 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-xxx-xxxx) at any time.

Participant's signature: _____ Date: _____

Researcher's signature: _____ Date: _____

FIGURE 4.4 Example of an Informed Consent Form (actual phone numbers have been removed)

Source: 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, CO. Adapted with permission.

- 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 (1988) in a research project for her doctoral dissertation, is presented in Figure 4.4. 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 regarding *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. 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 want 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 be meaningfully 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 institutional 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 assessing 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 aren't being asked to do something they ordinarily wouldn't do—and if they aren't 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 used different sections of a library and the frequency with which people hiked along certain trails in a national park. Both of these examples involved behaviors within the scope of participants' normal activities.

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, the person should be given a pseudonym—and other trivial, irrelevant details that might give away the person's identity should be changed—in order to protect that person's privacy.

In this age of the Internet, researchers must also take precautions to ensure that computer hackers cannot access participants' individual data. Our advice here is simple: Don't communicate 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 information about how to decode the attachments 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. Knowledge about participants' individual performances should be revealed *only* to any co-researchers who have a significant role in the research investigation unless participants have specifically granted permission *in writing* that it can 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

Unless complete objectivity can be assured, researchers should acknowledge any personal beliefs, attitudes, and theoretical orientations that might bias their observations, assessments, and interpretations. Researchers must also 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. Plain and simple, such an action constitutes scientific fraud.

Within this context, we ask you to recall the 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—to be blunt—makes you a thief. 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 and unethical. Honest researchers don't hesitate to acknowledge their indebtedness to others.

Institutional Review Boards

Historically, some researchers had few (if any) scruples about the harm they inflicted on certain people or animals. Among the most notorious were German doctors who conducted horrific experiments on concentration camp prisoners during World War II—experiments that sometimes resulted in death or permanent disabilities. Other researchers, too, have exposed people or animals to conditions that created significant physical or psychological harm, with virtually no oversight by more ethical colleagues. Fortunately, safeguards are now in place in many countries to keep inappropriate experimentation in check.

In the United States, in Canada, and among members of the European Union, any college, university, or research institution must have an **institutional review board (IRB)**[10] 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 projects to be sure that the procedures are not unduly harmful to participants, that participants' privacy and anonymity are ensured, and that appropriate procedures will be followed to obtain participants' informed consent. Typically, an IRB will require documentation that a research project includes getting *written* consent from participants. In some situations, however, the board will grant a *waiver* of such documentation, especially if a study poses minimal, if any, risk to participants. Waivers are sometimes appropriate for online surveys (see Chapter 6). Waivers may also be granted if (a) the only risk to participants is a possible breach of confidentiality and (b) a signed consent form would increase this risk. Thus, waivers may be appropriate for minimal-risk studies with particularly vulnerable populations, such as undocumented immigrants or victims of domestic violence, who may be reticent to sign a consent form linking them with the study. If the IRB grants a waiver in these studies, participants would still be given informed consent forms, but they wouldn't be required to sign the forms.

It is important to note that the research study is reviewed at the proposal stage. *A proposal must be submitted to and approved by the IRB before any data are 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 (a) quickly declare it *exempt from review,* (b) give it an *expedited review,* or (c) bring it before the board for a *full review.* In any case, the researcher cannot begin the study until either the board has given its seal of approval or 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

[10] Some institutions use a different label (e.g., *Internal Review Board, Committee for Protection of Human Subjects*).

Tufts University (tnemcirb.tufts.edu), the University of South Carolina (orc.research.sc.edu), or the University of Texas (utexas.edu/research/rsc). You can find other helpful sites on the Internet by using a search engine (e.g., Google or Bing) and such keywords as *IRB, human participants,* and *human subjects*.

Universities and other research institutions have review boards for animal research as well. Any research project 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 that it's possible to do so. For examples of research institutions' IACUC policies and procedures, we refer you to the University of Maryland (umresearch.umd.edu/IACUC) and the University of Arizona (orcr.arizona.edu).

Many novice researchers view IRB and IACUC reviews as a bothersome hurdle to jump in their efforts to carry out a successful research project. We authors can assure you that members of these boards want to encourage and support research—*not* impede it—and typically work hard to make their proposal reviews as quick and painless as possible. Also, they can provide helpful advice to ensure that a proposed study doesn't needlessly jeopardize the welfare of the study's participants.

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 the Internet. Following are examples of organizational websites with ethical codes related to research in their disciplines:

American Anthropological Association (aaanet.org)
American Association for Public Opinion Research (aapor.org)
American Counseling Association (counseling.org)
American Educational Research Association (aera.net)
American Psychological Association (apa.org)
American Sociological Association (asanet.org)
Society for Conservation Biology (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 (a) attended to the nature and availability of the data you require; (b) decided whether a quantitative, qualitative, or mixed-methods methodology might be most suitable for your research problem or question; (c) possibly identified valid, reliable ways of assessing characteristics, behaviors, or other phenomena central to your problem or question; and (d) 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 you must definitely be realistic and practical regarding what you can accomplish in a single project. 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 practicality. Completing the following checklist can help you wisely plan and accurately evaluate the research project 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 aren't 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. If relevant, how available are any already-existing data you need?
- _____ Readily available
- _____ Available with permission
- _____ Available with great difficulty or rarely available
- _____ Unavailable

_____ 7. Also, if you plan to use existing data, 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 you collect any new data? If so, how do you plan to gather it?
_____ Observation _____ Questionnaire _____ Test _____ Rating scale
_____ Photocopying of records _____ Interview (with or without audio recording)
_____ Specialized machine/device _____ Computer technology
_____ Other (explain): _____

_____ 9. Is special equipment or are special conditions necessary for gathering or processing the data? _____ Yes _____ No
If your answer is yes, specify: _____

_____ 10. 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?

_____ 11. What is the estimated cost in time and money to gather the data?

_____ 12. What evidence do you have that the data you gather will be valid and reliable indicators of the phenomena you wish to study?

Overall Assessment

_____ 13. 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 difficulty answering some of the questions in the checklist? For instance, did you have trouble estimating how much time it would take you to gather your data? Did you realize that you might need to create your own questionnaire, test, or other assessment 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 and refine particular procedures, assessment 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 address your overall research problem or question.

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, acquiring existing assessment instruments or developing new assessment strategies, setting up any experimental interventions you have planned, collecting data, analyzing and interpreting the data, 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've 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. "Such a waste!" we thought then . . . and continue to think now.

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 listed in the preceding paragraph. If you think you can accomplish all these things within 2 or 3 months, you're 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.

The most effective strategy we can suggest here is to *develop a research and writing schedule and try to stick to it.* Figure 4.5 provides a workable format for your schedule. In the left-hand column, list all the specific tasks you need to accomplish for your research project (writing a 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 final column to check off each step as you complete it.

Using Project Management Software and Electronic Planners

Project management software is available both commercially (e.g., Milestone Planner, ToDoList) and as freeware that can be downloaded from the Internet (e.g., Freedcamp, GanttProject). You can use such software to organize and coordinate the various aspects of a research project. For example, the software will let you outline 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 school principals. Coordinating the efforts of the many observers, teachers, students, and administrators is a complex task that can be easily laid out and scheduled with project management software.

Task to Complete	Estimated Amount of Time Needed	Target Date for Completion	Task Completed (indicate with a ✓)

FIGURE 4.5 Establishing a Schedule for Your Project

You might consider electronically organizing your schedule even if you don't expect your research project to be as multifaceted as the one just described. For example, you might use the *calendar* application that comes with your laptop or smartphone, or you might download day-planning freeware from the Internet (e.g., My Daily Planner, Planner Pro). With such applications 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 authors have found that a concrete schedule goes a long way toward helping us complete a seemingly humongous task. In fact, this is exactly the approach we have taken when we have written 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.

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 resolving your research problem.

MyLab Education Self-Check 4.6

MyLab Education Application Exercise 4.3: Identifying Potential Risks and Research Plans

APPLICATION EXERCISES

Similar exercises are available on Pearson MyLab Education.

1. You are interested in finding out whether business-run childcare or independent childcare better prepares preschoolers for kindergarten. You choose only 5-year-old children entering kindergarten who have been in the same childcare arrangement for at least 6 months to participate. You decide to measure the preparedness for kindergarten by using standard kindergarten entrance exam questions and kindergarten CTBS (a standardized test) scores given during the spring of the school year.

   i. What is the research topic or question?
   ii. Is this a qualitative or quantitative approach to the research question?
   iii. On reflection, you decide that you would also like to know if there are differences in the social skills and behavior of children in business-run childcare compared with independent childcare, but you are uncertain about exactly what skills and behavior you might measure. Describe the approach you would take to address this problem.
   iv. In what important ways would the data collected during the first part of your study differ from the data collected during the later part of your study?

2. Dr. Smithson conducts a study comparing sleeping habits and work productivity of participants. He finds that participants who wake up no later than 6:00 a.m. are more productive workers than participants who wake up later than 6:00 a.m. In his conclusions, he asserts that people are more likely to be good workers if they wake up early rather than later.

   i. Did Dr. Smithson's study have internal validity?
   ii. He decides to do a second study. How can he ensure that his second study has high internal validity? Evaluate at least three possible options, and explain your reasoning.
   iii. Imagine that Dr. Smithson instead wants to examine the perspectives of this group of participants on work-life balance and the sorts of resources and supports that enhance work productivity, and he designs and conducts a qualitative study. What approaches could he employ to enhance the credibility of the study?

3. The internal review board (IRB) is charged with identifying potential risks to participants; these are often risks that are unnoticed by the researchers as they plan a study. Identify concerns that the board might have regarding each of the following examples. Keep in mind that research risks include

physical, mental, and social risks beyond those found in typical daily activities.

i. Samuel wants to investigate how children respond to negative feedback. In his design, children complete difficult puzzles. Some of the children are informed erroneously that they are worse than their peers at solving these problems. Others are told that they perform better than their peers.

ii. Denise studies worker burnout in a CPA firm. She proposes to study workers firsthand by posing as a new clerical employee and working in an office facing many difficult issues for a year. Only in this way can she really appreciate what it is like for workers "in the trenches," and she feels she can get honest answers from workers with burnout.

iii. As a kinesiologist interested in the relation between physical activity and academic performance, Danny proposes to limit the exercise of one group of children in his study while allowing the other group to exercise as they normally would.

iv. Jane is a college professor interested in students' beliefs about the typical instructional strategies of professors. As part of her ongoing research agenda, she requires students in her research methods course to fill out questionnaires.

FOR FURTHER READING

Planning Your Research Design

Bordens, K. S., & Abbott, B. B. (2014). *Research design and methods: A process approach* (9th ed.). New York, NY: McGraw-Hill.

Burns, A. C., Veeck, A., & Bush, R. F. (2017). *Marketing research* (8th ed.). Hoboken, NJ: Pearson.

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. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Los Angeles, CA: Sage.

Ercikan, K., & Roth, W.-M. (2006). What good is polarizing research into qualitative and quantitative? *Educational Researcher, 35*(5), 14–23.

Firestone, W. A. (1987). Meaning in method: The rhetoric of quantitative and qualitative research. *Educational Researcher, 16*(7), 16–21.

Gergen, K. J., Josselson, R., & Freeman, M. (2015). The promises of qualitative inquiry. *American Psychologist, 70,* 1–9.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14–26.

Maxfield, M. G., & Babbie, E. R. (2015). *Research methods for criminal justice and criminology* (7th ed.). Stamford, CT: Cengage Learning.

Maxwell, J. A. (2010). Using numbers in qualitative research. *Qualitative Inquiry, 16,* 475–482.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Los Angeles, CA: Sage.

Neuman, W. L. (2011). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Hoboken, NJ: Pearson.

O'Cathain, A. (2010). Assessing the quality of mixed methods research: Toward a comprehensive framework. In A. Tashakkori & C. Teddlie (Eds.), *Mixed methods in social & behavioral research* (2nd ed., pp. 531–555). Los Angeles, CA: Sage.

Punch, K. F. (2014). *Introduction to social research: Quantitative and qualitative approaches* (3rd ed.). Los Angeles, CA: Sage.

Sahu, P. K. (2013). *Research methodology: A guide for researchers in agriculture science, social science and other related fields.* New Delhi, India: Springer.

Tashakkori, A., & Teddlie, C. (Eds.). (2010). *Sage handbook of mixed methods in social and behavioral research* (2nd ed.). Los Angeles, CA: Sage.

Thomas, J. R., Nelson, J. C., & Silverman, S. J. (2015). *Research methods in physical activity* (7th ed.). Champaign, IL: Human Kinetics.

Vogt, W. P., Gardner, D. C., & Haeffele, L. M. (2012). *When to use what research design.* New York, NY: Guilford Press.

Wood, M. J., & Ross-Kerr, J. C. (2011). *Basic steps in planning nursing research: From question to proposal* (7th ed.). Sudbury, MA: Jones & Bartlett.

Measurement

Campbell, D. T., & Russo, M. J. (2001). *Social measurement.* Los Angeles, CA: Sage.

Earickson, R., & Harlin, J. (1994). *Geographic measurement and quantitative analysis.* Upper Saddle River, NJ: Prentice Hall.

Fried, H. O., Knox Lovell, C. A., & Schmidt, S. S. (Eds.). (2008). *The measurement of productive efficiency and productivity growth.* New York, NY: Oxford University Press.

Miller, D. C., & Salkind, N. J. (2002). *Handbook of research design and social measurement* (6th ed.). Los Angeles, CA: Sage.

Thorndike, R. M., & Thorndike-Christ, T. (2010). *Measurement and evaluation in psychology and education* (8th ed.). Columbus, OH: Pearson.

Ethics

Bankowski, Z., & Levine, R. J. (Eds.). (1993). *Ethics and research on human subjects: International guidelines.* Albany, NY: World Health Organization.

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). Los Angeles, CA: Sage.

Comstock, G. (2012). *Research ethics: A philosophical guide to the responsible conduct of research.* Cambridge, England: Cambridge University Press.

Eiserman, W. C., & Behl, D. (1992). Research participation: Benefits and considerations for the special educator. *Teaching Exceptional Children, 24,* 12–15.

Erwin, E., Gendin, S., & Kleiman, L. (Eds.). (1994). *Ethical issues in scientific research: An anthology.* New York, NY: Garland.

Garrett, J. R. (Ed.). (2012). *The ethics of animal research: Exploring the controversy.* Cambridge, MA: MIT Press.

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.

Israel, M. (2015). *Research ethics and integrity for social scientists* (2nd ed.). Los Angeles, CA: Sage.

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.

Macrina, F. L. (2014). *Scientific integrity: Text and cases in responsible conduct of research* (4th ed.). Washington, DC: American Society for Microbiology.

Mazur, D. J. (2007). *Evaluating the science and ethics of research on humans: A guide for IRB members.* Baltimore, MD: Johns Hopkins University Press.

Mertens, D. M., & Ginsberg, P. (Eds.). (2008). *The handbook of social research ethics.* Los Angeles, CA: Sage.

National Bioethics Advisory Commission. (2015). *Ethical and policy issues in research involving human participants.* Bethesda, MD: Author.

Panter, A. T., & Sterba, S. K. (Eds.). (2012). *Handbook of ethics in quantitative methodology.* New York, NY: Routledge.

Rhodes, C. S., & Weiss, K. J. (Eds.). (2013). *Ethical issues in literacy research.* New York, NY: Routledge.

Roberts, L. W. (2006). Ethical principles and practices for research involving human participants with mental illness. *Psychiatric Services, 57,* 552–557.

Sieber, J. E., & Tolich, M. B. (2013). *Planning ethically responsible research* (2nd ed.). Los Angeles, 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":

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 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.
4. This is an *ordinal* scale that reflects varying levels of quality. There's no indication that the four categories each reflect the same range of quality, and a true zero point (no road at all) isn't represented by the categorization scheme.
5. Don't be misled 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 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 Assessment Strategies":

1. The test lacks *content validity:* It doesn't reflect the content domain that instruction has focused on—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 short time interval (2 weeks at most), even though their general opinions about climate change may have 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 assess a single characteristic: matrimonial harmony.
7. *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 assessment technique doesn't *appear* to be a good measure. Although face validity isn't "true" validity, a lack of it can sometimes negatively impact participants' cooperation in a research project.
8. In this case the questionnaire's *construct validity* is suspect. Open-mindedness and religious tolerance are hypothesized internal characteristics that can be inferred and assessed only indirectly through observable patterns in people's behaviors. Here the behaviors being observed are merely responses to a questionnaire; such responses might not be good indicators of how the study's participants actually think about and behave toward members of the religious group in question. (A *social desirability effect* may be at work here; once again we refer you to Chapter 6 for a discussion of this concept.)

This page intentionally left blank

Chapter 5 Writing the Research Proposal

Researchers typically communicate their thoughts, objectives, plans, and methods for others to read, discuss, and react to. The formal mechanism that initiates such a dialogue is the research proposal. As a point of departure, it must be a precise document from the first word to the last.

Learning Outcomes

5.1 Describe three general characteristics of a good research proposal, and identify strategies for organizing and writing the first draft of a proposal.

5.2 Describe common weaknesses in research proposals.

5.3 Identify strategies for effectively revising and strengthening your proposal.

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, described, inspected, and, in almost every instance, approved by others. Graduate students conducting research for a thesis or dissertation must get the approval of an advisor and, in the case of a dissertation, a doctoral committee. Researchers seeking grant funding must get approval from the university or organization for which they work, and the project must be deemed worthy of funding by the grant-awarding agency. Researchers who plan to work with human subjects must get the approval of an institutional review board (IRB), and those who plan to work with animals must get approval from an institutional animal care and use committee (IACUC) (see Chapter 4). Such approvals are usually secured through the submission of a *research proposal.* The proposal lays out the problem or question at the core of the proposed project, describes exactly how the research will be conducted, and outlines in precise detail the resources and strategies the researcher will use to achieve the desired results.

A proposal is as essential to successful research as a building contractor's plans are to the construction of a new home. No one would start building a house by digging a hole for the foundation. Before one turns a shovelful of earth, many questions must be addressed, many decisions made. Will the house be two stories, 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 roof be pitched and shingled? What kind of heating system will be installed? 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 contractor needs a floor plan of the entire structure, showing to the inch exactly where each room, door, window, and closet will be located. The contractor also needs a series of elevations of the proposed structure, showing each wall to scale as it will appear when completed. Finally, the contractor requires a set of specifications for the building, indicating exactly what lumber will 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—including codes regarding plumbing, wiring, and distance from the street. A permit provides a means of ensuring that a new building meets these codes. The point is this: *Permission is essential to the plan.*

Like the contractor who presents detailed plans for a building, the researcher develops a written proposal for a research project. In this proposal, the general problem or question is clearly stated, often in conjunction with a purpose statement. Furthermore, any subproblems and *a priori* hypotheses are presented, all key concepts are defined, delimitations and possible limitations are carefully spelled out, and the potential significance of the proposed study—why it is important—is explained. The researcher then specifies all anticipated details of acquiring, organizing, analyzing, and interpreting the data. The researcher sets forth the resources at hand for carrying out the research, including his or her qualifications (and those of any assistants), the means by which the needed data will be obtained, and any essential equipment and facilities. Nothing is overlooked. All questions that might arise in the minds of those who review the proposal are anticipated and answered. Any unresolved matter is a weakness in the proposal and can jeopardize its approval.

Sometimes novice researchers think that a proposal is merely a necessary formality and thus don't give it the serious consideration it deserves. They try to describe their 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're seeking financial support for a project from a funding agency or seeking approval for a thesis or dissertation from a faculty committee, a clear, well-written proposal is essential. Nothing else can substitute for an explicit description of both problem and procedure.

CHARACTERISTICS OF A PROPOSAL

Good research requires that those who undertake it be able to think clearly, without confusion. The proposal will demonstrate—for better or for worse—whether the author has this ability. When reviewers receive a proposal that is unfocused, poorly organized, and lacking in critical details, they're apt to get 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. More often than not, the perceived qualifications of a researcher rest 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 shouldn't be cluttered with extraneous and irrelevant material. Whatever doesn't contribute directly to the delineation of the problem and its solution must be omitted. Anything else is a distraction. Remember the building contractor's drawings: clean, clear, and economical. They contain all the information that's necessary, not one detail more.

A straightforward statement of the specific problem or question to be investigated should be presented within the first three or four paragraphs, along with relevant empirical, conceptual, or theoretical contexts. In some cases the best approach is to begin—in the very first sentence—with

a purpose statement that clearly spells out the problem or question. At other times a researcher can more effectively convince reviewers of the proposed project's importance by first presenting the larger context within which the problem rests—a context that provides a clear-cut rationale for the project—and only then identifying the research problem or question itself.

With the exception of certain kinds of qualitative studies that, by nature, require considerable self-observation and self-reflection (e.g., see the discussion of *autoethnographies* in Chapter 8), a proposal should not include a discussion of why the researcher became personally interested in the problem or feels a burning desire to investigate it. Nor does it need explanations of why the researcher decided *not* to study certain other topics. Those who will review a proposal are typically not interested in such autobiographical excursions. Indeed, these excursions may suggest to readers that you can't separate essentials from irrelevancies; thus, they will neither enhance your stature as a researcher nor recommend you as one who can think in a clear and focused manner.

A Proposal Is Not a Literary Production

A contractor's drawings of a future house are not works of art. Similarly, a proposal is not a creative production that strives to engage readers with complex characters, vivid imagery, and a spellbinding plot. The purpose of both is simply to communicate clearly. Just as a contractor'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 exact. The proposal provides a chance to show how clearly and completely the researcher can state a research problem, delineate the collection of relevant data, and explain how those data will be interpreted in an effort to solve the problem.

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—as well as current articles in popular magazines—and you'll 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's 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. Statement of the problem
   B. Subproblems and hypotheses

C. Definitions of terms
D. Assumptions
E. Delimitations and limitations
F. Importance of the study
II. Review of the related literature
III. 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 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. Qualifications of the researcher and any assistants
V. Outline of the proposed study (steps to be taken, timeline, etc.)
VI. References
VII. Appendices

Proposals for qualitative studies often use a somewhat different format. The following format is an example of an outline you might use for a qualitative proposal:

I. Introduction
A. Purpose of the study
B. General context for the study
C. Guiding questions
D. Delimitations
E. Significance of the study
II. Methodology
A. Conceptual (or theoretical) framework
B. Type of design and its underlying assumptions
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 ensuring credibility
III. Management plan, timeline, and feasibility
IV. References
V. Appendices

A key rule governs the writing of both a proposal and a final research report: *Information should be presented in such a manner that it leads readers clearly and logically from one facet of the research project to the next.*

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 for the largest organizational units; 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 nonboldface 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 were to use APA 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 line space (that is, an empty line) immediately preceding each one. Also, you should not end a page with a first- or second-level heading; instead, place it at the top of the subsequent page so that the new section it announces follows immediately below.

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're a student, with any graduate school requirements.

Above all, you should be *consistent* in how you format your headings. We authors have seen too many proposals, theses, and dissertations in which headings of equal importance appear in ALL CAPITALS and at other times 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 a 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 word-processing software.*** In today's world, this first guideline should go without saying, but we're saying it anyway. The key issue is when you should bring this software into your writing process. 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've 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!).

We urge you to 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, ¿*), 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 institution or agency 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 doesn't 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 shouldn't 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—such as organizing thoughts, following a logical sequence, expressing ideas clearly and succinctly, using acceptable grammar and punctuation, and spelling words correctly—may far exceed that capacity. In other words, you may very well *not* be able to do it all at once!

In your 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 immediately think of the right word to use somewhere, 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're 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'll need several such bits of information. You can track them all down later, *after* you have finished your first draft.

4. ***Explicitly present your research problem or question early in your proposal, perhaps within the context of a purpose statement.*** Obviously, your potential reviewers must know what kinds of data you plan to collect, how you plan to collect them, and how you will analyze them. But these reviewers can make sense of your proposed methods only if you *first* tell them what particular problem or question you want your data to address.

5. ***Provide a relevant context for your research problem or question.*** A good proposal places the research problem or question within a specific context that helps readers understand why the problem or question to be investigated needs a 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 reflects 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 later, perhaps in the second chapter.

6. ***Convince readers of the importance of your project.*** You must convince your readers that your planned project isn't 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 don't want to get emotional in your presentation, you nevertheless want to generate interest in what you're doing: You want your readers to *want to know* what your project's outcome will be.

As illustrated in the proposal outlines presented earlier in the chapter, researchers often 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 in the introductory discussion of the 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 can 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's very much like a contract to which you and your reviewers will ultimately agree. Accordingly, leave nothing unsaid, no potential question unanswered.

8. ***Communicate that you have an open mind about the outcomes of your study.*** Researchers often embark on research studies with the hope that they'll uncover evidence in support of particular *a priori* 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 be predicted with 100% accuracy ahead of time, then what's 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* . . . ," "The proposed research project is designed to investigate the *possible* effects of . . . ," or "This project will *explore* the role 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're using a quantitative approach, you will need to specify your sample, procedures, and assessment strategies in the utmost detail. If you're using a qualitative approach, you will probably be making some sampling and procedural decisions after the study is under way. 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, a researcher may need certain kinds of records. In such a situation, the researcher should know their exact location either online or in physical space. 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 aren't available. In your proposal, 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. ***Explain how you will use the data to address your research problem.*** Even though you haven't yet collected your data, you will nevertheless need to describe how you intend to organize, analyze, and interpret them so that you can address your research problem or question. Spelling out the treatment and interpretation of the data can be a tedious, time-consuming process. But the alternative—presenting only a broad sweep, describing only a general 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 consider Figure 5.1, which shows an excerpt from an early draft of a doctoral student's proposal for a dissertation about labor relations. 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 in Figure 5.1, 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 the figure. 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?*

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.

Unfortunately, the student isn't explaining how he intends to interpret 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 professional employees toward certain aspects of management policy"? These are critical questions that should be answered *in the proposal.*

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 management lukewarm in their support? What keywords did they use in their responses? What did the overall category responses indicate about 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:

a. Restate the subproblem.
b. Clearly identify the data that relate to the subproblem.
c. Explain fully and unequivocally how you intend to analyze and interpret the data to resolve the subproblem.

More generally, *your plan for the treatment of your data should be so specific and unequivocal that any other qualified person could carry out your research project based solely on the information you present in 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 project. 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 assessment 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 offer an easy way to present any necessary details that aren't central to the points you're trying to make. Simply refer to each appendix as it's 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 D)." If you have two or more appendices, 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.

13. ***Be careful that you don't plagiarize other sources, even in subtle ways.*** We remind you of how we have previously defined *plagiarism* in Chapter 3: It 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. Even in the appendices, you must give credit to your sources; for example, if you use or modify an assessment instrument that a previous researcher has used, you must give credit to that other researcher and, if relevant, identify places where you have made changes. As this book goes to press, you can find a useful tutorial called "How to Recognize Plagiarism" on Indiana University's website (https://www.indiana.edu/~academy/firstPrinciples/choice.html).

MyLab Education Self-Check 5.1

GUIDELINES Revising Your Proposal

Your first draft will almost certainly *not* be your final proposal. We remind you of a point first made in Chapter 1: *Anticipate that you will almost certainly have to write multiple drafts.* Here we offer 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 thus won't necessarily read what you *actually wrote.*

2. ***Read a paper copy—rather than an electronic version—of your first draft.*** As previously mentioned in Chapter 1, paper copies often reveal problems in a text that somehow escape our attention on the computer screen. We authors 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 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 your computer or blank sheets of paper nearby for major rewrites.

FIGURE 5.2 Commonly Used Editing Marks

or or	delete; take it out
	close up; make one word
or > or	caret; insert here
#	insert a space
stet	let marked text stand as originally written
tr	transpose, change order the
¶	begin a new paragraph
sp	spell out (change 5 lbs. to five pounds)
cap	change to capitals (CAPITALS)
lc	change to Lowercase (lowercase)
bf	change to boldface (**boldface**)
∨	superscript or superior (as in πr^2)
∧	subscript or inferior (as in H_2O)
∧	comma
∨	apostrophe
o	period
or	quotation marks
	parentheses

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 or subtopic 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 noteworthy difference between them. It's 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 *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. Unpredictably moving back and forth among

the different labels might lead a befuddled reader to conclude that the proposed research project will have six groups instead of two!

4. ***Aggressively 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 master's student:

> It appears to be the case that not only is note taking superior for recall immediately after a 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 did 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 assessment 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 their length, of course, but break up overly long, contorted sentences into shorter, more succinct ones. Be alert to how and where you use adjectives, adverbs, and modifying phrases. Misplaced words and phrases can wreak havoc with the thoughts you want to communicate. As a simple, nonacademic example, consider the misplaced modifier in this 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., merriam-webster.com)—can help you find the exact word you need.[2] Never use a long word where a short one will do. When you have a choice, use one- or two-syllable words rather than longer ones; for example, there's often no reason to use the word *utilize* when the word *use* would serve the same purpose.[3] Use professional jargon only when you need it to relate your ideas to existing concepts, theories, and literature in the discipline.

7. ***Check carefully for errors in grammar, punctuation, spelling, and formatting.*** Now is the time to attend to nitty-gritty 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 might be equally careless and sloppy.

Your word-processing software can certainly be helpful in this respect. For instance, *use the grammar checker.* Grammar checkers can search for word expressions, clichés, multiple negation,

---

[1]In such a study, the researcher must not incorrectly tell sugar pill recipients that they're 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 don't need to be informed about which kind of pill they're taking, as such information might affect their subjective *perceptions* of pain and thereby distort the results of the study.

[2]In our experiences, we authors have found a large, hardcover thesaurus—an actual book—to be much more comprehensive and helpful than a word-processing tool or online resources.

[3]A recent Internet search of "use vs. utilize" identifies websites in which various authorities make subtle distinctions between the two words, but, curiously, they don't all make the *same* distinctions. Also, some website authors suggest that using *utilize* reflects nothing more than a writer's superficial attempt to convey a high intelligence level.

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 pointed out in Chapter 1, a spell checker typically does nothing more than check each word to see if it's a match with a word in the English language or in some other list you specify. It won't tell you whether you have used the *right* words in every case. So even if you take advantage of the spell checker, always, *always* follow up by reading your document, word for word, to be sure that every word is correctly spelled. If you're a poor speller, then ask someone else—a good speller, obviously—to proofread 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 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 behavior").
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 readers that you haven't taken the proposal-writing task seriously.

Speaking of *it's* versus *its*, you should 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. Ellis's research").

8. ***Make sure that items in bulleted lists and the like are parallel in structure.*** In writing a research proposal, a sequence of bullets, several sequentially numbered items, or a similar list 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. Sticklers for grammar (including some proposal reviewers) generally frown upon mixing incomplete and complete sentences within a single list.

For example, in any list of definitions of terms, your definitions should consistently be either (a) nonsentence phrases or (b) complete sentences. As an illustration, in a proposal for a study intended to investigate different forms of aggression in secondary school students, nonsentence-phrase definitions might take the following forms:

Physical aggression. Action that can potentially cause bodily injury (e.g., hitting or kicking).

Psychological aggression. Action intended to cause mental anguish or reduce self-esteem (e.g., ostracism, cyberbullying).

In contrast, in their complete-sentence forms, the same concepts might be defined as follows:

Physical aggression. Physical aggression (e.g., hitting or kicking) is an action that can potentially cause bodily injury.

Psychological aggression. Psychological aggression (e.g., ostracism, cyberbullying) is an action intended to cause mental anguish or reduce self-esteem.

Careful attention to such parallelism in format 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 main text, footnotes, and appendices 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, personal communication, or some other entity to which you're giving credit. Furthermore, every item in the reference list should be cited somewhere in the proposal.[4] As you learned in Chapter 3, the formats for citations and reference lists should be consistent with the style manual typically used in your particular academic discipline.

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're going to do.

11. ***Print out your revised proposal and carefully read it once again.*** Look critically at each thought as it stands on paper. Do the words say exactly what you want them to say? Read slowly, line by line. See whether one word will carry the burden of two or more. Throw out superfluous words.

12. ***Seek the feedback of others, and take it seriously when writing subsequent drafts.*** We cannot stress this point enough. No matter how hard you try, you can't be as objective as you would like to be when you read your own writing. So ask people to read and critique what you've 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. If you're writing a dissertation proposal, your doctoral committee will almost certainly request some revisions to what you have planned and written. If you're writing a master's thesis, your advisor and any advisory committee members will probably make requests and recommendations.

One final comment: ***Get used to writing!*** Researchers write continuously—sometimes to communicate with others, at other times to facilitate their own thinking. For the latter purpose, either paper or a word processor can be effective for personal brainstorming sessions. Take time to sit back and use a pen or keyboard to help you clarify your thoughts and ideas.

[4]In contrast to a reference list, a *bibliography* can include many potentially useful sources related to a topic. Virtually all research proposals and reports provide citation-only reference lists rather than bibliographies.

FIGURE 5.3 Common Weaknesses in Research Proposals

Weaknesses Related to the *Research Problem (or Question):*

- 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 theoretical or conceptual context.
- The problem is unimportant or unlikely to yield new information.
- One or more *a priori* hypotheses are ill-defined, doubtful, or unsound; they are 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. (This issue may be irrelevant to some action research and design-based research projects, in that they are specifically conducted to address local problems or unique products.)
- The project is unrelated to the funding agency's purposes and reasons for sponsoring new research.

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 procedures and assessment strategies are inappropriate for the research problem (e.g., proposed measurement instruments may have poor reliability and validity).
- Appropriate controls for the chosen methodology are either lacking or inadequate.
- The equipment to be used is outdated or inappropriate.
- Data analyses have not received adequate consideration, are too simplistic, or are unlikely to yield clear-cut and credible results.
- Potential limitations of the project, even if unavoidable, are not adequately addressed.

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 relevant scholarly literature.
- The proposal does not adhere to the appropriate style manual.
- The proposal has grammatical and/or spelling errors.

Sources: Based in part on Allen, 1960; Cuca & McLoughlin, 1987; Dahlberg, Wittink, & Gallo, 2010; Davitz & Davitz,1996; Wong, n.d.

PRACTICAL APPLICATION Strengthening Your Proposal

Not all research proposals get approved. Rejections are common for proposals requesting funding from a private organization or government agency. But some proposals to conduct nonfunded research get turned down as well, usually for one or more 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've 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

Evaluating an Early Draft of a Research Proposal

Check each item—or have a friend or colleague in your discipline check it—to be sure that your proposal exhibits *none* of the following characteristics:

For Any Research Proposal

_____ 1. The statement of the main problem (or question) 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 address the subproblem or the overall research problem.

_____ 4. The proposal lacks sharpness. It is not logically organized. Without clear sections that delineate the various aspects of the research project, it rambles. Readers have trouble isolating the discussion of the main research problem or question, the subproblems, the related literature, the methodology, the interpretation of the data, and other important parts of the proposal.

_____ 5. The proposal uses words and phrases that are too general, ambiguous, or imprecise to be useful for evaluation. Without more clarification, such phrases as "tests will be given" and "statistical procedures will be conducted" are largely meaningless.

_____ 6. The format of the proposal deviates from the guidelines set forth by the approval group or funding agency.

_____ 7. Some cited sources do not appear in the reference list; alternatively, they are incompletely or incorrectly cited.

For a Proposal to a Funding Agency

_____ 8. The problem does not address the research area outlined by the funding agency.

_____ 9. The section of the proposal explaining the study's importance is not specific enough for the funding agency to see the study's relationship to the purpose for which a grant is requested.

_____ 10. A clear and explicit budget statement outlining program expenditures is lacking, or the summary of estimated costs is ambiguous and indefinite.

_____ 11. Items included in the budget are disallowed by the terms of the grant.

_____ 12. The proposal is too ambitious for the grant money available.

MyLab Education Self-Check 5.2

MyLab Education Application Exercise 5.1: Organizing a Research Proposal

MyLab Education Application Exercise 5.2: Planning for Data Interpretation

FINAL THOUGHTS ABOUT PROPOSAL WRITING

When drawing up a contract, an attorney meticulously includes all of the rights and obligations of the parties identified 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 or question, the methodology, the data, and any other factors critical to conducting the investigation must be clearly set forth and mutually agreed on before the actual research activity can begin.

Any thesis or dissertation project must begin with a proposal, and any project involving human subjects or nonhuman animals must get IRB or IACAC approval, respectively, before it ever gets off the ground. But even when a proposal isn't mandatory, it's 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.

Most faculty advisors will want to review a thesis or dissertation proposal periodically as it's being developed. 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.

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.

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 paper.*

MyLab Education Self-Check 5.3

MyLab Education Application Exercise 5.3: Reviewing and Revising a Research Proposal

A SAMPLE RESEARCH PROPOSAL

We conclude this chapter by presenting an example of a successful research proposal—in this case, a proposal for a doctoral dissertation at the University of Northern Colorado. 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. We present the proposal itself on the left and add our commentary 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 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.

Comments

The author begins by reminding the reader of the purpose of her proposed research. The repetition of the research problem at this point, although 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 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.

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.

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 a 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.

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.)

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.

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 (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).

In quantitative research, the heading "Procedure" is more commonly used than "Data Collection," but the latter is acceptable as well.

Here the author describes her procedures regarding informed consent.

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

Here she describes the steps she will take to ensure participants' right to privacy.

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.

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.

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.

The author will vary the order in which participants respond to the instruments, presumably as a way of determining whether reading and responding to one instrument affects how a participant responds to any subsequent instruments.

The author is taking steps to increase the reliability *of the instrument by* standardizing *its administration.*

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.

The author will administer the TAASS to some participants twice so that she can determine its test–retest reliability.

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 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.

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.

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)."

(The author then describes the specific analyses she plans to conduct and how they relate to her research problem.)

Note: From a research proposal submitted in 1998 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.

APPLICATION EXERCISES

Similar exercises are available on Pearson MyLab Education.

Consider this draft of a research proposal, written by a student. Read her proposal carefully, and answer application exercises 1 through 3.

Use of Smartphones by Children and its Impact on Emotion Regulation

Introduction

The easy access to smartphones and similar gadgets among children of all ages and the heavy reliance of parents on using them as a source of entertainment for the children has led to rising concerns about the adverse effects of smartphone use at such a young age. Smartphones provide easy access to games and educative apps, that are replacing traditional forms of play and learning aids, such as drawing books, at a fast pace. Smartphones are also a source of engaging in social media, a major source of entertainment in today's world. This study aims to look at the relationship between smartphone use by young children, aged 9–11 years old and their emotional regulation.

Review of the Literature

An emphasis on the rising physiological as well as psychological problems as a result of smartphone use among individuals has only recently emerged as a burgeoning field of study. Most of these studies have focused on the consequences of smartphone use and addiction among adults, while studies on children are few. Investigating the role of smartphone use on problems being faced by children is important as excessive smartphone use among children has already been implicated in problematic behaviours and decreased emotional intelligence (Baek, Shin, & Shin, 2018; Cho & Lee, 2017), decreased executive functioning (Huber et al., 2018), paediatric dry eye disease (Moon, Kim, & Moon, 2016), short sleep duration and obesity (Chahal et al., 2012).

Several factors have been implicated which have been seen to influence the use of smartphones among children. A major factor is the screen time of the parents themselves with uninhibited use of smartphones by the parents leading to increased use of smartphones by their children as well (Lauricella, Wartella, & Rideout, 2015). Hence, the self-evaluative skills of parents with regard to mindful use of smartphones have been emphasized to prevent the adverse consequences of smartphone use among children (Cho & Lee, 2017). Other factors implicated include parental attitudes and overall family environment (Lauricella et al., 2015) along with low social support (Leung, 2006).

With regard to the specific construct of emotion regulation, it has been seen to have a significant influence on the social and emotional competence of children and is a vital skill to master for the holistic development of a child. Emotion regulation has been associated with positive outcomes in children such as academic success and productivity (Graziano et al., 2007), emotional understanding and language skills (Eisenberg, Sadovsky, & Spinrad, 2005) and reduction in problematic behaviours (Havighurst et al., 2010) among children.

Emotion regulation is achieved using two major strategies. The first strategy is Cognitive Reappraisal, which aims at reinterpreting an emotion-eliciting situation in a way that alters the meaning of the situation and thereby, changes the emotional impact of that situation (Gross & John, 2003). The second strategy is Expressive Suppression, which either hides, inhibits or reduces the expression of an ongoing emotion being felt by the individual (Gross & John, 2003). Cognitive Reappraisal has been found to contribute to better psychological health in the long term as opposed to Expressive Suppression which has been implicated in negative mental health outcomes.

Some of the prominent signs of emotional intelligence in children include the abilities to name emotions—their own and of others, being able to understand why they are feeling a certain way and being able to understand the impact of their actions on others feelings (Salovey & Mayer, 1990). Over reliance on and use of technology, a major aspect of which is the use of smartphones, has been seen to erode skills of emotion regulation and self-awareness. Smartphone use, by creating a bubble world, divorced from real-life situations and interactions, has been seen to result in a distance from one's emotional self. Instead of engaging with their own feelings and those of others', children instead remain vested in artificial choices. Smartphone use has also been seen to increase impulsivity, which makes effective emotion regulation more difficult (Grover, 2017). Socio-emotional skills are essential to navigate successfully in the world, and the time spent staring at a lifeless screen deprives children of the opportunity to work on learning and honing those skills.

The present review of literature has focused on three major ideas. The first is the association of increased smartphone use among children with multiple negative physiological and psychological outcomes. The second is the importance of emotion regulation skills for the holistic development of children. The third is about how smartphone use by children acts to the detriment of developing effective emotion regulation skills, at a stage when it is essential. The present study aims to reduce the research gap by exploring whether a direct association between smartphone use by children and their use of one or the other emotion regulation strategies can be established.

Present Study

The aim of the present study is to establish whether the use of smartphone by children aged 9–11 years of age has an impact on the emotion regulation strategies used by them. For the purpose of this study, the two major emotion regulation strategies of Cognitive Reappraisal and Expressive Suppression are being focused upon. As elaborated in the review of literature, Cognitive Reappraisal has been seen to be a helpful emotion regulation strategy while Expressive Suppression has been seen to have potential harmful effects in the long run.

The hypothesis for this study is that increased time spent on smartphones by children will result in greater use of Expressive Suppression as a way to regulate emotions, thereby resulting in ineffective emotion regulation. Smartphone use, resulting in greater impulsivity, is likely to result in lower frustration tolerance (Grover, 2017). Moreover, a lack of self-reflexivity brought about by smartphone use is also likely to impede the use of alternative thinking sets, which is required for Cognitive Reappraisal to occur.

Children aged 9–11 years are being chosen for this study as this is the time they start exercising more control over their surroundings, being in a position to make a conscious choice about smartphone use. This is also the time they start using more stable and sophisticated emotion regulation strategies (Gullone et al., 2010). Hence, this seems the right age to study the relationship between these two variables.

A total of 100 children will be a part of the sample. These children will be recruited from a primary school setup and informed consent will be obtained from both the parents and the children. The children will be randomly selected from all the sections of a particular standard till the desired number of students for the study is achieved.

Smartphone use of the children will be tracked by the parents, to ensure greater accuracy in data. The number of hours spent by the child on a smartphone every day, for a total of 7 days, will be recorded and reported by the parents. This data will be recorded separately for each day. The researcher will also estimate the average time spent by each child on smartphones for the one week period.

In order to tap into the use of emotion regulation strategies by children, the Emotion Regulation Questionnaire for Children and Adolescents (ERQ-CA) will be administered to the children at the end of the week, once their usage of smartphone over a one week period has been mapped. The ERQ-CA has been developed by Gross and John (2003) for use with children and adolescents. It has 10 statements in total, with 6 statements tapping into the use of Cognitive Reappraisal and 4 statements tapping into the use of Expressive Suppression.

For the purpose of data analysis, correlations will be computed to determine associations, if any, between the two variables of smartphone use and emotion regulation strategies. The time spent on smartphones by children will be sorted into different categories and a one-way ANOVA will be used to assess whether there are significant differences among the different conditions (categories) when it comes to the use of emotion regulation strategies. Furthermore, differences if any, on the basis of gender and socioeconomic status, will be assessed using independent samples t-test.

The study has implications for understanding the impact of smartphone use by children on their emotion regulation skills, and hence on their overall emotional intelligence. In case a significant association is found, it will have implications for finding alternative ways of enhancing emotion regulation skills among children, as well as for limiting the use of smartphones by children. In contemporary urban society where smartphone use is on a drastic rise and smartphone addiction is becoming common, the study will add value to this as-yet less explored area of research.

Bibliography

Baek, H. W., Shin, Y. M., & Shin, K. M. (2014). Emotional and behavioral problems related to smartphone overuse in elementary school children. *Journal of Korean Neuropsychiatric Association, 53*(5), 320–326.

Chahal, H., Fung, C., Kuhle, S., & Veugelers, P. J. (2013). Availability and night-time use of electronic entertainment and communication devices are associated with short sleep duration and obesity among Canadian children. *Pediatric Obesity, 8*(1), 42–51.

Cho, K. S., & Lee, J. M. (2017). Influence of smartphone addiction proneness of young children on problematic behaviors and emotional intelligence: Mediating self-assessment effects of parents using smartphones. *Computers in Human Behavior, 66*, 303–311.

Eisenberg, N., Sadovsky, A., & Spinrad, T. L. (2005). Associations of emotion-related regulation with language skills, emotion knowledge, and academic outcomes. *New Directions for Child and Adolescent Development, 2005*(109), 109–118.

Graziano, P. A., Reavis, R. D., Keane, S. P., & Calkins, S. D. (2007). The role of emotion regulation in children's early academic success. *Journal of School Psychology, 45*(1), 3–19.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*(2), 348.

Grover, S. (2017, July 14). How Technology Lowers Emotional Intelligence in Kids. Psychology Today. https://www.psychologytoday.com/us/blog/when-kids-call-the-shots/201707/how-technology-lowers-emotional-intelligence-in-kids

Gullone, E., Hughes, E. K., King, N. J., & Tonge, B. (2010). The normative development of emotion regulation strategy use in children and adolescents: A 2-year follow-up study. *Journal of Child Psychology and Psychiatry, 51*(5), 567–574.

Havighurst, S. S., Wilson, K. R., Harley, A. E., Prior, M. R., & Kehoe, C. (2010). Tuning in to kids: Improving emotion socialization practices in parents of preschool children–findings from a community trial. *Journal of Child Psychology and Psychiatry, 51*(12), 1342–1350.

Huber, B., Yeates, M., Meyer, D., Fleckhammer, L., & Kaufman, J. (2018). The effects of screen media content on young children's executive functioning. *Journal of Experimental Child Psychology, 170*, 72–85.

Lauricella, A. R., Wartella, E., & Rideout, V. J. (2015). Young children's screen time: The complex role of parent and child factors. *Journal of Applied Developmental Psychology, 36*, 11–17.

Leung, L. (2006). Stressful life events, motives for Internet use, and social support among digital kids. *CyberPsychology & Behavior, 10*(2), 204–214.

Moon, J. H., Kim, K. W., & Moon, N. J. (2016). Smartphone use is a risk factor for pediatric dry eye disease according to region and age: a case control study. *BMC Ophthalmology, 16*(1), 188.

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality, 9*(3), 185–211.

1. Read the proposal carefully, and answer the following questions.
   - **i.** Identify the statement of the problem as mentioned in the Introduction of the study. When does the statement of the problem become clear to the reader?
   - **ii.** What can be added to the Introduction section to provide more information about the statement of the problem?
   - **iii.** Looking at the Review of Literature section, how well organized does it read to you? What are some of the good aspects of this section?
   - **iv.** What typical elements of a research proposal are missing from this student's proposal?

2. Read the proposal carefully, and answer the following questions.
   - **i.** What do you understand about the data collection procedure as elaborated in the research proposal under the Present Study section?
   - **ii.** Does the data collection procedure evoke any concerns for you with regard to its process and rigour?
   - **iii.** The author has conceptualized a quantitative study. How does the author plan to analyse the data using quantitative analysis?

3. Read the proposal carefully, and answer the following questions.
   i. Is the research proposal well-organized?
   ii. What is your analysis of the proposed methodology of the study under the Present Study section of the research proposal? What are the good aspects and what are the areas that can be improved?
   iii. The author has listed out the implications of the study. Do they form sufficient basis to convince the reader about the importance of the study?

FOR FURTHER READING

Bernstein, T. M. (1993). *The careful writer: A modern guide to English usage.* New York, NY: Free Press.

Booth, W. C., Colomb, G. G., Williams, J. M., Bizup, J., & FitzGerald, W. T. (2008). *The craft of research* (4th ed.). Chicago, IL: University of Chicago Press.

Dahlberg, B., Wittink, M. N., & Gallo, J. J. (2010). Funding and publishing integrated studies: Writing effective mixed methods manuscripts and grant proposals. In A. Tashakkori & C. Teddlie (Eds.), *Mixed methods in social & behavioral research* (2nd ed., pp. 775–802). Los Angeles, CA: Sage.

DeCuir-Gunby, J. T., & Schutz, P. A. (2016). *Developing a mixed methods proposal: A practical guide for beginning researchers.* Los Angeles, CA: Sage.

Johnson-Sheehan, R. (2008). *Writing proposals* (2nd ed.). New York, NY: 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.

Locke, L. F., Spirduso, W. W., & Silverman, S. J. (2014). *Proposals that work: A guide for planning dissertations and grant proposals* (6th ed.). Los Angeles, CA: Sage.

Miner, L. E., & Miner, J. T. (2013). *Proposal planning and writing* (5th ed.). Santa Barbara, CA: Greenwood.

Roberts, C. M. (2010). *The dissertation journey: A practical and comprehensive guide to planning, writing, and defending your dissertation* (2nd ed.). Los Angeles, CA: Corwin.

Rudestam, K. E., & Newton, R. R. (2015). *Surviving your dissertation: A comprehensive guide to content and process* (4th ed.). Los Angeles, CA: Sage.

Terrell, S. R. (2016). *Writing a proposal for your dissertation: Guidelines and examples.* New York, NY: Guilford Press.

Chapter 6 Descriptive Research

Our physical and social worlds present overwhelming amounts of information. But you can learn a great deal if you study a well-chosen sample from one of those worlds and draw reasonable inferences from your observations of this sample.

Learning Outcomes

6.1 Describe general characteristics and purposes of (a) observation studies, (b) correlational research, (c) developmental designs, (d) experience-sampling methods, and (e) survey research. Also, describe strategies you might use in each of these research methodologies.

6.2 Identify effective strategies for conducting a face-to-face, telephone, or video-conferencing interview.

6.3 Identify effective strategies for constructing and administering a questionnaire and for analyzing people's responses to it.

6.4 Explain possible uses of checklists, rating scales, rubrics, computer software, and the Internet in data collection.

6.5 Determine an appropriate sample for a descriptive study.

6.6 Describe common sources of bias in descriptive research, as well as strategies for minimizing the influences of such biases.

The term **descriptive research** encompasses a variety of methodologies that are best suited to examining and trying to make sense of a situation or event *as it currently exists* in the world. This general category of research designs includes both (a) designs aimed at characterizing the general nature of an observed phenomenon and (b) designs aimed at revealing possible associations among two or more phenomena. By and large, descriptive research does *not* involve changing or modifying a situation under investigation, nor is it intended to determine cause-and-effect relationships.

Much of our focus in this chapter is on quantitative approaches in descriptive research. Nevertheless, some of the things we say about observation studies, survey research, interviews, questionnaires, and sources of bias are relevant to qualitative research as well.

DESCRIPTIVE RESEARCH DESIGNS

In the next few pages, we describe observation studies, correlational research, developmental designs, experience-sampling methods, and survey research, all of which typically 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

As you will discover in Chapter 8, many qualitative researchers rely heavily on personal observations—typically of people or another animal species (e.g., gorillas, chimpanzees)—as a source of data. In *quantitative* research, however, an **observation study** is quite different. For one thing, an observation study in quantitative research might be conducted with plants rather than animals, or it might involve nonliving objects (e.g., rock formations, soil samples) or dynamic physical phenomena (e.g., weather patterns, black holes).

Also, a quantitative observation study tends to have a limited, prespecified focus. When human beings are the topic of study, the focus is usually 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, a researcher strives to be *as objective as possible* in assessing the behavior being studied. To maintain such objectivity, the researcher 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 later in the chapter).
- Have two or three people rate the same behavior independently, without knowledge of one another's ratings.
- Train the raters 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) provides an example of 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 also what people near them said) to a remote audio recorder. Each teacher's speech was recorded for 15 minutes on each of two different days. Following data collection, the recordings 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 specific 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 some of the richness and complexity of human behavior. In certain situations, then, it provides a quantitative alternative to such qualitative approaches as ethnographies and grounded theory studies (both described in Chapter 8).

Correlational Research

A **correlational study** examines the extent to which differences in one variable are associated with differences in one or more *other* variables. A **correlation** exists if, when one variable increases, another variable either increases or decreases in a somewhat predictable fashion. Knowing the value of one variable, then, enables us to *predict* the value of the other variable with some degree of accuracy.

In correlational studies, researchers gather quantitative data about two or more characteristics for a particular group of people or other appropriate units of study. When human beings are the focus of investigation, the data might be test scores, ratings assigned by one or more expert observers, or frequencies of certain behaviors. Data in animal studies, too, might be frequencies of particular behaviors, but alternatively, they could 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 researchers' assessment instruments can measure with some objectivity. Whatever the nature of the data, at least two different variables are assessed in order to determine whether and in what way these variables are interrelated.

Let's consider a simple example: As children grow older, most of them 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, knows the children's ages, and obtains reading achievement test scores for them that indicate an approximate "grade level" at which each child is 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 age and reading ability. Figure 6.1 presents this hypothetical scatter plot. Chronological age is on the graph's vertical axis (the *ordinate*), and reading grade 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.

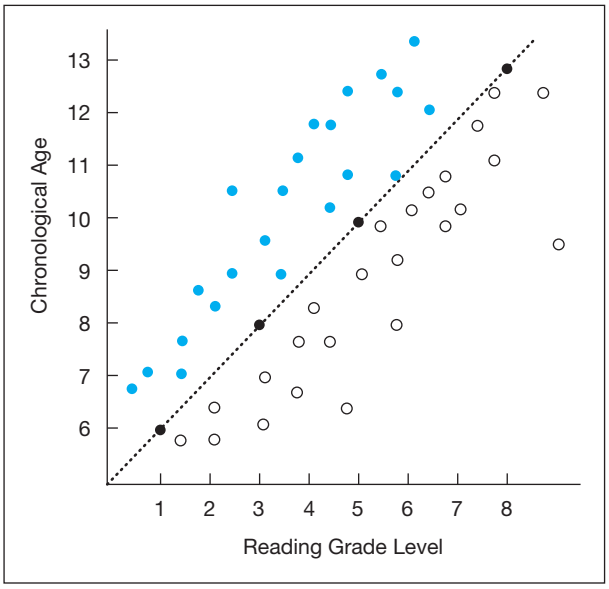


FIGURE 6.1 Example of a Scatter Plot: Correlation Between Age and Reading Level

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 roughly elliptical shape (as the dots in Figure 6.1 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 6.1—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 relative to their peers; these children are designated by colored 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 reading ability 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 importantly—we can *interpret* these data and give them meaning. The data tell us not only that children become better readers as they grow older—that's a "no-brainer"—but also that any predictions of children's future reading abilities based on age alone will be imprecise ones at best.

A Caution About Interpreting Correlational Results

When two variables are correlated, researchers sometimes conclude that one of the variables must in some way cause or 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.*

Let's take a silly example. A joke that seems to have gone viral on the Internet is this one:

> *I don't trust joggers. They're always the ones that find the dead bodies. I'm no detective . . . just sayin'.*

The tongue-in-cheek implication here is that people who jog a lot are more likely to be murderers than people who don't jog very much and that perhaps jogging *causes* someone to become a murderer—a ridiculous conclusion! The faulty deduction regarding a possible cause-and-effect relationship is crystal clear.

In other cases, however, it would be all too easy to draw an unwarranted cause-and-effect conclusion on the basis of correlation alone. For example, in a series of studies published in the journal *Psychological Science*, researchers reported several correlations between parenthood and psychological well-being: As a general rule, adults who have children tend to be happier—and to find more meaning in life—than adults who don't have children (Nelson, Kushlev, English, Dunn, & Lyubomirsky, 2013). Does this mean that becoming a parent *causes* greater psychological well-being? Not necessarily. Possibly the reverse is true—that happier people are more likely to *want* to have children and so take steps to have them either biologically or through adoption. Or perhaps some other factor is at the root of the relationship—maybe financial stability, a strong social support network, a desire to have a positive impact on the next generation, or some other variable we haven't considered.

The data may not lie, but our cause-and-effect interpretations of them may, at times, be highly suspect. Ideally, a good researcher isn't content to stop at a correlational relationship because *beneath the correlation* may lie some potentially 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 of the experimental studies described in Chapter 7 to test various hypotheses about what causes what.

Developmental Designs

In the preceding section, we presented a hypothetical example of how children's reading levels might correlate with their ages. 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 investigate 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 researcher might get measures of both academic achievement and social adjustment for a group of fifth-graders and then, 10 years later, find out what their high school GPAs were and whether they had actually graduated and earned a diploma. This researcher might also compute correlations between the variables assessed in the fifth grade and the students' high school performance variables; 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* (a.k.a. *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 assessed 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 assessed 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 assessment instrument, they might improve simply because of their *practice* with the instrument, even if the characteristic being assessed hasn't changed at all.

[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., see Kuhn, 1995; Kuhn, Goh, Iordanou, & Shaenfield, 2008).

But cross-sectional designs have their disadvantages as well. For one thing, the different age groups sampled may have grown up in distinctly different environmental conditions. For example, imagine that we want to find out whether critical 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, perhaps using a computer-administered multiple-choice test, assess their ability to critically analyze various scenarios. Now imagine that, in this study, the 20-year-olds obtain higher scores on our critical thinking test than the 70-year-olds. Does this mean that critical thinking ability declines with age? Not necessarily. At least two other possible explanations readily come to mind. First, instructional practices have changed in many ways over the past few decades, and the younger participants may possibly have had more practice using critical thinking skills than has been true for the older participants. Also, it's quite likely that the younger folks have had more experience taking computer-administered tests than the older folks have had. Such problems pose threats to the *internal validity* of this cross-sectional study: We can't eliminate other possible explanations for the results observed (recall our brief mention of internal validity in the section "Enhancing the Credibility of Your Findings" in Chapter 4).

A second disadvantage of a cross-sectional design is that we cannot compute correlations across different age groups in the study. Consider, again, the educational researcher who wants to use students' academic achievement and social adjustment in fifth grade to predict their high school success. If this researcher were to use a cross-sectional study, the fifth-graders and the post–high school participants would be two different groups of people—and thus there would be only one set of measures for each individual—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 lengthy period (this is the longitudinal piece). As an example, let's return to the issue of how people's critical thinking ability changes over time. Imagine that instead of conducting 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. At the beginning of the study, we give both groups a computer-administered test designed to assess critical thinking; then, 5 years later, we give the test a second time. If both groups improve over the 5-year time span, we might wonder if practice in taking computer-administered tests in general or practice in taking this *particular* test might partly account for the improvement. Alternatively, if the test scores increase for the younger group (now 25-year-olds) but decrease for the older group (now 70-year-olds), we might reasonably conclude that critical thinking ability *does* decrease somewhat in the later decades of life.

Like a longitudinal study, a cohort-sequential study 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 critical 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 might reasonably conclude that critical thinking ability is a relatively stable characteristic—that certain people currently engage and will continue to engage in more critical thinking than others do. Furthermore, we could add other variables to the study—for instance, the amount of postsecondary education that participants have had and the frequency with which they engage in activities that require critical thinking—and determine whether such variables mediate or moderate the long-term stability of critical thinking.

Cross-sectional, longitudinal, and cohort-sequential designs are used in a variety of disciplines, but as you might guess, they're 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 articles in such journals as *Child Development* and *Developmental Psychology* for ideas about specific research strategies.

Experience-Sampling Methods

An **experience-sampling method (ESM)** is an approach in which a researcher collects frequent and ongoing data about people as they live their normal, everyday lives.[2] Many ESM studies require some kind of technological device that people wear or carry with them. For example, a study might involve having participants (a) wear small electronic devices that measure their movements, heart rate, or blood pressure; (b) carry smartphones that regularly ask brief survey questions; or (c) carry cameras with which they take pictures of their immediate surroundings or current activities. In this high-tech age of Apple Watches, Fitbits, and similar timekeepers, some people (including one of us authors) are electronically keeping ongoing records of their daily physical exercise (e.g., how many steps they take, how many hours they spend in active movement); in a sense, these folks are conducting their own informal ESM research.

Experience-sampling methods have been successfully used in both quantitative and qualitative research projects. In fact, some of their earliest users were qualitative researchers who wanted to capture the essence of people's perceptions, feelings, and other subjective reactions to their immediate circumstances and experiences—a focus that should remind you of our discussion of *phenomenology* in Chapter 1. But increasingly, quantitative researchers have also been discovering the benefits of ESM as a means of collecting informative data about such diverse variables as people's physiological well-being, social environments, reactions to particular events, and mental or emotional states (Zirkel, Garcia, & Murphy, 2015).

A study by Shernoff and his colleagues using both ESM and coded researcher observations provides an example (Shernoff et al., 2016). With their study, these researchers wanted to identify possible relationships between (a) the quality of instructional activities in high school classes and (b) students' interest and mental engagement in those activities. To do so, they recruited students in seven different classrooms (representing lessons in six different content areas) and, at 25-minute intervals during a class period, prompted the participating students to complete a brief rating-scale-based survey regarding their interest and engagement in the activity just prior to the signal, as well as their perceptions of the activity and their general mood at the time. Meanwhile, two video cameras recorded each classroom lesson in its entirety, with one focused on the classroom teacher and the other focused on a small group of student participants. Two raters coded various "instructional episodes" within each video with respect to such characteristics as "authentic and challenging tasks," "clear goals," "supportive relationships," and "physical activity." In their statistical analyses, Shernoff and his colleagues found that certain aspects of instructional activities—primarily those that reflected environmental support for students' learning efforts—did indeed correlate with students' interest and engagement in a lesson.

Experience-sampling methods have several advantages that can enhance the overall quality of a research study. First, by assessing certain characteristics, behaviors, or other entities in real time (i.e., either during or immediately after their occurrence), a researcher might increase the accuracy and validity of the assessments. Second, by getting multiple assessments of a single characteristic for a single individual, a researcher gains data that might be useful in determining the test–retest reliability of an assessment instrument; furthermore, by averaging some of those multiple assessments (if appropriate), the researcher might *enhance* test–retest reliability. Third, ESM can be helpful if a researcher wants to collect longitudinal data as a means of investigating any short-term *changes* in a characteristic as environmental or behavioral variables change. In fact, ESM can be combined with any of the other methodologies we describe in this chapter. If you look once again at the preceding description of Shernoff and colleagues' study, you will see that it incorporated both observations and surveys and that its statistical analyses included computation of correlations.

[2]You may sometimes see the term *repeated measures* used within the context of ESM because such methods do, in fact, assess one or more variables over and over again. Repeated-measures variables are also used in some of the experimental and quasi-experimental designs described in Chapter 7.

Survey Research

In **survey research**, a researcher obtains information about one or more groups of people—perhaps about their behaviors, opinions, attitudes, or previous experiences—by asking them questions and either directly tabulating or systematically coding their answers. Often, it involves collecting data about a *sample* of individuals that is presumed to represent a much larger *population.*

Reduced to its basic elements, a *survey* is quite simple in design: The researcher (a) poses a series of questions to willing participants; (b) systematically classifies and codes any complex responses to open-ended questions; (c) summarizes both the coded responses and participants' cut-and-dried, *it's-clearly-this-or-that* responses with percentages, frequency counts, or more sophisticated statistical indexes; and then (d) draws inferences about a particular population from the sample participants' responses. A survey 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 views 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 a researcher to conduct than other types of research. Quite the contrary, a survey design makes critical demands on the researcher that, if not carefully addressed, 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 might generalize about the state of affairs for a longer time period. But we must keep in mind the wisdom of the Greek philosopher Heraclitus: There is nothing permanent but change.[3]

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 strictly online methods of data collection.

Face-to-Face and Telephone Interviews

In quantitative survey research, interviews tend to be *standardized*—that is, everyone is asked the same set of questions (recall the discussion of *standardization* in Chapter 4). In a **structured interview**, the researcher asks certain questions and nothing more. In a **semistructured interview**, the researcher may follow the standard questions with one or more individually tailored questions to get clarification or probe a person's reasoning; such an interview obviously has a qualitative element as well.

Face-to-face interviews have the distinct advantage of enabling a 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 often less expensive, and a researcher has potential access to virtually anyone on the planet who has a landline telephone or cell phone. Although response rates for phone interviews aren't as high as those for face-to-face interviews—many people are apt to be busy, annoyed at being bothered, concerned about using costly cell-phone minutes, or otherwise not interested in participating—they are considerably higher than response rates for mailed questionnaires. Unfortunately, a researcher

[3]There is apparently some doubt about whether Heraclitus said these exact words; however, he is well known for his belief that change is central to our world, and in *Kratylos*, Plato quotes him as saying, "Everything changes and nothing stands still."

conducting telephone interviews can't 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.

Midway between a face-to-face interview and a telephone interview is an interview conducted using Skype (skype.com) or other video-conferencing software. Such a strategy can be helpful when face-to-face contact is desired with participants in distant locations. However, participants must (a) feel comfortable using modern technologies, (b) have easy access to the needed equipment and software, and (c) be willing to schedule an interview in advance—three qualifications that can, like telephone interviews, lead to bias in the sample chosen.

Whether they're conducted face-to-face, over the telephone, or via video conferencing software, personal interviews allow a 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

Written questionnaires can be distributed to a large number of people, including those who live in faraway locations, potentially saving a researcher travel expenses or lengthy long-distance telephone calls. Also, participants can respond to questions with anonymity—and thus with some assurance that their responses won't come back to haunt them. Accordingly, some participants 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. For instance, when they're distributed by mail or e-mail, the majority of people who receive them don't respond to and return them—in other words, there may be a low **return rate**—and the people who do return them aren't necessarily representative of the originally selected sample. Asking people to complete a brief questionnaire when you see them *in person*, at, say, a shopping mall or popular university hangout, can encourage a greater percentage of individuals to complete the questionnaire—in other words, it can yield a relatively high response rate—but your sample will inevitably be limited to people who happen to be in the right place at the right time.

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, a researcher must identify *in advance* all of the questions that will be asked—and thereby eliminates other questions that *could* be asked about the issue or phenomenon in question. As a result, the researcher gains only limited, and possibly distorted, information—introducing yet another possible source of bias affecting the data obtained.

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.

MyLab Education Self-Check 6.1

MyLab Education Application Exercise 6.1: Selecting a Design

PLANNING FOR DATA COLLECTION IN A DESCRIPTIVE STUDY

Naturally, a quantitative descriptive study involves measuring one or more variables in some way. With this point in mind, let's return to a distinction first made in Chapter 4: the distinction between substantial and intangible phenomena. Measuring *substantial phenomena*—those that have an obvious basis in the physical world—is usually pretty straightforward. Some widely accepted measurement techniques also exist for studying *intangible phenomena*—concepts, abilities, and other abstract entities that can't be pinned down in terms of precise

physical qualities. For example, an economist might use gross domestic product (GDP) statistics as measures of a nation's economic growth, and a psychologist might use the *Stanford-Binet Intelligence Scales* to obtain IQ scores that to some degree reflect children's general cognitive abilities.

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 quantitative assessment instruments exist. In such instances, researchers often collect data through systematic observations, interviews, or questionnaires. In the following sections, we explore a variety of strategies related to these data-collection techniques.

PRACTICAL APPLICATION Using Checklists, Rating Scales, and Rubrics

Three tools that can facilitate quantification of complex phenomena are checklists, rating scales, and rubrics. A **checklist** is a list of behaviors or characteristics for which a researcher is looking. The researcher—or in many 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**.[4]

Checklists and rating scales can presumably be used in research related to a wide variety of phenomena, including those involving human beings, nonhuman animal species, 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 summertime 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 change 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 Richard Trahan, a sociologist at a nearby university, to address their research question: How do park visitors feel about the new bus system? Trahan 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 6.2. Based on people's responses, Trahan concluded that people were solidly in favor of the bus system (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

[4]Although we authors have often heard *Likert* pronounced as "lie-kert," Rensis Likert pronounced his name "lick-ert."

FIGURE 6.2 Excerpts from a Survey at Rocky Mountain National Park. Item 4 is a *Checklist.* Items 5 and 6 are *Rating Scales*

Source: 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

the instructional intervention, she read a series of statements and asked students to either agree or disagree with each statement by pointing to one of four faces. The statements and the rating scale that students used to respond to them are presented in Figure 6.3.

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 respondents straddle the fence by including a "No opinion" or other neutral response, and then design your scales accordingly.

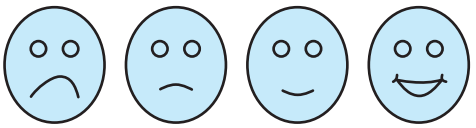
Whenever you use checklists or rating scales, you simplify and more easily quantify people's behaviors or attitudes. Furthermore, when participants *themselves* complete these things, you can collect a great deal of data quickly and efficiently. In the process, however, you don't get information about *why* participants respond as they do—qualitative information that might ultimately help you make better sense of the results you obtain.

An additional problem with rating scales is that people don't necessarily agree about what various points along a scale mean; for instance, they may interpret such labels as "Excellent" or "Strongly disapprove" in idiosyncratic ways. Especially when researchers rather than participants are evaluating certain behaviors—or perhaps when they're evaluating certain *products* that participants have created—a more explicit alternative is a **rubric**. Typically, a rubric includes two or more rating scales for assessing different aspects of participants' performance, with concrete descriptions of what performance looks like at various points along each scale. As an example, Figure 6.4 shows a possible six-scale rubric for evaluating several qualities in samples of people's nonfiction writing. A researcher could quantify the ratings by attaching numbers to the labels. For example, a "Proficient" score might be 5, an "In Progress" score might be 3, and "Beginning to Develop" might be 1.

FIGURE 6.3 Asking Elementary School Children About Science and Learning

Source: 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.



1. No | 2. Sort of No | 3. Sort of Yes | 4.Yes

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 of 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.

Such numbers would give the researcher some flexibility in assigning scores (e.g., a 4 might be a bit less skilled than "Proficient" but more skilled than just an "In Progress" label would indicate).

Keep in mind, however, that although rating scales and rubrics might yield numbers, a researcher can't necessarily add the results of different scales together. For one thing, rating scales often yield *ordinal* data rather than *interval* data, precluding even such simple mathematical calculations as addition and subtraction (see the section "Types of Measurement Scales" in Chapter 4). Also, combining the results of different scales into a single score may make no logical sense. For example, imagine that a researcher uses the rubric in Figure 6.4 to evaluate people's writing skills and translates the "Proficient," "In Progress," and "Beginning to Develop" labels into scores of 5, 3, and 1, respectively. And now imagine that one person gets scores of 5 on the first three scales (all of which reflect writing mechanics) but scores of only 1 on the last three scales (all of which reflect organization and logical flow of ideas). Meanwhile, a second individual gets scores of 1 on the three writing-mechanics scales and scores of 5 on the three organization-and-logical-flow scales. Both people would have total scores of 18, yet the quality of their writing samples would be quite different.

FIGURE 6.4 Possible Rubric for Evaluating Samples of People's Nonfiction Writing

Source: From *Essentials of Educational Psychology* (5th ed.), by J. E. Ormrod & B. D. Jones, 2018. Copyright 2018, Pearson. Used by permission.

Characteristic	*Proficient*	*In Progress*	*Beginning to Develop*
Correct spelling	Writer correctly spells all words.	Writer correctly spells most words.	Writer incorrectly spells many words.
Correct punctuation and capitalization	Writer uses punctuation marks and uppercase letters where, and only where, appropriate.	Writer occasionally (a) omits punctuation marks, (b) inappropriately uses punctuation marks, or (c) inappropriately uses uppercase/lowercase letters.	Writer makes many punctuation and/or capitalization errors.
Complete sentences	Writer uses complete sentences throughout, except when using an incomplete sentence for a clear stylistic purpose. Writing includes no run-on sentences.	Writer uses a few incomplete sentences that have no obvious stylistic purpose, or writer occasionally includes a run-on sentence.	Writer includes many incomplete sentences and/or run-on sentences; writer uses periods rarely or indiscriminately.
Clear focus	Writer clearly states main idea; sentences are all related to this idea and present a coherent message.	Writer only implies main idea; most sentences are related to this idea; a few sentences are unnecessary digressions.	Writer rambles, without a clear main idea, or writer frequently and unpredictably goes off topic.
Logical train of thought	Writer carefully leads the reader through his/her own line of thinking about the topic.	Writer shows some logical progression of ideas but occasionally omits a key point essential to the flow of ideas.	Writer presents ideas in no logical sequence.
Convincing statements/arguments	Writer effectively persuades the reader with evidence or sound reasoning.	Writer includes some evidence or reasoning to support ideas/opinions, but a reader could easily offer counterarguments.	Writer offers ideas/opinions with little or no justification.

PRACTICAL APPLICATION Computerizing Observations

One good way of enhancing your efficiency in data collection is to record your observations on a laptop, computer tablet, or smartphone as you're 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. Alternatively, you might download free or inexpensive data-collection software for your smartphone or computer tablet; examples are Magpi (magpi.com), Open Data Kit (opendatakit.org), and QuickTapSurvey (quicktapsurvey.com).

For especially complex observations, you might create a general *template* document in spreadsheet or word-processing software and then electronically "save" a separate version of the

document for each person, situation, or other entity you're 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.

For some types of observations, existing software programs can greatly enhance a researcher's accuracy and efficiency in collecting observational data. An example is CyberTracker (cybertracker.org), with which researchers can quickly record their observations and—using global positioning system (GPS) signals—the specific locations at which they make each observation. For instance, a biologist working in the field might use this software to record specific places at which various members of an endangered animal species or invasive plant species are observed. Furthermore, CyberTracker enables the researcher to custom-design either verbal or graphics-based checklists for specific characteristics of each observation; for instance, a checklist might include photographs of what different flower species look like or drawings of the different leaf shapes that a plant might have.

PRACTICAL APPLICATION Planning and Conducting Interviews in a Quantitative Study

In a quantitative study, interviews tend to be carefully planned in advance, and they're conducted in a similar, standardized way for all participants. Here we offer guidelines for conducting interviews in a quantitative study; some of them are also applicable to the qualitative interviews described in Chapter 8.

GUIDELINES Conducting Interviews in a Quantitative Study

Taking a few simple steps in planning and conducting interviews can greatly enhance the quality of the data obtained, as reflected in the following recommendations.

1. ***Limit questions to those that will directly or indirectly help you answer your research question.*** Whenever you ask people to participate in a research study, you're asking for their *time.* They're more likely to say *yes* to your request if you ask for only a reasonable amount of their time. If, instead, you want 2 or 3 hours from each potential participant, you're apt to end up with a sample comprised primarily of people who either (a) are especially interested in and concerned about your topic or (b) aren't terribly busy with other activities and commitments—potential sources of bias that can adversely affect the generalizability of your results.

2. ***As you write the interview questions, consider how you can quantify the responses, and modify the questions accordingly.*** Especially if you're conducting a quantitative study, you will, to a considerable degree, be coding people's responses as numbers and subsequently conducting statistical analyses on those numbers. You'll be able to assign numerical codes to responses more easily if you identify an appropriate coding scheme ahead of time.

3. ***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 hard to interpret.

4. ***Consider asking a few questions that will elicit qualitative information.*** 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. If you ultimately want to assign numbers to people's responses to open-ended questions—perhaps because you want to include them in statistical analyses—you might use

either checklists or rating scales to quantify their content. And don't forget that numbers can be used simply as labels for distinctly different categories of responses (recall the discussion of *nominal scales* in Chapter 4).

5. ***Consider how you might use a computer to streamline the process.*** Some computer software programs allow you to record interviews directly onto a laptop computer and then transform these conversations into written text; examples are Dragon Naturally Speaking (nuance.com/dragon) and SurveyGizmo (surveygizmo.com). Alternatively, if interviewees' responses are likely to be quite short, you might either (a) use a multiple-choice-format checklist to immediately categorize them or (b) directly type them into a spreadsheet or word-processing program.

6. ***Pilot-test the questions.*** 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 systematic data collection. Asking a few volunteers to answer your questions in a pilot study can help you identify the questions' shortcomings.

7. ***Courteously introduce yourself to potential participants, and explain the general purpose of your study.*** You're more likely to gain potential participants' cooperation if you're friendly, courteous, and respectful and if you explain—up front—what you're hoping to learn in your research. The goal here is to motivate people to *want* to help you out by giving you a little bit of their time.

8. ***Get written permission.*** Recall the discussion of *informed consent* in the section on ethical issues in Chapter 4. All participants in your study (or, in the case of children, their parents or other legal guardians) should agree to participate in advance and in writing.

9. ***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've established rapport and gained a person's trust. You might also preface a sensitive topic with a brief statement suggesting that violating certain laws or social norms—although not desirable—is fairly commonplace (Creswell, 2015; Gall, Gall, & Borg, 2007). For example, you might say something like this: "Many people admit that they have occasionally driven a car while under the influence of alcohol. Have you ever driven a car when you probably shouldn't have because you've had too much to drink?"

10. ***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?"—may yield the additional information you need.

MyLab Education Self-Check 6.2

PRACTICAL APPLICATION Constructing and Administering a Questionnaire

Questionnaires may seem to be quite simple, but 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 many suggestions that can help you make your use of a questionnaire both fruitful and efficient. We have divided our suggestions into three categories: constructing a questionnaire, using technology to facilitate questionnaire administration and data analysis, and maximizing your return rate.

GUIDELINES Constructing a 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 short as possible and solicit only information that is essential to your research effort. You should evaluate each item by asking yourself two questions: "What do I intend to do with the information I'm requesting?" and "Is it absolutely essential to have this information to address part of my 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're asking for people's *time*, a precious commodity for many folks these days. People are more likely to respond to a questionnaire—and to do so promptly—if they perceive it to be quick and easy to complete (McCrea, Liberman, Trope, & Sherman, 2008).

Open-ended questions that ask people to respond in considerable depth can be time-consuming and mentally exhausting for both the participants and the researcher. Furthermore, the usefulness of responses to open-ended items rests entirely on participants' skill in expressing their thoughts in writing. Those who write in the style of "Yes/no, and I'll tell you exactly why" are few and far between. Some respondents may ramble, engaging in discussions that aren't focused or don't answer the questions. Furthermore, when asking 10 or more questions that all require lengthy responses, your respondents will think you're 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. People from some cultural backgrounds 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 try to avoid words that have imprecise meanings, such as *several* and *usually.*

5. ***Give a rationale for any items whose purposes may be unclear.*** We cannot say this enough: You are asking potential participants 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 its 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's 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 more likely to be a smoker than a nonsmoker; notice how "None" is the last option listed. A second assumption is that a person smokes the same number of cigarettes each day, but for many smokers, this assumption isn't viable; for instance, they may smoke when they're at home rather than at work, or vice versa. How are the people in this group supposed to answer the question?

Had the questionnaire's author considered the assumptions on which the question was predicated, the author might first have asked questions such as the following:

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.*** Consider this 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 the respondent 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 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 they believe will be socially acceptable, rather than truly accurate ones, in order to present a favorable impression (more about this issue later in the chapter). To allow for such a possibility, you may want to ask the same question two or more times—using different words each time—at various points in your questionnaire. For example, consider the following two items, appearing in a questionnaire as Items 2 and 30. (Their distance from each other increases the likelihood that participants will answer the second without recalling how they answered the first.) Notice how one individual has answered them:

2. Check one of the following choices:
   __X__ 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 be eager to be among the first to try them.
   __X__ 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 chooses the option generally thought to be moderately conservative rather than liberal. Such an inconsistency might lead you to question whether the respondent really *is* a liberal thinker or only wants to be seen as one.

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 a willingness to try new teaching techniques and in other ways take

reasonable risks in their instructional strategies. 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).

When several items assess the same characteristic—*and* when the responses can reasonably be presumed to reflect an interval (rather than ordinal) measurement scale—responses to those items might be combined into a single score. But a researcher who uses a counterbalancing approach can't 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 wants 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 are consistent—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.[5]

Ideally, preliminary data on internal consistency reliability are 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 questions are clear and will effectively solicit the desired information. At a minimum,

[5]Two 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 multinumber rating scales, such as the 5-point scale for the risk-taking items).

you should give your questionnaire to several friends or colleagues to see whether they have difficulty understanding any items. Have them actually complete 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're likely to get and, by making appropriate revisions, can increase the chances that, in your actual study, the data you obtain will be of sufficient quality to help you address your overall research problem or 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 *under*rate their abilities (Chipman, 2005; Lundeberg & 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're thinking as they read and respond to particular items—is one step toward determining whether a questionnaire has validity for its purpose. Some academic disciplines (e.g., psychology and related fields) insist that a researcher use additional strategies to determine a questionnaire's validity, especially when the questionnaire is intended to assess complex, intangible constructs (e.g., personality, motivation, attitudes). We refer you to the section "Determining and Potentially Enhancing the Validity of an Assessment Instrument" in Chapter 4 for a refresher on four other strategies that might be relevant to your purpose: creating a table of specifications, taking a multitrait–multimethod approach, striving for consistency with a particular conceptual framework, and consulting with a panel of experts.

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, objectivity, relevance, and probability of favorable reception and return. Have you mentally put yourself in the shoes of a person who might be asked to invest time on your behalf? If you received such a questionnaire from a stranger, would *you* agree to complete it? These questions are important and should be answered impartially.

Above all, you should make sure that *every question is essential for you to address your research problem.* A table similar to Table 6.1 can help you evaluate 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), put each questionnaire item in a box 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, throw it out!

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 the research effort seriously and has high regard for prospective participants.

TABLE 6.1 Guide for the Construction and Evaluation of Questionnaire Items

Write your questions in the spaces below.	*Why are you asking this question? How does it relate to your research problem?*
(Create as many rows as you need)	

GUIDELINES Using Technology to Facilitate Questionnaire Administration and Data Analysis

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. One possibility is to use a dedicated website both to recruit participants and to gather their responses to survey questions; we address this strategy in a Practical Application feature a bit later in the chapter. Following are several additional suggestions for using technology to make the use of a questionnaire more efficient and cost-effective.

1. ***When participants are in the same location that you are, have them respond to the questionnaire directly on a laptop or tablet.*** Electronic questionnaires can be highly effective if participants feel comfortable with computer technology. When participants enter their responses directly into 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 problem. One recent study with college students indicated that a tablet-based approach yielded acceptance rates, completion rates, response lengths, and internal consistency similar to those for their paper-and-pencil equivalents (Ravert, Gomez-Scott, & Donnellan, 2015).

2. ***When participants are at diverse locations, use e-mail to request participation and obtain participants' responses.*** If the people you want to survey have easily obtainable e-mail addresses and are regularly online, 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.

3. ***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.

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 items, have yes-or-no answers, or involve 5-point rating scales. You will probably want 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. ***Use a computer database to keep track of who has responded and who has not.*** An electronic spreadsheet or other database software program provides an easy way of keeping track of people's names and addresses, along with information regarding (a) which individuals have and have not yet received your request for participation, (b) which ones have and have not responded to your request, and (c) which ones need a first or second reminder letter or e-mail message. Also, many spreadsheet programs include templates for printing mailing labels.

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, perhaps, 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 thus 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 writing reports. 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, but it was *not* convenient for the people to whom she sent the questionnaire. Her response rate—and her study!—suffered as a result. So consider the characteristics of the sample you're 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's 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 an impression of uncluttered ease? Are the places for responses adequate and clearly indicated? Is the tone courteous, and is the request reasonable?

3. ***Motivate potential respondents.*** Give people a reason to *want* to respond. If you're especially fortunate, you 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 or e-mail message that accompanies your questionnaire.

One potentially effective strategy is to send a letter or e-mail message soliciting people's cooperation *before* actually sending them the questionnaire. For example, Figure 6.5 shows an example of a letter that a researcher might use to gain people's cooperation in responding to a paper-and-pencil 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, a cover letter is written on the institution's official letterhead stationery. Alternatively, an e-mail request for participation might include an eye-catching banner with the institution's name and logo.
- Rather than saying "Dear Sir or Madam," the letter is personalized for the recipient.

FIGURE 6.5 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 15 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 Handwritten Signature

Your Name

FIGURE 6.6 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

- The letter describes the potential value of the study, both for the individual and for alumni in general, hence giving a potential respondent a reason to *want* to respond.
- The letter assures the individual that his or her cooperation will not place an 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, such as that shown in Figure 6.6—a quick and easy first step for potential participants—the researcher gains people's commitment to completing 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 the event that a potential respondent 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 6.5 with the brief note in Figure 6.7 that was sent to one of us authors and that, unfortunately, is all too typical of students' first attempts at drafting a cover letter or e-mail request. A focus only on the researcher's needs in an introduction of this sort may be another reason for the poor return rate of questionnaires in some research projects.

The cover letter or e-mail request 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.

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. At either the beginning or end of your questionnaire, you might provide a box to check to indicate the desire for a summary, together with places for a name and either a postal or e-mail

FIGURE 6.7 An Unenticing Request for Participation in a Research Study

X Y Z UNIVERSITY
Campus Station

Dear Sir:

I am a graduate student at X Y Z University, and the enclosed questionnaire is sent to you in the hope that you will assist me in obtaining information for my master's thesis.

I would appreciate your early reply since I am attempting to get my degree this June.

Yours truly,

John Doe

John Doe

address. If anonymity is important, a mailed questionnaire might include a separate, stamped 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.*** Many experts 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?

Some Internet websites designed to facilitate the use and analysis of online questionnaires include features that enable follow-up requests that ensure potential participants' anonymity (we say more about these websites shortly). To address the problem in a mailed paper-and-pencil questionnaire, 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's time to send a follow-up letter, they send it only to 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 6.5 and 6.6. 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 she didn't receive a reply within 3 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 6.8.

FIGURE 6.8 ■ 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 15 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 Handwritten 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, especially in questionnaire studies.

MyLab Education Self-Check 6.3

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

In recent years, some researchers have collected descriptive 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, maintained by John Krantz, Professor of Psychology at Hanover College (psych.hanover.edu/Research/exponnet.html). As this edition of the book goes to press, the site is hosting research projects on such diverse topics as aggression in intimate relationships, digital games as possible methods of instruction, critical thinking skills in college students, and people's moral reasoning regarding the human rights of refugees. Dr. Krantz checks to be sure that each project has been approved by the appropriate institutional review board (IRB) and incorporates informed consent procedures. There is no fee for using the site.

Commercial websites for data collection are available as well. Probably the most popular one is SurveyMonkey (surveymonkey.com), which charges a modest monthly fee. This website provides templates that make questionnaire design easy and enable a researcher to present a variety of item types (e.g., multiple-choice items, rating scales). Commercial data-collection websites typically also include features for (a) communicating with a preselected sample of participants (e.g., through e-mail invitations), (b) ensuring participants' anonymity, (c) statistically analyzing the data collected, and (d) downloading the results.

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 should be just as rigorous as they would be if you were collecting data through face-to-face contacts or the postal service. In some instances—more specifically, when your IRB determines that your data collection method will pose minimal, if any, risk to participants—the board will waive the need to present an informed consent form. Otherwise, participants should 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. Also, participants' responses should remain as confidential as they would in any research study. The *protection from harm* ethical standard can be especially troublesome in an online study, because it may be virtually impossible to discover that a participant has found a task or question extremely stressful or upsetting and needs some sort of follow-up intervention. If you are currently a university student, your research advisor and university's IRB 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 enables a researcher to zero in on a predetermined sample 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 won't be representative either of a particular group of people or of the overall population of human beings. 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 an issue for any researcher, but it is especially so for a researcher who wants to draw inferences about a large population. In the following section, we look at strategies for selecting an appropriate sample.

MyLab Education Self-Check 6.4

MyLab Education Application Exercise 6.2: Developing Measurement Instruments

CHOOSING A SAMPLE IN A DESCRIPTIVE STUDY

Any 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 senior citizens, 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 senior citizens, elementary school teachers, or national park visitors, we are talking about *very large* groups of people; for example, more than 4½ million people visited Rocky Mountain National Park in 2016.

In such situations, researchers typically don't study the entire population of interest. Instead, they select a subset, or **sample**, of the population. However, they can use the results obtained from their sample to make generalizations about the entire population only if *the sample is truly representative of the population.* Here we're talking about a research study's *external validity*—quantitative researchers' term for the extent to which the study's findings can be reasonably generalized beyond the specific participants in the study (see Chapter 4).

When stating 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.

Based on this purpose study alone, we must conclude that the researcher proposes to survey "the attorneys"—all of them! The American Bar Association consists of approximately 400,000 attorneys distributed over more than 3.5 million square miles. Surveying all of them would be a gargantuan undertaking.

A researcher who intends to survey only a subset of a population should say so, perhaps by using such qualifying words as *selected, representative, typical, certain,* or *a random sample of.* For example, the researcher who wants to study the philosophical perspectives of American Bar Association members might begin the problem statement by saying, "The purpose of this research is to survey the legal philosophies of a random sample of attorneys . . . ." Careful researchers say precisely what they mean.

The specific sampling procedure used depends on both the purpose of the study and parameters of the population being studied. Ideally, *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're 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. We discuss this and other possible sources of bias later in the chapter.

Sampling Designs

Different sampling designs may be more or less appropriate in different situations and for different research questions. Here we consider eight approaches to sampling, which fall into two major categories: probability sampling and nonprobability sampling.

Probability Sampling

In **probability sampling**, the sample is chosen from the overall population by *random selection*—that is, it's selected in such a way that each member of the population has an equal chance of being chosen. 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 two, we should be able to scoop up a small amount 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 ethnic background, education level, social standing, wealth, 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 the same characteristics that exist in the larger population, and we will find them in roughly the same proportions.

There are many possible methods of choosing a random sample. For example, we could assign each person in the population a unique 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 and Internet websites also provide means of picking random numbers (e.g., search for "random number generator").

A popular paper-and-pencil method of selecting a random sample is to use a **table of random numbers**, which you can easily find on the Internet and in many statistics textbooks. Figure 6.9 presents an excerpt from such a table. Typically, a table of random numbers includes blocks of digits that can be identified by specific row and column numbers. For instance, the excerpt in Figure 6.9 shows 25 blocks, each of which includes 50 digits arranged in pairs. Each 50-digit block can be identified by both a row number (shown at the very left) and a column number (shown at the very top). To ensure a truly random sample, the researcher identifies a starting point in the table *randomly*.

How might we identify a starting entry number? Pull a dollar bill from your wallet. The one we pull as we write this book has the serial number L45391827A. We choose the first 2 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, as shown in Figure 6.9.

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 license plate, the last two digits of a friend's social security number, or the stock quotations page in a newspaper. Not all of these suggested sources

FIGURE 6.9 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

reflect strictly random numbers; instead, some numbers may appear 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.

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 2-digit numbers. If it is to be more than 99 but fewer than 1,000, we will need 3 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 2-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'll assume that the total population consists of 90 individuals from which we will select a sample of 25; thus, we'll need random numbers of 2 digits each. Beginning in the upper left-hand corner of the designated block, 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) doesn't apply because there are only 90 people altogether. Our next choice is 52, we ignore 93, and then we choose 80. Going on to the subsequent column, we choose 88, 76, 02, 07, and 47. We continue with additional columns to the right, ignoring the numbers 91–99, 00, and any numbers we've already selected, until we get a 25-person sample.

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 one summer (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 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 thus we would take equal samples from each of the three grades. Our sampling method would look like that in Figure 6.10.

Stratified random sampling has the advantage of guaranteeing equal representation of each of the identified strata. It is most appropriate when the strata are roughly equal in size in the overall population.

Proportional Stratified Sampling Proportional stratified sampling is appropriate when various strata are *different* in size. For example, imagine a small town that has 1,000 Jewish residents, 2,000 Catholics, and 3,000 Protestants. (For purposes of this example, we ignore the fact that the town almost certainly has members of other religious groups as well, plus, of course, people with no religious affiliation.) A local newspaper publishes a section dealing with interfaith news, religious events, and syndicated articles of interest to the religious community in general. The editor decides to conduct a survey in order to obtain certain information and opinions from the paper's readers.

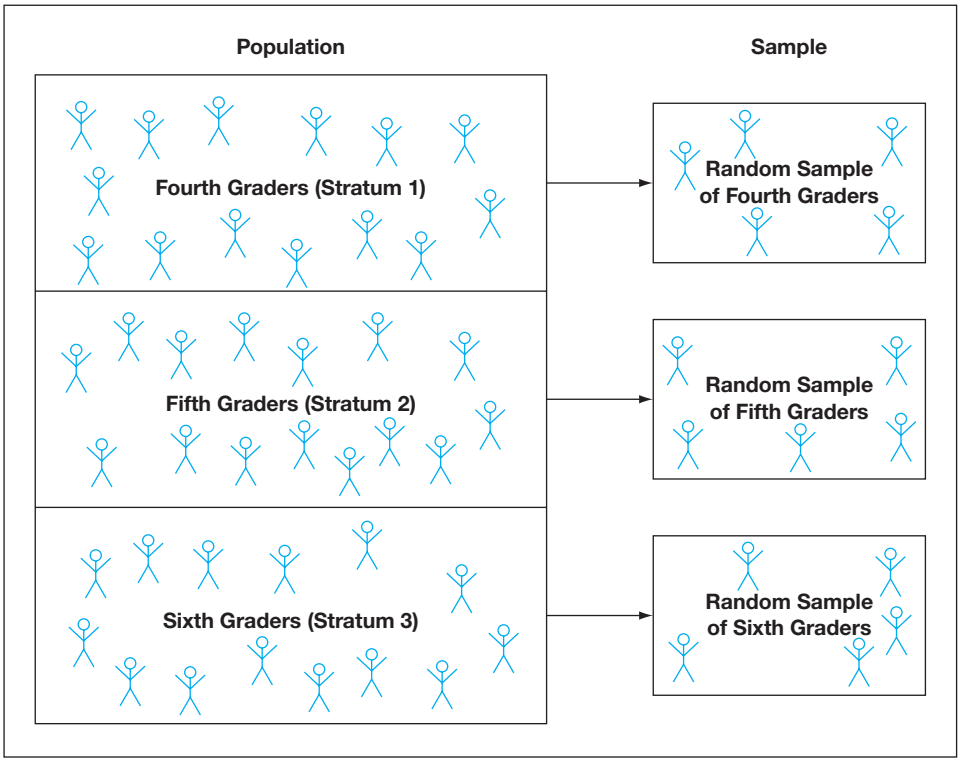
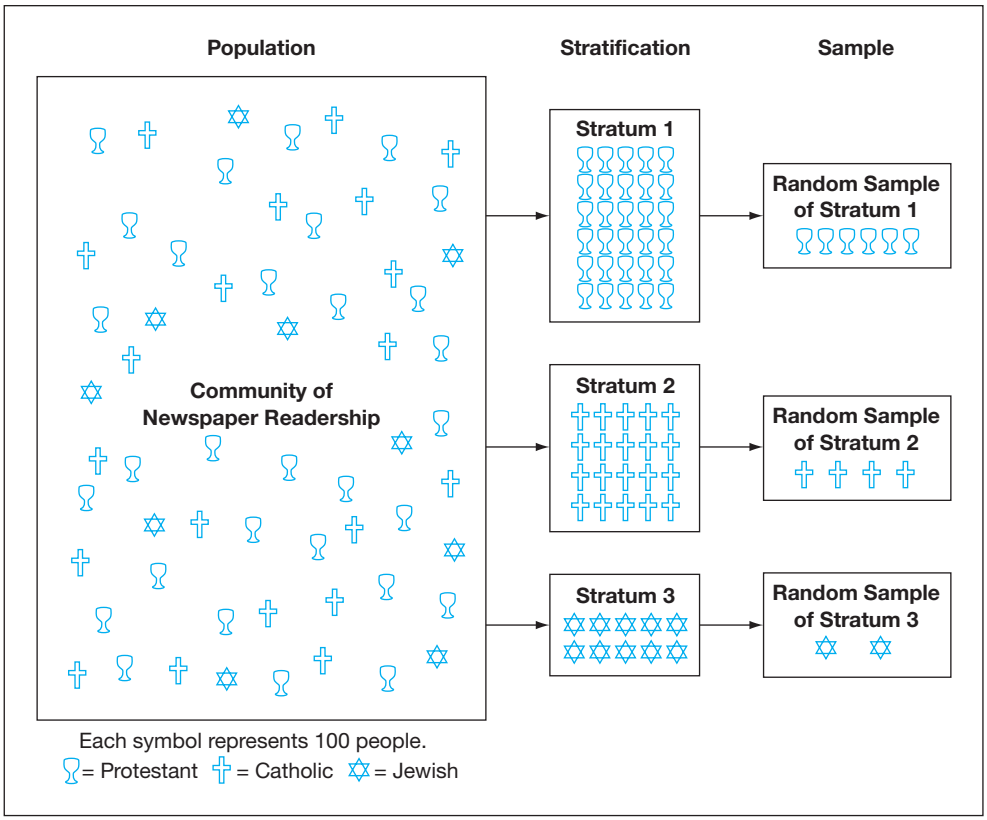


FIGURE 6.10 An Example of Stratified Random Sampling

FIGURE 6.11 An Example of Proportional Stratified Sampling



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. 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 6.11 represents this type of sampling.

Cluster Sampling Sometimes the population of interest is spread over a large area, such that it isn't feasible to make a list of every population member. Instead, we might obtain a map of the area showing political boundaries or other subdivisions. We can then subdivide the area into smaller units, or *clusters*—perhaps precincts, school districts, or counties. In cluster sampling, clusters should ideally be similar to one another in makeup, with each cluster containing an equally heterogeneous mix of individuals.

A subset of the clusters is randomly selected, and the members of these clusters comprise our sample. For example, imagine that we want to learn the opinions of Jewish, Catholic, and Protestant residents in a fairly large community. We might divide the community 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 6.12.

Systematic Sampling Systematic sampling involves choosing individuals—or perhaps clusters—according to a predetermined sequence, with the sequence being determined 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 return to the 12 clusters shown in Figure 6.12. Half of the cell numbers are odd, and the other half are even. Using a systematic sampling approach, 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

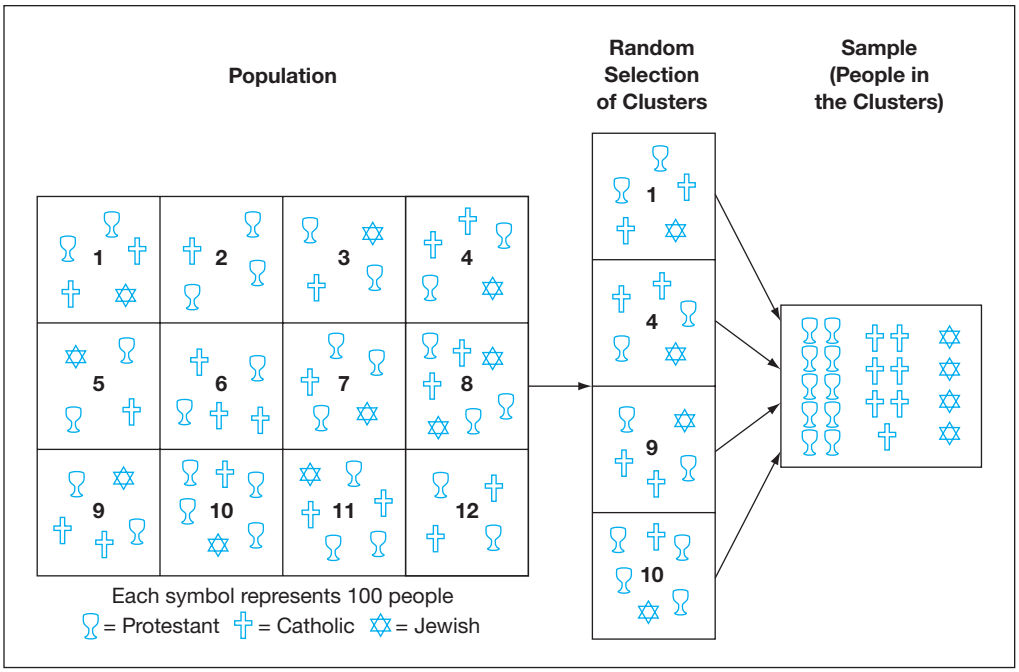


FIGURE 6.12 An Example of Cluster Sampling

first even-numbered digit, which is 2, and select the systematically sequential clusters 4, 6, 8, 10, 12. Figure 6.13 illustrates this process.

Each of the sampling designs just described is uniquely suited to a particular kind of population; thus, you should consider the nature of your population when selecting your sampling technique. Table 6.2 identifies the various kinds of populations for which different probability sampling techniques might be appropriate.

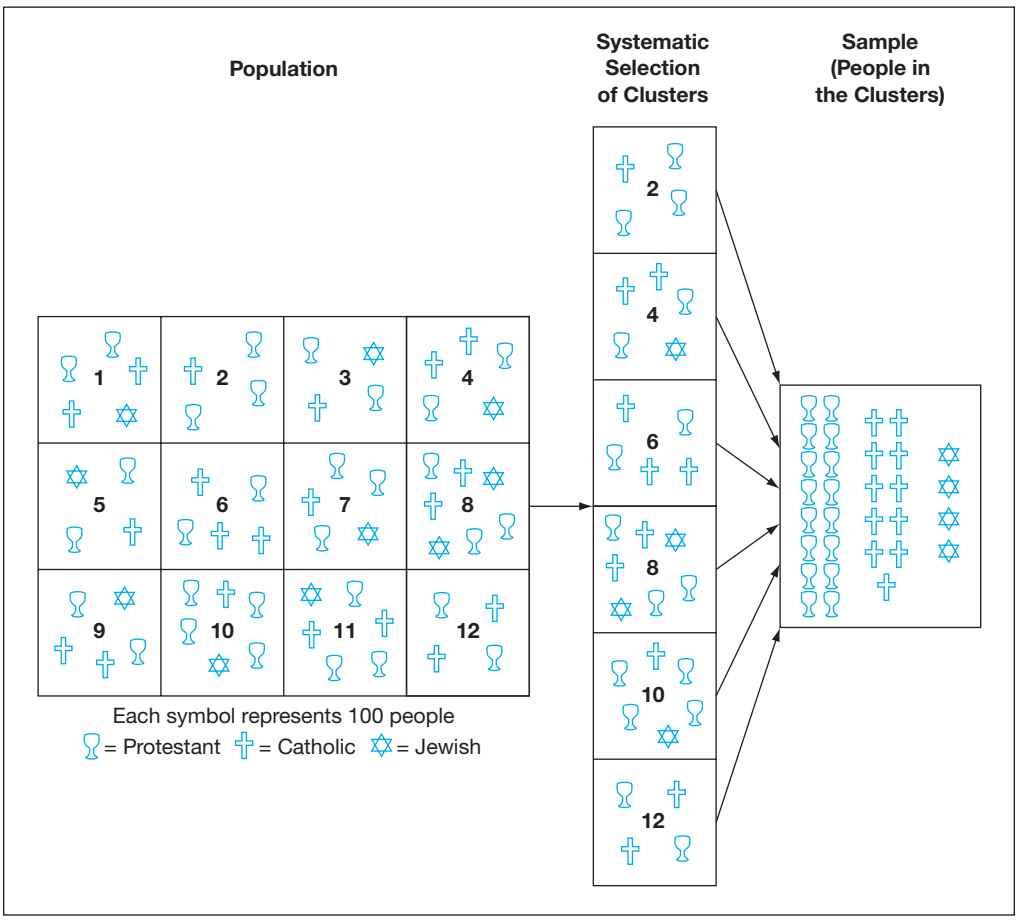


FIGURE 6.13 An Example of Systematic Sampling

TABLE 6.2 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 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 (30%), Protestant (30%), Jewish (15%), Muslim (5%), other religious group (5%), or unaffiliated (15%)	• 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 15 leading airline terminals (It is assumed that all 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 presumed to be 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—sometimes called *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 arriving on the scene by mere happenstance.

A good example of convenience sampling is the use of college students to study and draw conclusions about characteristics of people in general—perhaps people's reactions to various kinds of television commercials, their ability to think "outside the box" when trying to solve challenging problems, or their willingness to change their opinions when confronted with credible contradictory evidence. College students hardly provide a representative sample of *all* human beings, or even of, say, citizens of a particular country, but they are certainly handy research participants for professors, graduate students, and other members of a university community. Furthermore, the vast majority of them are legally considered to be adults; typically, then, they can sign informed consent forms for themselves.

A high-tech alternative to on-site convenience sampling is using one or more social media sites to recruit participants for a research study. For example, in the social sciences, researchers are increasingly using Facebook to convince people to complete a survey or make their posts available for inspection and systematic analysis. This approach is certainly a form of convenience sampling in the sense that, thanks to the Internet, even people who live far distances from the researcher are readily available for study. But the sample obtained could hardly be construed as adequately representing a larger population that includes people without Facebook accounts. And remember, many posts and other information available through social media are *not* truly public information; for instance, most Facebook posts are theoretically available only to people who have officially been identified as "friends." We refer you to an article by Kosinski, Matz, Gosling, Popov, and Stillwell (2015) for guidance about using Facebook both effectively and ethically.

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. Suppose, for example, that you're 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. You would be doing essentially the same thing if, of all the individuals you recruited via Facebook to complete a survey, you used data only for the first 20 African Americans and 20 European Americans who submitted responses, with racial-group membership determined by one or more demographic questions in your survey.

Quota sampling regulates only the size of each group within the sample. In every other respect, the selection of the sample is nonrandom and, in most cases, convenient.

Purposive Sampling This particular sampling strategy is used not only in some quantitative descriptive studies but also in most qualitative studies and many mixed-methods studies (see Chapter 8 and Chapter 9, respectively). Thus, we include it as a key term here, complete with an entry in the book's glossary. Regardless of research methodology, **purposive sampling** involves choosing participants or other units of study, as the name implies, for a particular *purpose.* For instance, we might choose people whom we've determined to be "typical" of a group or, instead, 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 helpful in predicting the final outcomes.

Purposive sampling can be quite appropriate for certain kinds of research problems and questions. However, a researcher should always provide a rationale for the selection criteria used.

Sampling in Surveys of Very Large Populations

Nowhere is sampling more critical than in surveys of large populations. Sometimes a researcher reports that *x%* of people believe such-and-such, that *y%* do so-and-so, or that *z%* 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.*

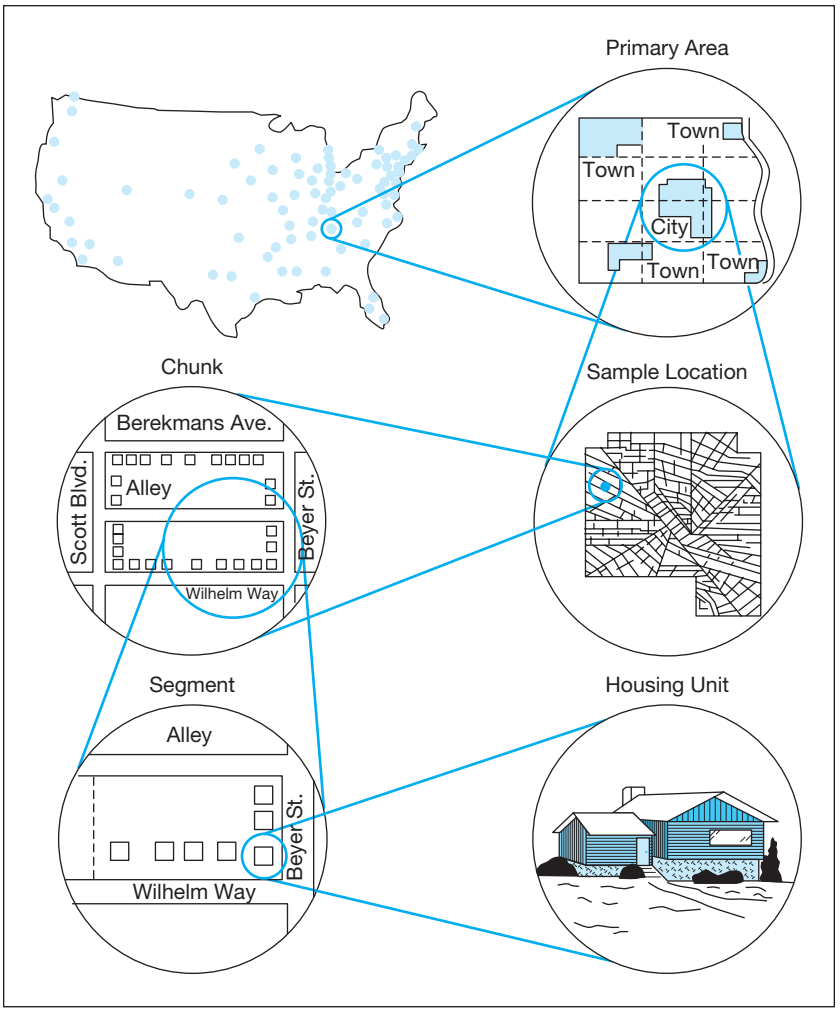
But now imagine that a researcher wants to conduct a survey of a country's *entire adult population.* How can the researcher possibly hope to get a random, representative sample of such a large group of people? The Survey Research Center of the University of Michigan's Institute for Social Research has successfully used a *multistage sampling of areas*, described in its now-classic *Interviewer's Manual* (1976):

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.[6] A predetermined number of these areas is 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 or knock on the door is unanswered, the researcher returns at a later date and tries again.

[6] You should note that the manual has been written for surveys conducted in the United States, where every state is divided into counties. Researchers who are surveying people in "county-less" nations can use other existing political units that are fairly uniform in size.

FIGURE 6.14
Multistage Sampling

Source: 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.



As you may have deduced, the approach just described is a multistage version of cluster sampling (see Figure 6.14). 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.

PRACTICAL APPLICATION Identifying a Sufficient Sample Size

A basic rule in sampling is this one: *The larger the sample, the better.* But such a generalized rule isn't very helpful to a researcher who must make a practical decision about a specific research situation. Mills and Gay (2016, p. 147) 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.
- 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 is necessary than if the population is fairly homogeneous. Important, too, is the degree of precision with which the researcher wants to draw conclusions or make predictions about the population under study. And in quantitative research, a larger sample size increases the *power* of a statistical analysis; you will learn more about this concept in Chapter 11.

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 and on many Internet websites (e.g., search "determining sample size").

PRACTICAL APPLICATION Choosing an Appropriate Sample in Descriptive Research

Select a particular population, and reflect on its overall composition and characteristics. Then, in the following checklist, analyze the population, and identify a defensible strategy for obtaining a sample from this population.

CHECKLIST

Considering the Nature of the Population When Identifying Your Sampling Procedure

_____ 1. On the following line, describe the particular population you have chosen:

_____

_____ 2. Now answer the following questions with respect to the *composition* 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?	____	____
e. Is the population so diverse that its members cannot easily be categorized into layers or clusters?	____	____

_____ 3. Can you identify a feasible means of drawing a *representative sample* from this population? If so, describe your procedure on the following lines:

_____

_____

_____

	YES	NO
a. Refer to Table 6.2. Is your sampling procedure appropriate for the characteristics of the population?	____	____
b. Will your procedure ensure that your sample will be chosen by chance?	____	____

_____ 4. If nonprobability sampling is more appropriate for studying the population, what specific procedure should you use?

_____ Convenience sampling at one or more physical locations

_____ Convenience sampling through social media

_____ Quota sampling (from either on-site or media-based recruiting)

_____ Purposive sampling

_____ Other (explain): _____

_____ 5. What are the strengths of the sampling procedure you have identified?

_____

_____

_____ 6. What are the weaknesses inherent in this sampling procedure?

_____

_____

MyLab Education Self-Check 6.5

MyLab Education Application Exercise 6.3: Selecting a Sampling Technique

COMMON SOURCES OF BIAS IN DESCRIPTIVE STUDIES

In this and preceding chapters, we have occasionally mentioned that a particular research strategy might in some way *bias* the results. In general, **bias** in a research study is any influential factor that distorts the data obtained or conclusions drawn. Bias can creep into a research project in a variety of subtle ways. For example, when conducting an interview, a researcher's tone of voice in asking questions might predispose a participant to respond in one way rather than in another, or the researcher's personality might influence a participant's willingness to reveal embarrassing facts.

Most sources of bias in descriptive research fall into one of four categories, each of which we examine now.

Sampling Bias

A key source of bias in many descriptive studies is **sampling bias**—any factor that yields a nonrepresentative sample of the population being studied. For example, imagine that a researcher wants to conduct a telephone survey of people living in a small town in the

United States—more specifically, a certain town in the state of New Hampshire. The researcher first needs to find out which one or more exchange numbers are used exclusively in that town's telephone numbers—that is, which three-digit series might appear after New Hampshire's area code (603) (e.g., 1-603-###-5555). Looking on the Internet, she finds only one exchange number used for the town's landline and mobile phones: 211.[7] Then, being an especially persistent researcher, she dials every conceivable phone number for this exchange; this would be 10,000 attempted phone calls altogether. When a dialed number reaches an actual person, she tries to gain the person's informed consent and then administer her survey. Would her strategy yield a representative sample of the town's residents? No, not at all! The researcher's survey data would be obtained only from people who (a) own either a landline or a mobile phone (this criterion would eliminate many very-low-income and homeless individuals), (b) are currently near their phones when she calls, (c) haven't become so annoyed with telephone surveys and automated "spam" calls that they don't answer their phones unless they know who is calling, or (d) refuse to participate once they know the reason for the call. And the researcher hasn't accounted for (e) town residents who purchased mobile phones in a different town (and thus have an exchange number other than 211) or (f) people who bought their mobile phones in the town but don't currently live there.

Likewise, as noted in earlier sections of the chapter, studies involving online interviews or Internet-based questionnaires are apt to be biased—this time in favor of computer-literate individuals with easy access to the Internet. Studies involving mailed questionnaires frequently fall victim to bias as well. For example, suppose that a questionnaire is sent to 100 working adults in the United States, 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 say that they have been audited, whereas 35 are from people who respond that they have never been audited. The researcher might therefore conclude that 50% of U.S. taxpayers are likely to be audited by the IRS at one time or another.

The researcher's generalization isn't necessarily 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 honesty and moral integrity. Perhaps for this reason, some individuals in the researcher's sample may not have wanted to admit to an audit and therefore discarded the questionnaire. 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 facts. Perhaps, instead of a 50-50 split, an estimate of 60% (people audited) versus 40% (people not audited) is more accurate. The data the researcher has obtained don't enable the researcher to make such an estimate, however.

The examples just presented illustrate various ways in which bias can creep into the research sample. In the cases of telephone- and Internet-based data collection, *sample selection* itself was biased because not everyone in the population had an equal chance of being selected; some 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, either physically or on the Internet, 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, the return rate in a questionnaire study is 50% or less, and the greater the percentage of *non*respondents, 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.

Nonrespondents to mailed questionnaires might be different from respondents in a variety of ways (Rogelberg & Luong, 1998). They may have illnesses, disabilities, or language barriers

[7]As this book goes to press in 2017, an Internet search tells us that the exchange number 211 isn't actually used in the state of New Hampshire, hence our use of it in this hypothetical example.

that prevent them from responding, and, on average, they have lower educational levels. In contrast to sampling bias in mailed surveys, many people who are hard to reach *by phone* are young working adults who are *more* educated than the average individual (Witt & Best, 2008).

Even when potential participants' ages, health, language skills, educational levels, and computer literacy are similar, they can differ widely in their *motivation* to participate in a study: Some might have other priorities, and some might worry that a researcher has sinister intentions. Participants in longitudinal studies may eventually grow weary of being "bothered" time after time. Also, a nonrandom subset of them might die before the study is completed.

We return to Step 5 in the University of Michigan's Survey Research Center's procedure for obtaining a sample in a very large population: "If a doorbell or knock on the door 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)

Instrumentation Bias

By **instrumentation bias**, we mean the ways in which particular assessment strategies slant the obtained results in one direction or another. For instance, in our earlier discussion of questionnaires, we mentioned that a researcher must choose certain questions—and by default must omit *other* questions. The same is true of structured interviews: By virtue of the questions asked, participants are encouraged to reflect on and talk about some topics rather than other ones. The outcome is that some variables are included in a study, and other potentially important variables are overlooked.

As an example, imagine that an educational researcher is interested in discovering the kinds of goals that students hope to achieve when they're at school. Many motivation researchers have speculated that students might be concerned about either (a) truly mastering classroom subject matter, on the one hand, or (b) getting good grades by any expedient means, on the other. Accordingly, they have designed and administered rating-scale questionnaires with such items as "I work hard to understand new ideas" (reflecting a desire to master a topic) and "I occasionally copy someone else's homework if I don't have time to do it myself" (reflecting a desire to get good grades). But in one study with middle school students (Dowson & McInerney, 2001), researchers instead asked open-ended questions related to what things students thought were most important for them to accomplish at school. Many students focused not on a desire to do well academically but instead on *social* goals, such as being with and helping classmates and avoiding behaviors that might adversely affect their popularity.

Response Bias

Whenever we gather data through interviews or questionnaires, 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. To the extent that people describe their thoughts, beliefs, and experiences inaccurately, **response bias** is at work. For example, people's descriptions of their attitudes, opinions, and motives 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 thus may be colored by recent events, the current context, or flawed self-perceptions (McCaslin, Vega, Anderson, Calderon, & Labistre, 2011; Schwarz, 1999). Furthermore, some participants may intentionally or unintentionally misrepresent the facts in order to give a favorable impression—

a source of bias known as a **social desirability effect** (e.g., Uziel, 2010). 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. And when we ask people about *past* events, behaviors, and perspectives, interviewees must rely on their memories, and human memory is rarely as accurate as a video recorder might be. 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., Schwarz, 1999; Wheelan, 2013). The experience-sampling methods (ESMs) described earlier in the chapter provide one means of minimizing memory-based distortions.

Researcher Bias

Finally, we must not overlook the potential effects of a researcher's expectations, values, and general belief systems, which can predispose the researcher to study certain variables and not other variables, as well as to draw certain conclusions and not other conclusions. For example, recall the discussion of philosophical assumptions in Chapter 1: Researchers with a *positivist* outlook are more likely to look for cause-and-effect relationships—sometimes even from correlational studies that don't warrant conclusions about cause and effect!—than postpositivists or constructivists.

Ultimately, we must remember that *no human being can be completely objective.* Assigning numbers to observations helps a researcher quantify data, but it doesn't necessarily make the researcher any more objective in collecting or interpreting those data.

PRACTICAL APPLICATION Acknowledging the Probable Presence of Bias in Descriptive Research

When conducting research, it's almost impossible to avoid biases of one sort or another—biases that can potentially influence the data collected, the conclusions drawn, or both. Good researchers demonstrate their integrity by admitting, without reservation, that certain biases may well have influenced their findings. For example, 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. Here we offer guidelines for identifying possible sampling biases in questionnaire research. We then provide a checklist that can help you pin down various biases that can potentially contaminate descriptive studies of all sorts.

GUIDELINES Identifying Possible Sampling Bias in Questionnaire Research

Rogelberg and Luong (1998) have suggested several strategies for identifying possible bias in questionnaire research. Following are three especially useful ones.

1. ***Carefully scrutinize the questionnaire for items that might be influenced by factors that frequently distinguish respondents from nonrespondents.*** For example, ask yourself questions such as these:

- Might some people be more interested in this topic than others? If so, would their interest level affect their responses?
- How much might people's language and literacy skills influence their ability and willingness to respond?

- Are people with high education levels likely to respond differently to certain questions than people with less education?
- Might younger people respond differently than older ones?
- Might people with full-time jobs respond differently than people who are retired and unemployed? (Fully employed individuals may have little or no free time to complete questionnaires, especially if they have young children.)
- Might healthy people respond differently than those who are disabled or chronically ill do? (Healthy people are more likely to have the time and energy to respond.)

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, compare their answers to 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 phone 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.

CHECKLIST

Identifying Potential Sources of Bias in a Descriptive Study

_____ 1. Do you have certain expectations about the results you will obtain and/or the conclusions you are likely to draw? If so, what are they?

_____

_____

_____ 2. Do you have any preconceived notions about cause-and-effect relationships within the phenomenon you are studying? If so, what precautions might you take to ensure that you do *not* infer causal relationships from cross-variable correlations you might find?

_____

_____

_____ 3. How do you plan to identify a sample for your study? What characteristics of that sample might limit your ability to generalize your findings to a larger population?

_____

_____

_____ 4. On what specific qualities and characteristics will you be focusing? What potentially relevant qualities and characteristics will you *not* be looking at? To what degree might omitted variables be as important or more important in helping you understand the phenomenon you are studying?

_____

_____

_____ 5. Might participants' responses be poor indicators of certain characteristics, attitudes, or opinions? For example:

- Might participants say or do things in order to create a favorable impression?
  _____ Yes _____ No
- Might you be asking them questions about topics they haven't really thought about before?
  _____ Yes _____ No
- Will some questions require participants to rely on their memories of past events?
  _____ Yes _____ No

If any of your answers are *yes*, how might such sources of bias influence your findings?

_____

_____

MyLab Education Self-Check 6.6

INTERPRETING DATA IN DESCRIPTIVE RESEARCH

In our discussion of descriptive research methods in this chapter, we have focused largely on strategies for acquiring data. But 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 or question in light of data that relate to the problem or question.
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 paramount and should never be neglected.

A descriptive study is often a very "busy" research method. The researcher must decide on a population; choose a technique for sampling it; identify or create one or more valid means of collecting desired information; minimize the potential for bias in the study; and then actually collect, record, organize, and analyze all the necessary data. These activities can be so complex and time-consuming that the researcher eventually loses sight of the overall research problem or question that instigated the research effort.

Activity for activity's sake can be seductive and emotionally satisfying, especially when it involves collecting large bodies of data and presenting them in verbal descriptions, numerical tables, and graphic displays. Yet all research activity is subordinate to the research problem or question itself. 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's 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 you need to address your research problem and its subproblems?

- What procedures should you use to obtain the needed information? How can you best implement those procedures?
- If applicable, how can you get a sample that will reasonably represent the overall population about which you are concerned?
- How can you control for possible bias in your collection of data?
- What will you do with the data once you have collected them? How might you best organize and prepare them for analysis?
- Above all, in what ways might you reasonably interpret your data? What conclusions might you reach from your 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 hypothesized 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 hypothesized 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 3

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 are the extent of criminal history and its relation 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—the 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.

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 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:

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.

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.

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

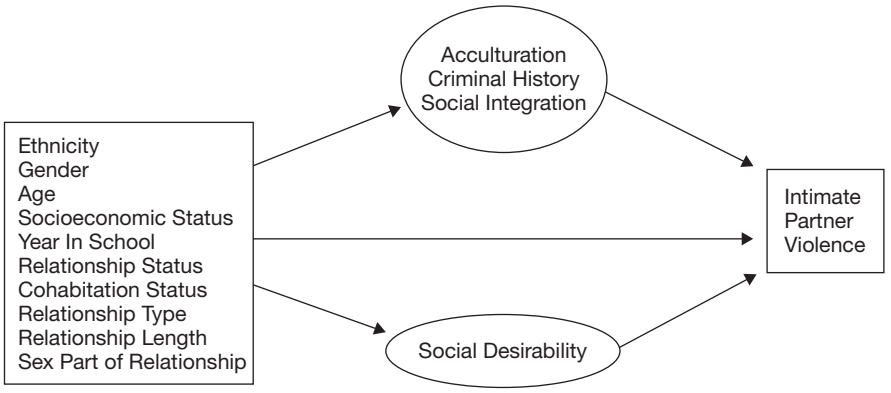


Figure 1.1 effectively condenses and summarizes the researcher's hypotheses. Also, it graphically demonstrates that four variables—acculturation, criminal history, social integration, and social desirability—are hypothesized to be mediating variables *in the relationship between demographics and violence.*

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.

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.

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.

The author, whose home town is El Paso, Texas, has numerous acquaintances at both institutions and so can easily gain access to these students. He must, of course, seek approval from the institutional review boards at the two institutions, as well as at the institution where he is completing his doctorate.

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 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.

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 final sample size (348), this approach is reasonable.

A debriefing form was given to each participant as he or she turned in his or her 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 who voluntarily participated in the study were offered extra credit points by their professors.

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.

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.

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: 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.

APPLICATION EXERCISES

Similar exercises are available on Pearson MyLab Education.

1. A school district is interested in tracking its graduates' satisfaction with career training and the information they receive about postgraduate employment and education. Specifically, the district is interested in getting information about the following:
   - Satisfaction with career guidance counseling
   - Satisfaction with vocational and technical training
   - Importance of different educational activities in high school
   - Educational and career opportunities explored upon graduation

   i. Identify a way to study this issue as a cross-sectional survey study.
   ii. Identify a way to study this issue as a longitudinal survey study.
   iii. Identify a way to study this issue as a correlational study.

2. A local aeronautics firm offers a summer internship program for college students majoring in astrophysics, engineering, and other natural sciences. The program director wants to know if the summer program leads to student growth, both

personal and professional, and decides to gather data at the start of the summer and then again at the end of the summer.

i. What kind of study is the program director conducting?
ii. The program director identifies a number of questions that he'd like to ask, including whether students feel they are good scientists, whether this opportunity helped them to learn, and what their personal lives are like outside the lab. Given his general topics, do you recommend that he conduct face-to-face interviews or give out questionnaires? Explain your answer.
iii. The program director would like students to complete a rating scale as part of his interview or questionnaire, and he proposes the following:
Please rate the following statements about personal growth on a scale from 1 (strongly disagree) to 7 (strongly agree).
a. I usually have the patience to deal with the frustration that arises when experiments fail.
b. I have reduced the number of glasses of wine I drink each week.
c. I am sure that my research area is something I want to continue to pursue during graduate study.
d. I have some doubts about whether my scientific skills are good enough.
Review each of his questions for clarity and potential bias.

3. For each of the following situations, describe an appropriate procedure for selecting a sample, and explain how to choose a sample using that procedure.

i. There are 100 kindergarten classes in the school system, with an average of 20 children per class. You want a sample of 200 children.
ii. There are 220 attorneys in the local bar association, and you want a random sample of 40 participants.
iii. There are 2,400 registered voters in the county, and there are approximately twice as many Republicans as Democrats. You want a sample of voters, and you want to ensure proportional representation.

FOR FURTHER READING

Alreck, P. L., & Settle, R. B. (2003). *The survey research handbook* (3rd ed.). New York, NY: McGraw-Hill.

Bourque, L. B., & Fielder, E. P. (Eds.). (2002a). *How to conduct self-administered and mail surveys* (2nd ed.) Los Angeles, CA: Sage.

Bourque, L. B., & Fielder, E. P. (2002b). *How to conduct telephone surveys* (2nd ed.). Los Angeles, CA: Sage.

Daniel, J. (2011). *Sampling essentials: Practical guidelines for making sampling choices.* Los Angeles, CA: Sage.

Delandshere, G., & Petrosky, A. R. (1998). Assessment of complex performances: Limitations of key measurement assumptions. *Educational Researcher, 27,* 14–24.

Fink, A. (2017). *How to conduct surveys: A step-by-step guide* (6th ed.). Los Angeles, CA: Sage.

Fowler, F. J., Jr. (2014). *Survey research methods* (5th ed.). Los Angeles, CA: Sage.

Friedman, H. H., & Amoo, T. (1999). Rating the rating scales. *Journal of Marketing Management, 9*(3), 114–123.

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., Marvasti, A. B., & McKinney, K. D. (Eds.). (2012). *The SAGE handbook of interview research: The complexity of the craft* (2nd ed.). Los Angeles, CA: Sage.

Gwartney, P. A. (2007). *The telephone interviewer's handbook: How to conduct standardized conversations.* San Francisco, CA: Jossey-Bass.

Harari, G. M., Lane, N. D., Wang, R., Crosier, B. S., Campbell, A. T., & Gosling, S. D. (2016). Using smartphones to collect behavioral data in psychological science: Opportunities, practical considerations, and challenges. *Perspectives on Psychological Science, 11,* 838–854.

Henry, G. T. (1990). *Practical sampling.* Los Angeles, CA: Sage.

Laursen, B., Little, T. D., & Card, N. A. (Eds.). (2013). *Handbook of developmental research methods.* New York, NY: Guilford Press.

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.). (2010). *Problems and methods in longitudinal research: Stability and change.* New York, NY: Cambridge University Press.

Marsden, P. V., & Wright, J. D. (Eds.). (2010). *Handbook of survey research* (2nd ed.). Bingley, United Kingdom: Emerald Group.

Menard, S. (Ed.). (2007). *Handbook of longitudinal research: Design, measurement, and analysis.* Burlington, MA: Academic Press.

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.

Saris, W. E., & Galhofer, I. N. (2014). *Design, evaluation, and analysis of questionnaires for survey research.* Hoboken, NJ: Wiley.

Schuman, H., & Presser, S. (1996). *Questions and answers in attitude surveys: Experiments on question form, wording, and context.* Los Angeles, CA: Sage.

Sue, V. M., & Ritter, L. A. (2012). *Conducting online surveys* (2nd ed.). Los Angeles, CA: Sage.

Valliant, R., Dever, J. A., & Kreuter, F. (2013). *Practical tools for designing and weighting survey samples.* New York, NY: Springer.

Zirkel, S., Garcia, J. A., & Murphy, M. C. (2015). Experience-sampling research methods and their potential for education research. *Educational Researcher, 44,* 7–16.

Chapter 7 Experimental, Quasi-Experimental, and Ex Post Facto Designs

We assess progress by noting the amount 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 direct influence on the phenomenon we are studying. To the extent that many potentially causal factors all vary at once in an entangled, confounded manner, we learn little or nothing about what causes what.

Learning Outcomes

7.1 Identify examples of independent and dependent variables, and describe several strategies for controlling for confounding variables in experimental studies.

7.2 Recognize examples of pre-experimental, experimental, quasi-experimental, ex post facto, and factorial designs, as well as meta-analyses.

7.3 Describe potential biases in these various designs and how they might affect a study's internal or external validity.

In the descriptive designs described in the preceding chapter, we make no systematic attempt to determine the underlying causes of the phenomena being studied. But sometimes we *do* want to know what leads to what. We can most convincingly pin down cause-and-effect relationships by using an **experimental design**. In such a design, we consider many possible factors that might cause or influence a particular behavior, characteristic, or other phenomenon. We then try 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, for all intents and purposes, we can 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 might 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 treatment but does *not* change for the other group, and if everything else about the two groups has remained the same in the meantime, we can reasonably conclude that the treatment brought about the change we observed. Because we have systematically *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 for an observed change. Yet all of the designs in the chapter have one thing in common: clearly identifiable independent and dependent variables.

We have previously introduced you to independent and dependent variables in Chapter 2, but because these concepts guide so much of our discussion in this chapter, a brief refresher might be in order. An **independent variable** is one that the researcher studies for its possible effects on one or more other variables. In many of the designs described in this chapter, the researcher directly manipulates and controls at least one independent variable. In contrast, a **dependent variable** is a variable that is potentially influenced by an independent variable; that is, its value *depends* to some degree on the value of the independent variable. In other words, the hypothesized relationship is this:

As an example, let's look at a dissertation written by Nancy Thrailkill (1996), who 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: (a) she described certain historical figures and events in such a manner that they were easy to imagine and visualize (*imagery* condition), (b) she included attention-grabbing phrases in the lecture (*attention* condition), or (c) she did neither of these things (*control* condition). In the following examples from Thrailkill's dissertation, 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 <u>the dimpled, baby-faced young blond</u> 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. <u>LISTEN TO ME NOW.</u> 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 set the less experienced <u>junior officer</u> 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. (Thrailkill, 1996, p. 62, underscores added)

After presenting different parts of the lecture under the three different conditions, Thrailkill assessed the students' recall of 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 that 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 ("control"). 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. Thrailkill'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

In Chapter 4, we stressed the importance of planning and conducting a *credible* study; in other words, the study should be designed and executed in such a way that other people can judge its findings as being reasonably accurate and trustworthy. We also mentioned that many quantitative researchers use the term *internal validity* when referring to such credibility. More specifically, the **internal validity** of a research study is the extent to which its design and the data it yields allow the researcher to draw defensible conclusions about cause-and-effect and other relationships within the data.

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. 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 North America. 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 3 months. "Humor boosts sales," the researcher concludes.
2. 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 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.
3. 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.

Did you detect anything 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 differ from each other in several ways (e.g., the presence of teenagers, the amount of action) in addition to humor. And we shouldn't overlook the fact that the humorous commercial airs during the hot summer months, when people are more likely to consume cold soft drinks.

In the second study, notice that the researcher looks for *volunteers* to use the new method for teaching reading. Are the volunteer teachers different in some way from the nonvolunteers? Are they better educated or more motivated? Do they teach with more enthusiasm and energy because they *expect* the new method to be more effective? Or do the volunteer

teachers happen to be teaching in schools in which children have had a better head start in reading skills before beginning school? Perhaps the children in the volunteers' classrooms earn higher scores on the achievement test not because the instructional method has been more effective, but because, as a group, they have previously been read to more frequently by their parents or have gone to more academically oriented preschools.

The third study uses only one group of participants, so there are no between-group differences to consider. As you will discover later in the chapter, certain kinds of single-group research designs can certainly help a researcher pin down cause-and-effect relationships. But this particular study has many weaknesses. One especially noteworthy problem is the fact that the typists *know* they are participating in a research study; they also know the nature of the researcher's hypothesis. Sometimes the participants in a research study change their behavior simply because they know they're in a research study and are getting extra attention as a result. This effect, known as the **Hawthorne effect**,[1] is an example of *reactivity*, a more general phenomenon previously described in Chapter 4; to refresh your memory, reactivity is a change in people's behavior simply because they're aware that they're being observed. But other explanations for the second study's results are possible as well. Perhaps the typists type more because they like the researcher and want to help him support his hypothesis. Perhaps the music energizes the typists for a few weeks simply because it creates 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 hasn't considered the number of people who were working before and after the music started. Perhaps productivity has increased simply because two people in the typing pool have just returned from vacation!

Whenever we either (a) compare two or more groups that are or might be different in ways *in addition to* the particular treatment or intervention we are studying or (b) assess the performance of only a single group both before and after an intervention, we are likely to have **confounding variables** in a study. Confounding variables can make it extremely difficult to draw conclusions about cause-and-effect relationships because we can't pin down exactly *what* is the cause of any pattern in the data observed after an intervention. Thus, confounding variables threaten a study's internal validity.

In a now-classic book chapter, Campbell and Stanley (1963) identified several potential threats to the internal validity of an experimental study; we describe them in Figure 7.1.

Controlling for Confounding Variables

Researchers use a variety of strategies to control for confounding variables in experimental studies. Following are several common ones.

1. ***Keep some things constant.*** When a factor is the *same* for everyone, it can't possibly account for any differences observed. Oftentimes researchers ensure that various experimental treatments are imposed in the same or very similar environments. They may also seek research participants who share a certain characteristic, such as age, gender, 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 (recall the discussion of this issue in Chapter 4).

2. ***Include a control group.*** Earlier we questioned the internal validity of a study in which an industrial psychologist found a 30% increase in typing productivity during a month in which the typists could hear classical music as they did their work. The seemingly impressive increase may or may not have been due to the classical music. There were

[1]The effect owes its name to the Hawthorne Works, an industrial complex in Illinois where the effect was first observed.

FIGURE 7.1 Potential Threats to the Internal Validity of 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. In a now-classic book chapter, Campbell and Stanley (1963) have identified the following potential threats to internal validity, which can be present either singly or in combination:

1. *History:* At some point after the pre-treatment assessment of the dependent variable but before its post-treatment assessment, an uncontrolled outside event brings about a change in the dependent variable. For example, a noteworthy event in the local community or on the national scene might change participants' knowledge, abilities, or emotional states in ways that affect the post-treatment assessment.
2. *Maturation:* A change in participants' characteristics or abilities is simply the result of the passage of time. For example, children might make normal developmental gains in eye–hand coordination or cognitive abilities.
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 an assessment instrument is administered or scored from one time to the next. For example, a researcher might have one research assistant rate participants' performance before an experimental intervention but have a different research assistant judge their performance after the intervention. Any observed change might be the result of the two assistants' differing standards for rating the performance. (The threat to internal validity just described reflects a problem with *interrater reliability*; see Chapter 4.)
5. *Statistical regression:* In a common phenomenon known as **statistical regression**, people who score extremely high or low on an assessment instrument at one time are likely to score in a less extreme manner on the same instrument at a later time; that is, extreme scorers tend to "drift" toward more average performance during a subsequent assessment. 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 assessment—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 administer one treatment to students enrolled in an 8:00 a.m. class and administer another treatment to students enrolled in a 2:00 p.m. class. 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 a study at proportionally different rates. For example, one group in a study might lose 25% of its members before the final assessment, 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 just listed. 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, for some reason, early risers are more likely to drop out than students who like to sleep in a bit. In this situation, it becomes virtually impossible to disentangle the possible effects of an experimental intervention from the 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 used in contemporary research literature.

too many possible confounding variables—nature of the documents being typed, personnel changes, numbers of people out sick or on vacation during the no-music and music months, even just the knowledge that an experiment was being conducted—that might also have accounted for the 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.

To avoid any group-specific influences that participants' awareness of being in an experiment might have, a researcher sometimes gives the people in a control group a **placebo** that might, on the surface, *appear* to be influential but in reality should *not* be influential. For example, a researcher studying the effects of a new arthritis medication might give some participants a particular dosage of the medicine and give others a 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 adolescents), the researcher must seriously weigh (a) the benefits of the new knowledge that can be gained by a control group receiving no treatment—or perhaps, instead, a less-intensive version of the experimental intervention—against (b) the lives that might 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, your institutional review board, and your conscience.

3. ***Conduct a double-blind experiment.*** In a **double-blind experiment**, participants aren't the only ones who are uninformed about their group membership. Individuals who are administering varying interventions or a control condition (e.g., research assistants, teachers, clinicians) are equally in the dark, thus further decreasing the likelihood that people's expectations for outcomes might influence the *actual* outcomes. For example, in a study in which chronic arthritis sufferers are given either (a) medication that is hypothesized to reduce arthritis symptoms or (b) a similar-looking placebo, neither the participants nor the clinicians distributing the medication should know which people are in which group.

4. ***Randomly assign people to groups.*** In Chapter 6 we spoke at length about 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 apt to be different from one another in many ways that are relevant to the variables under investigation. For example, let's return to that study in which a researcher wants to investigate the effectiveness of a new method of teaching reading. The first-graders in the study almost certainly differ from one another in intelligence, motivation, educational opportunities at home,

and other factors that would affect their performance on a reading 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 a researcher to make comparisons among two or more groups—are based on the assumption that group membership is randomly determined and that any pretreatment group differences are the results of chance alone.

5. ***Use one or more pretests to assess equivalence before the treatment(s).*** 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, employees at various companies, 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, related 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 might possibly have an effect on the dependent variable. For instance, a researcher comparing two different treatments for a particular illness might match patients according to age, gender, and duration and intensity of the illness. Likewise, a researcher comparing the achievement test scores of students in two different instructional programs might identify pairs of same-age and same-gender students who have similar IQ scores. In either case, the researcher doesn't 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 should, in the final research report, explain in what way(s) the participants in the study have been matched, for example, by saying, "Pairs were matched on the basis of age, gender, and previous grade point average."

One problem with assessing before-treatment equivalence with pretests is that the researcher rules out *only the variables that have actually been assessed and then determined to be equivalent across groups.* The design doesn't rule out other influential factors that the researcher hasn't assessed and perhaps hasn't even considered.

6. ***Expose participants to all experimental treatments.*** 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 in a study is known as a **within-subjects variable**; you may also see the term **repeated-measures variable** used in such situations.[2]

As an example, let's return to Thrailkill's (1996) dissertation involving three different lecture methods and their possible effects on recall of lecture content. Thrailkill's sample consisted of volunteer students who were enrolled in three sections of an undergraduate psychology class, and she planned to give the lecture just three times, once to each section. The lecture was about a little-known American Civil War battle that participants were unlikely to have learned about in school; thus, participants' prior knowledge about the battle was a constant—they all had *zero* prior knowledge—rather than a confounding variable. Thrailkill divided the lecture into three parts of

[2]As previously mentioned in a footnote in Chapter 6, the term *repeated-measures* is sometimes used as an alternative label for an experience-sampling method (ESM). In our reading of the literature, we have typically seen this term used—rather than the term *within-subjects*—in research reports describing qualitative or descriptive quantitative studies, especially when the studies involve structured questionnaires, open-ended questions, or other self-assessment instruments. Use of the term *within-subjects* is largely restricted to reports describing experimental studies.

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		
	First Part	*Middle Part*	*Last Part*
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.

Thrailkill 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 or, in some cases, with just a single individual; we describe general designs involving within-subjects variables later in the chapter.

7. ***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* (SEM) 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.*

MyLab Education Self-Check 7.1

MyLab Education Application Exercise 7.1: Selecting and Controlling Variables

TYPES 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. Many 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 strategies—some more effective than others—for identifying possible cause-and-effect relationships.[3]

[3]Designs 1 to 6 and Designs 8 to 11 are based on those that Campbell and Stanley (1963) 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.*

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 each contain one of four notations:

Tx: Indicates that a *treatment* (reflecting an independent variable) is presented.

Obs: Indicates that an *observation* (reflecting a dependent variable) is made; typically involves a relatively objective *assessment* of the dependent variable.

—: Indicates that nothing occurs at a particular time.

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 most of the design tables we present will include only one or two groups (perhaps one treatment group and one control group), it's entirely possible for a design to have several 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 starting points that get you thinking about how you might best tackle your own research problem.

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 an observation (Obs)—typically an assessment of some sort—is made to determine what effect the treatment may have had. This design is shown in the following table:

Group	Time →	
Group 1	Tx	Obs

The design has low internal validity because it's impossible to determine whether participants' assessment results are due to the experimental treatment per se. Many other variables may have influenced participants' post-treatment performance, such as physiological maturation or experiences elsewhere in the participants' general environment. Perhaps the characteristic or behavior assessed after the treatment existed *before* the treatment as well. The reality is that with a single observation, we have no way of knowing whether the situation has actually changed, let alone whether it has changed as a result of the intervention.

One-shot allegedly "experimental" case studies may be at the root of many common misconceptions. For example, imagine that we see a child sitting on the ground on a cool,

rainy day. The next day the child has a sore throat and a cold. We conclude that sitting on the damp earth caused the child to catch a cold. Thus, the design of our "research" thinking is something like this:

Exposure to cold, damp ground (Tx) → Child has a head cold (Obs)

Such "research" may also "support" such superstitious folk beliefs as these: If you walk under a ladder, you'll have bad luck; Friday the 13th is a day of catastrophes; a horseshoe nailed above the front door brings good fortune to one's home. 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 qualitative case study design described in Chapter 8. Qualitative case studies require 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, a researcher 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 assessment, then (b) is administered the experimental treatment, and finally (c) is assessed 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 8 weeks, and finally administers an alternate form of the same standardized reading test. If the students' test scores improve over the 8-week period, the teacher might conclude—perhaps correctly, but perhaps not—that the simultaneous-reading-and-listening intervention was the cause of the improvement.

Now suppose an agronomist crossbreeds two strains of corn. She finds that the resulting hybrid strain is more disease-resistant and has a better yield than either of the two parent types. She concludes that the crossbreeding 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 occurred. However, we haven't ruled out other possible explanations for the change. In the case of the elementary school teacher's study, improvement in reading scores may be due to other activities within the classroom curriculum, to more practice taking the reading test, or simply to the fact that the students are 8 weeks older. In the case of the agronomist's experiment, changes in rainfall, temperature, or soil conditions might actually be 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 treatment; the control group is not. After the treatment, both groups are assessed and their respective performances 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.

Designs 1, 2, and 3 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 to the three very simple designs just described, **experimental designs** offer a considerable 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: Control-Group Pretest–Posttest Design

In a control-group pretest–posttest design, people or other units of study (e.g., members of a particular plant or animal species) are randomly assigned to either an experimental group or a control group. The experimental group is observed (assessed), 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 control-group pretest–posttest design is as follows:

	Group	Time →		
Random Assignment	Group 1	Obs	Tx	Obs
	Group 2	Obs	—	Obs

Simple as it is, this design 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 forms 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 assessing participants before administering the experimental treatment may, in and of itself, influence how participants respond to the treatment. For example, perhaps the first assessment increases people's motivation: It makes them want to benefit from any intervention they receive—another instance of the *reactivity* effect.

To address the question *What effect does pretesting have?,* Solomon (1949) proposed an extension of the control-group pretest–posttest 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 who are not pretested has a particular advantage. If the researcher finds that in the final assessment, 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* (i.e., generalizability) 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: Control-Group Posttest-Only Design

Some life situations defy pretesting. You can't pretest the forces in a thunderstorm or a hurricane, nor can you pretest growing crops. Additionally, sometimes 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 control-group posttest-only 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 *absolutely 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

In a strictly **within-subjects design**, all participants are exposed to all experimental treatments and any control conditions.[4] Note that the design name uses the term *subject* rather than the term *participant.* The former word, *subject,* was the preferred terminology when this design was introduced in the 20th century. Furthermore, *subjects* has a broader meaning than *participants*, in that it can be used to refer to a wide variety of populations—perhaps people, dogs, or laboratory rats.[5]

[4]An alternative name for a within-subjects design is a *repeated-measures design.* Increasingly, however, repeated measures are also being used in nonexperimental studies (see an earlier footnote in this chapter), hence our reluctance to use the term here.
[5]We authors suspect that the gradual shift from *subjects* to *participants* in recent decades is due, in part, to the fact that the former term subtly suggests that a researcher might be *subjecting* people to an adverse experimental treatment without their consent. In contrast, the term *participants* connotes more voluntary involvement and engagement in a research study. For a good discussion of the shift, see Roediger (2004).

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 investigate the effects of graphic illustrations in instructional software programs designed to teach science concepts in online college courses. The software created for the study defines and describes 20 scientific concepts with similar precision and depth. In addition, the software illustrates 10 of those concepts (chosen randomly) with pictures or diagrams. After students have completed the software curriculum, 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, graphic illustrations help students learn science more effectively. In other words, the researcher has identified a cause-and-effect relationship: Illustrations enhance science learning.

For a within-subjects design to work, the two or more treatments must be such that their effects are fairly localized and unlikely to "spread" beyond specifically targeted behaviors. Such 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 wouldn't make sense to use a within-subjects design to study the effects of two different intervention techniques for reducing the number of felony convictions in adolescence: If a single group of adolescents receives both treatments and then shows a significant reduction in felonies, we might suspect that either treatment could have had a fairly broad impact; we wouldn't know whether one of the treatments was more effective than the other.

Ideally, too, the two different treatments are administered repeatedly, one after another, in a balanced but somewhat random order. For example, in the instructional software 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 program.

With the last point in mind, let's return once again to Thrailkill's (1996) dissertation involving a lecture about the American Civil War. 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, Thrailkill administered first one treatment (e.g., attention), then another (e.g., imagery), and finally the third (e.g., control). Had she limited her study to a single group, she could not have ruled out an alternative influential factor—*when* in the lecture the information appeared—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 she 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 have emphasized the importance of *randomness*, either in the selection of group members in a multiple-groups study or in the presentation of various 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 they can't completely rule out some alternative explanations for the results they obtain. They must take into consideration whatever variables and explanations they haven't controlled for 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 control-group pretest–posttest design (Design 4). Like Design 3, it involves two groups to which participants haven't been randomly assigned. But it incorporates the pretreatment assessments 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 randomly assigned group membership, there's no guarantee that the two groups are similar in every respect prior to the experimental treatment or intervention—no guarantee that any differences between the groups are due entirely to chance. However, an initial assessment (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 the treatment.

Identifying matched pairs in the two groups is one way of strengthening the nonrandomized control-group pretest–posttest design. For instance, if we are studying the effect of a particular preschool program on children's later achievement test scores, we might identify pairs of children—each pair including one child who is enrolled in the preschool program and one who is not—who are the same age and gender and have similar (and recently assessed) IQ scores before the program begins. Although we can't 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 enhancing 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., assessing 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. The design thus looks something like the following:

Group	Time →								
Group 1	Obs	Obs	Obs	Obs	Tx	Obs	Obs	Obs	Obs

In such studies, the observations made prior to the treatment are collectively called **baseline data**.

This 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) provides 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 simple time-series design might remind you of the *experience-sampling methods (ESMs)* previously described in Chapter 6 because both methodologies employ repeated measures of the variable(s) of interest. Remember, however, that ESM studies are typically *descriptive* studies—that is, no experimental treatments are imposed midway through data collection.

The major weakness of a simple time-series 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 7.1. If this other event is actually the cause of the change, any conclusion that the treatment has brought about the change will obviously be incorrect.

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 general design takes the following form:

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 9). 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, 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—although not entirely eliminating—the probability that outside effects might bring about any changes observed. The intervening experimental treatment is sometimes present and sometimes absent, and the dependent variable is assessed at regular intervals. Thus, we have the following design:

Group	Time →							
Group 1	Tx	Obs	—	Obs	Tx	Obs	—	Obs

As an illustration of this design, imagine that, as researchers in the field of social media, five of us are interested in whether people can reinforce one another for—and thereby increase the frequency of—personal posts in cyberspace. We decide to address this research question by investigating the potential reinforcing effects of "like" reactions to people's Facebook posts. Fortunately, all five of us have Facebook accounts and at least 50 Facebook "friends" whom we have in common. On the first Monday in March, we all react with a "like" to as many of our common friends' Facebook posts as possible; then, on Tuesday through Sunday, we "like" nothing, but we count the number of new posts that those friends create. On the following Monday, we do nothing, but, as before, we count the number of new messages posted on Tuesday through Sunday. During the third week, once again, we like as many posts as possible on Monday and count new posts on Tuesday through Sunday. The fourth week involves no liking but continued data collection. Our design would look as follows:

Group	Time →							
Facebook friends	"Like"	Count posts	—	Count posts	"Like"	Count posts	—	Count posts

Design 12: Alternating-Treatments Design

A variation of 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 manner:

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.[6]

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 likely to have long-lasting and perhaps fairly general effects. Furthermore, if an experimental treatment is apt 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 →					
Group 1	*Baseline* →		*Treatment* →			
	—	Obs	Tx	Obs	Tx	Obs
Group 2	*Baseline* →				*Treatment* →	
	—	Obs	—	Obs	Tx	Obs

[6]Some researchers use the term *ABAB design* in reference to either a reversal time-series design or an alternating-treatments design. The letter A designates one experimental treatment, and the letter B designates either a different treatment or a nontreatment condition.

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 5-day intervention in which a woman visited children's classrooms 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 different grade levels over a 3-week period; a random selection process resulted in her visiting first-grade classes one week, second-grade classes the following week, and 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 (for the first- and second-graders) after the interventions. Neither the children nor the observers had any knowledge about the nature and purpose of the study; hence, this was a *double-blind experiment.*

The data the observers collected are depicted in Figure 7.2; numbers of risky behaviors on the slide are shown with the lighter dots, whereas those on the climbing equipment

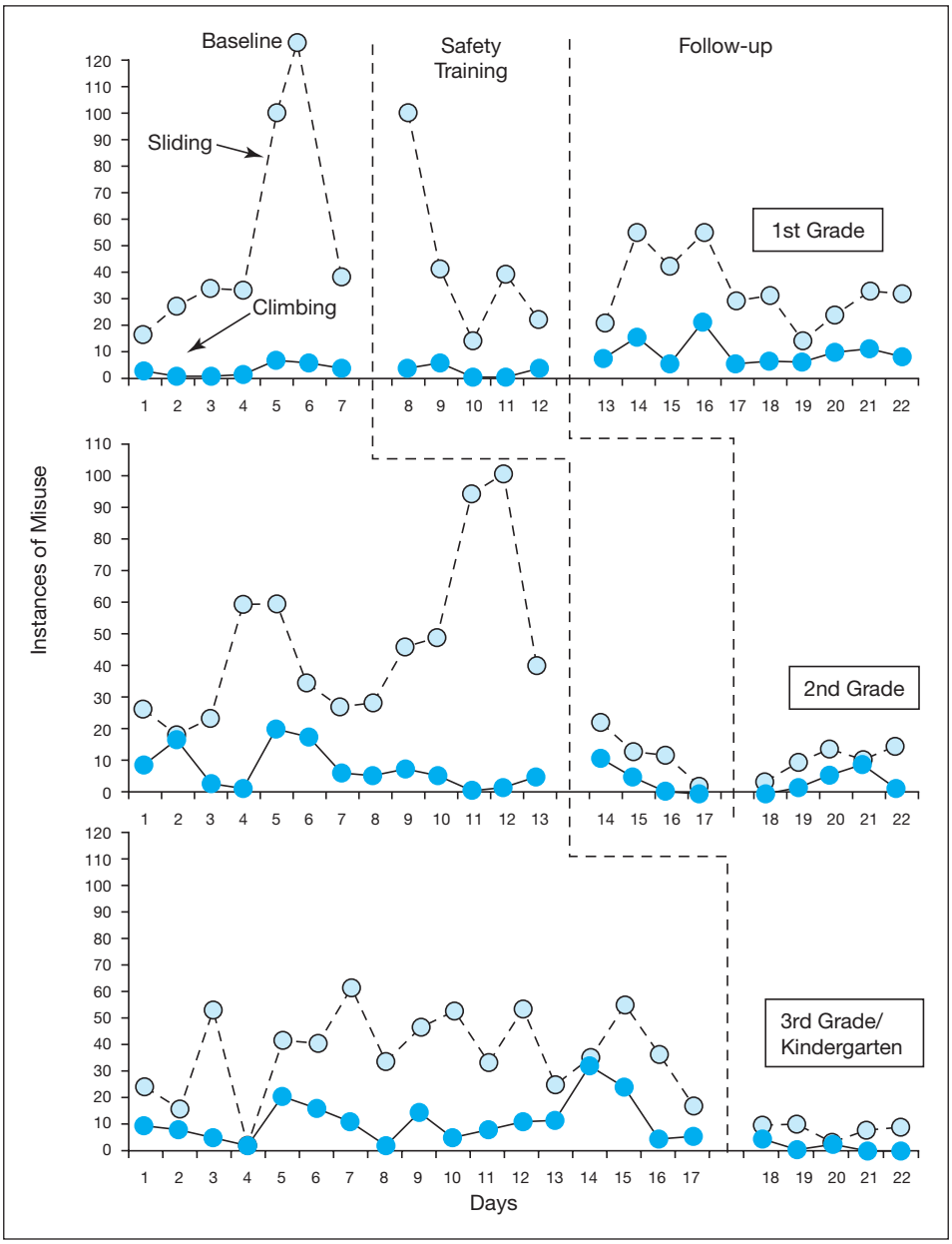


FIGURE 7.2 Instances of Risky Behavior on Slides and Climbers by Grade Level; Third-Graders and Kindergartners Shared a Single Recess

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.

are shown with the darker dots. 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 infrequent 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.

Using Designs 11, 12, and 13 in Single-Case Intervention Research

In the hypothetical single-group Facebook study previously used to illustrate a reversal time-series design, the group was a collection of commonly shared friends who didn't necessarily know or interact with one another. But reversal time-series, alternating-treatments, and multiple-baseline designs might also be used with either a single individual or a single intact group—an approach now commonly known as **single-case intervention research**.[7]

A study by Deaver, Miltenberger, and Stricker (2001) illustrates how a researcher might combine two of the designs—reversal and multiple-baseline—with a single child. A 2-year-old girl named Tina had been referred for treatment because she often twirled her hair with her fingers so vigorously that she pulled some of it out, and on one occasion she wrapped her hair around a finger so tightly that the finger began to turn blue and the hair had to be removed with scissors. Tina engaged in such behavior primarily when she was alone (e.g., at naptime); hence, there was no parent or other adult present to discourage it. The researchers identified a simple treatment—putting thin cotton mittens on Tina's hands—and wanted to document its effect. They videorecorded Tina's behaviors when she was lying down for a nap in either of two settings, her bedroom or her daycare center, and two observers independently counted the number of hair-twirling incidents as they watched the videos. Initially, the observers collected baseline data. Then, during separate time periods for the bedroom and daycare settings, they gave Tina the mittens to wear during naptime. After reversing back to baseline in both settings, they had Tina wear the mittens once again. The percentages of time that Tina twirled her hair in the two settings over the course of the study are presented in Figure 7.3.

In both the bedroom and daycare settings, the researchers alternated between baseline and treatment; this is the *reversal* aspect of the study. Furthermore, they initiated and then later reinstituted the treatment at different times in the two settings; this is the *multiple-baseline* aspect of the study. Figure 7.3 consistently shows dramatic differences in hair twirling during baseline versus mittens conditions, leading us to conclude that the mittens, rather than some other factor, were almost certainly the reason for the disappearance of hair twirling.

We now ask you to recall that the word *case* in single-case intervention research can also refer to an *intact group*. The playground study by Heck et al. (2001) involved three relatively intact groups (the first-graders, the second-graders, and the kindergartners and third-graders who played at the same time); we might actually call their study an example

[7]You may also see such terms as *single-case design* and *single-case experiment*. In previous editions of this book, we have used the term *single-subject design* in reference to quasi-experimental studies in which data are collected on only one individual, but *single-case design* is now the preferred term—and also a more inclusive one.

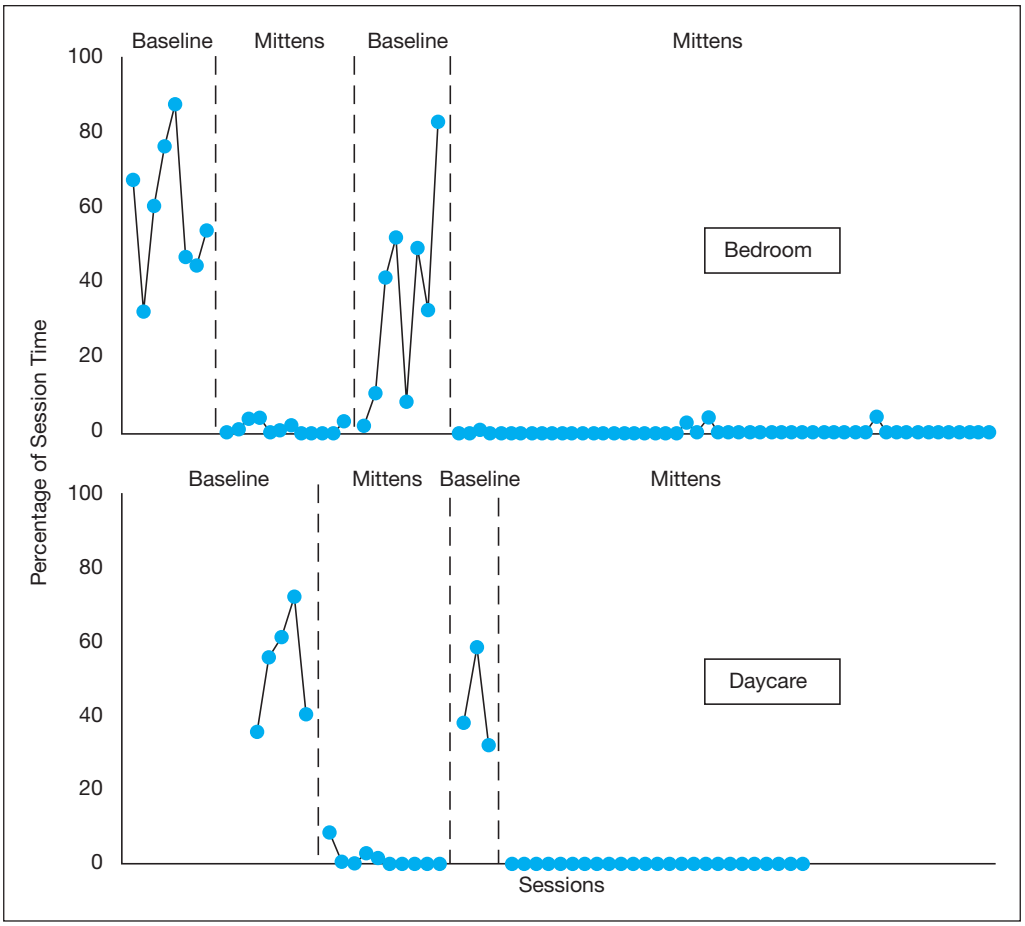


FIGURE 7.3
Percentage of Session Time in Which Hair Twirling Was Observed Both in the Bedroom and at Daycare

Reprinted from "Functional Analysis and Treatment of Hair Twirling in a Young Child," by C. M. Deaver, R. G. Miltenberger, & J. M. Stricker, 2001, *Journal of Applied Behavior Analysis, 34,* p. 537. Reprinted with permission of the Society for the Experimental Analysis of Behavior, Inc.

of *multiple-case intervention research* with a multiple-baseline design. In that study, only overall group data were provided. But especially in applied fields (e.g., clinical psychology, counseling, education, and social work), researchers may want to determine whether and how much one or more interventions affect the behaviors of *both* an overall group *and* of the specific individuals within the group. As a simple example, one team of researchers was interested in documenting the effectiveness of a 2-hour professional development workshop for teachers regarding how to enhance secondary students' understandings of new concepts (Bulgren, Lenz, Schumaker, Deshler, & Marquis, 2002, Study 2). More specifically, the workshop focused on how to increase students' awareness of both similarities and differences between concepts—for instance, between (a) mammals and birds, (b) minerals and rocks, or (c) simple and complex machines. Ten teachers who volunteered to participate in the study were observed during several class lessons both before and after attending one of two workshop sessions, with the observers using a point-based "Implementation Checklist" to determine how frequently each teacher used the various instructional strategies presented in the workshop. Five of the teachers attended the first workshop, and the other five attended the second workshop, thereby reflecting a multiple-baseline approach. The design of the study was essentially the one depicted in Table 7.1. The study's findings were clear-cut: Before the intervention, the teachers rarely, if ever, explicitly contrasted the specific characteristics of, say, mammals versus birds or minerals versus rocks. Yet after the 2-hour workshop, they used many of the newly acquired compare-and-contrast strategies in teaching students new concepts, typically reaching or exceeding an 85% "mastery" level on the checklist during their lessons.

TABLE 7.1 ■ Single-Case Design Used in Study 2 in Bulgren, Lenz, Schumaker, Deshler, and Marquis (2002)

Teacher	Time →								
	Baseline →			*Treatment* →					
Teacher 1	Obs	Obs	Obs	Tx	Obs	Obs	Obs	Obs	Obs
Teacher 2	Obs	Obs	Obs	Tx	Obs	Obs	Obs	Obs	Obs
Teacher 3	Obs	Obs	Obs	Tx	Obs	Obs	Obs	Obs	Obs
Teacher 4	Obs	Obs	Obs	Tx	Obs	Obs	Obs	Obs	Obs
Teacher 5	Obs	Obs	Obs	Tx	Obs	Obs	Obs	Obs	Obs
	Baseline →					*Treatment* →			
Teacher 6	Obs	Obs	Obs	Obs	Obs	Tx	Obs	Obs	Obs
Teacher 7	Obs	Obs	Obs	Obs	Obs	Tx	Obs	Obs	Obs
Teacher 8	Obs	Obs	Obs	Obs	Obs	Tx	Obs	Obs	Obs
Teacher 9	Obs	Obs	Obs	Obs	Obs	Tx	Obs	Obs	Obs
Teacher 10	Obs	Obs	Obs	Obs	Obs	Tx	Obs	Obs	Obs

EX POST FACTO DESIGNS

In many situations, it is 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 lethal new virus, withhold important instruction, ask parents to abuse their children, or change a person's personality to compare the effects of these factors on the dependent variables in one's research problem.

Ex post facto designs[8] (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 chickenpox and those who were not—and then calculate the percentage of reported cases of chickenpox 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.

[8]Ex post facto designs are also known as *causal-comparative* designs. However, as B. 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.

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 isn't possible, the researcher cannot draw firm conclusions about cause and effect. The problem here is that the experimenter can't control for confounding variables that might provide alternative explanations for any group differences observed.

Although an ex post facto study lacks the control element—and thus doesn't enable definite conclusions about cause and effect—it is nevertheless a legitimate research method. 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 came to light 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

On the surface, Design 14 is similar to the static group comparison (Design 3) described in the previous section on pre-experimental designs. The critical 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 hasn't been responsible for imposing it. A simple 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 had an experience as well, albeit an experience quite different 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 definitively show that those other (dependent) variables were actually caused by those (independent) 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 a **factorial design**.

Design 15: Two-Factor Experimental Design

In its simplest form—one involving two independent variables, which we will call *Variable 1* and *Variable 2*—such a design might look 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 performances of Groups 1 and 2 with those 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 performances but that Group 1 outperforms the other three. Such a result might 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 described, participants are randomly assigned to groups in a true experimental study. But it's 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 Assign-ment	Group 1a	Tx_a	Obs
			Group 1b	Tx_b	Obs
Group 2	Exp_b	Random Assign-ment	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 might influence a particular dependent variable *and* how a previous experience or pre-existing characteristic might interact

with that manipulation. If the two variables significantly *interact* in their apparent effects on the dependent variable—for example, a particular intervention might be beneficial for males but not females, or vice versa—the *ex post facto* variable is essentially a *moderating variable* that influences the effects of the experimental intervention. (Refer back to Chapter 2 if you need a refresher about the nature of moderating variables.)

In a variation of Design 16, 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 geographical principles 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 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 geographical principles; for instance, a river originated in the plains and ran *up* into a mountain range, and various transportation networks didn't interconnect in ways that they normally do. The two different maps reflected one of our independent variables: the 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: academic background. In one-on-one sessions, we asked the professors to study each of the two maps aloud for three 2-minute intervals (we audiorecorded what they said during the study sessions) and then, after each of the three intervals, to draw as much of the map as they could remember.

Thus, if we call the two maps Tx_a (logical map) and Tx_b (illogical map)—each of which was shown to participants three times—our design looks like the following:

Group	Time →											
Geographers	Tx_a	Obs	Tx_a	Obs	Tx_a	Obs	Tx_b	Obs	Tx_b	Obs	Tx_b	Obs
Sociologists	Tx_a	Obs	Tx_a	Obs	Tx_a	Obs	Tx_b	Obs	Tx_b	Obs	Tx_b	Obs
Educational psychologists	Tx_a	Obs	Tx_a	Obs	Tx_a	Obs	Tx_b	Obs	Tx_b	Obs	Tx_b	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, academic 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 academic 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 manner, with few attempts to make sense of what they saw on the maps.

Table 7.2 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,

TABLE 7.2 Summary of Experimental and Ex Post Facto Designs

Name of the Design	*Goal 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)	Group / Time → Group 1 | Tx | Obs	Shows a before-and-after sequence but cannot substantiate a cause-and-effect relationship.
2. One-group pretest–posttest design	To show that change occurs after a treatment	Group / Time → Group 1 | Obs | Tx | Obs	Provides a measure of change but yields no conclusive results about the cause of the change.
3. Static group comparison	To show that a group receiving a treatment behaves differently than a group receiving no treatment	Group / Time → Group 1 | Tx | Obs Group 2 | — | Obs	Fails to determine pretreatment equivalence of groups.
		True Experimental Designs	
4. Control-group pretest–posttest design	To show that change occurs following, but only following, a particular treatment	Group / Time → Random Assignment: Group 1 | Obs | Tx | Obs Group 2 | Obs | — | Obs	Controls for many potential threats to internal validity.
5. Solomon four-group design	To investigate the possible effect of pretesting	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 might affect the final outcome observed.
6. Control-group posttest-only design	To determine the effects of a treatment when pretesting cannot or should not occur	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.
7. Within-subjects design	To compare the relative effects of different treatments for the same participants	Group / Time → Group 1 | Tx_a | Obs_a Tx_b | Obs_b	Useful only when effects of each treatment are temporary and localized.
		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	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 assessed.

(continues)

TABLE 7.2 Summary of Experimental and Ex Post Facto Designs *(continued)*

Name of the Design	*Goal of the Research*	*Graphic Depiction*	*Comments on the Design*
		Quasi-Experimental Designs *(continued)*	
9. Simple time-series design	To show that, for a single group, change occurs during a lengthy period only after the treatment has been administered	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.
10. Control-group time-series design	To bolster the internal validity of the preceding design with the addition of a control group	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.
11. Reversal time-series design	To show, in a single group or individual, that a treatment consistently leads to a particular effect	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.
12. Alternating-treatments design	To show, in a single group or individual, that different treatments have different effects	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.
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	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.
		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	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.
		Factorial Designs	
15. Two-factor experimental design	To study the effects of two experimenter-manipulated variables and their possible interaction	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.
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	Group / Time → 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	Requires a larger sample size than two-group studies; random assignment to the experimenter-manipulated variable is essential.

as stated earlier, this is not an exhaustive list of possible designs. You can combine and expand on the 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.

CONCEPTUAL ANALYSIS EXERCISE Identifying Quantitative 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 16 design categories just described, 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 "For Further Reading" section at the end of the chapter.

1. Two researchers want to see if a particular training program is effective for 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' "memories" of 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 U.S. 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).

PRACTICAL APPLICATION Determining Possible Cause-and-Effect Relationships

The research designs described in this chapter differ considerably in the degree to which they control for potential confounding variables—variables that threaten a study's internal validity—and thus they also differ in 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 overall research design 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 study or single-case intervention 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(s)? _____

_____

_____

_____ 7. To what extent might each of the following factors threaten the study's internal validity? If any of these factors poses a potential threat, how is the researcher minimizing or eliminating its influence? (Refer to Figure 7.1.)

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:

_____

_____

_____

_____

MyLab Education Self-Check 7.2

MyLab Education Application Exercise 7.2: Planning for an Experimental Design

MyLab Education Application Exercise 7.3: Planning for a Quasi-Experimental Design

META-ANALYSES

Remember, we can conclude that a cause-and-effect relationship exists between an independent variable and a dependent variable only if we have (a) directly manipulated the independent variable and (b) 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's 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*: 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 in 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 previously conducted analyses. In particular, the researcher combines the results of many experimental and/or ex post facto studies to determine whether they lead to consistent conclusions. A meta-analysis is primarily a statistical technique, and thus we describe this procedure in greater depth in Chapter 11.

CONDUCTING EXPERIMENTS ON THE INTERNET

In Chapter 6 we mentioned that some researchers now conduct quantitative 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.[9] 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 thus 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 impact 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.[10] 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 3 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

[9]As noted in Chapter 6, this website is maintained by John Krantz, Professor of Psychology at Hanover College (psych.hanover.edu).
[10]In case you're curious, Norwegians seem to be especially partial to spicy Mexican food, at least in comparison with other Europeans.

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. (If you've noticed a possible problem with *attrition* in the study, give yourself a pat on the back! We'll address this problem shortly.)

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 study. The suggestions for ethical practices presented in Chapter 6 for online questionnaires are equally applicable to online experiments (see the Practical Application "Using the Internet to Collect Data for a Descriptive Study" in Chapter 6).

Remember, too, that the sample you get in an online study will hardly be representative of the overall population of human beings; for instance, it's likely to consist largely of college-educated, computer-literate people who enjoy participating in research studies. An additional problem is that you can't 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 investigating a topic such as very-long-term memory (as Cepeda and colleagues did 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 studies often begin with specific *a priori* hypotheses; quasi-experimental and ex post facto studies may begin with them as well. 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. These procedures are briefly described in Chapter 11.

Yet your analyses need not be restricted *only* to the testing of the hypotheses you've put forth before collecting your data. Oftentimes a study will yield additional results—results that are unexpected yet intriguing—that merit analysis. There's no reason why you can't examine these findings as well, perhaps statistically, perhaps not. You must be careful not to go overboard, however, as you will learn in the discussion of *Type I errors* in Chapter 11.

Finally, don't forget that all-important final step: *interpreting* your results. It isn't enough to report whether you did or didn't find evidence to support your hypotheses. You must also explain *why* you found what you did and *whether or in what ways* your findings answered your research problem and its subproblems.

PRACTICAL APPLICATION Acknowledging the Probable Presence of Bias in Experimental Research

Despite the tight controls in many experiments—and in some cases *because* of such controls—one or more forms of bias can wiggle their ways into the data or into interpretations of the data. Some of these biasing factors, such as group selection procedures, statistical regression, and differing attrition rates, can adversely affect the *internal validity* of a study (look once again at Figure 7.1). For example, as you were reading about the memory-for-trivia experiment in the earlier section on Internet-based experiments, you might have wondered if the dropout (attrition) rate was higher for participants with longer between-session delays, and indeed it was (Cepeda et al., 2008). Were participants who had poor memories more likely

to drop out over the long run than participants who had good memories? If so, by Session 3, the people who remained in spread-out-studying treatment groups might simply have had better memories *in general* than people who remained in close-together-studying treatment groups. To determine the extent to which the differing attrition rates for various treatment groups might jeopardize the study's internal validity, Cepeda and colleagues collected basic demographic data at the beginning of Session 1. In their data analyses, they found no significant differences in any demographic variables or in Session 1 performance between participants who completed all three sessions and those who did not—thus lending support to their premise that the members of the various treatment groups were similar in all ways *except* for the differing study intervals.

Other biasing factors can negatively impact the *external validity* (generalizability) of an experimental study. For example, when conducting their online study of studying and memory, Cepeda and his colleagues not only had a biased sample—computer-literate individuals who volunteered to participate—but they also had participants study *trivial facts* in an *atypical learning situation.* Would their conclusions apply to more important learning tasks in everyday learning contexts? Their results alone can't answer this question.

Furthermore, a researcher's expectations and hypotheses are likely to bias a study from the get-go: By focusing only on certain variables and by assessing those variables in predetermined ways, no researcher can discover the "whole truth" about the phenomenon under investigation—if there *is* a general, underlying truth. Try as we might, we human beings can't be completely objective and unbiased in our well-intended efforts to make better sense of our physical, biological, social, and psychological worlds. But we can—and *must*—be honest about our biases so that others can realistically evaluate the merits of our work.

CHECKLIST

Identifying Potential Sources of Bias and Potential Threats to External Validity in an Experimental, Quasi-Experimental, or Ex Post Facto Study

The previous checklist, "Looking for Confounding Variables," can help you identify biasing factors that can adversely affect the internal validity of a study. Following are additional questions to keep in mind when looking for possible biases in an experimental, quasi-experimental, or ex post facto study, as well as for factors that might adversely affect a study's external validity.

_____ 1. What one or more *a priori* hypotheses about cause-and-effect relationships has the researcher formed about the phenomenon being studied? What potentially important variables might these hypotheses be leading the researcher to ignore?

_____ 2. How is the sample being selected for the study? Is this sample likely to be representative of the population about which generalizations are being made? Why or why not?

3. If the study involves human beings, what are participants being asked to do? To what degree are these tasks similar to real-world activities?

_____ 4. In what setting is the study being conducted? To what degree is it similar to settings in everyday life?

_____

_____

_____ 5. With your answers to Questions 2, 3, and 4 in mind, describe the extent to which the results of the study should be generalizable to diverse individuals, tasks, and settings.

_____

_____

_____

MyLab Education Self-Check 7.3

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. Kinnick wanted to determine whether a particular method of teaching concepts (one described by theorists Tennyson and Cocchiarella) might be more effective for teaching fetal monitoring skills than the method traditionally used in nursing education programs. In Kinnick's dissertation, the purpose statement is 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 was not one of the designs we have specifically described in this chapter. Instead, it involved 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 →		
Random Assignment†	Group 1	Tx_1	Obs	Obs
	Group 2	Tx_2	Obs	Obs
	Group 3	Tx_3	Obs	Obs

We now present excerpts from the methodology chapter of the researcher's dissertation. Our comments and observations appear on the right-hand side.

DISSERTATION ANALYSIS 4

Methodology

(After an introductory paragraph outlining the chapter's contents, the author describes the sample used in the study: students enrolled in maternity nursing courses at two universities. 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 plus 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.

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. . . .

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 study 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).

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.

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.

Notice that the author's notion of what is "traditional" instruction is based on her own experiences, and she says so here.

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.

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.

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.)*

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.

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.

By "interrogatory presentation," the author means asking questions to assess students' understanding of, and ability to apply, what they have learned. The "transparencies" to which the author refers were a low-tech equivalent to the PowerPoint slides one might present on a computer screen in a contemporary college classroom.

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.)

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.

Because the author conducted the study at two universities, she followed the necessary human research review procedures at both institutions.

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.

Notice how the author randomly assigned participants to instructional methods as one means of controlling for confounding variables.

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 posttest 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 students on their individual patients.

Statistical Analysis

(The author continues with a discussion of the statistical analyses she used to compare the performance of the three groups.)

Note: 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.

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 presumably helped 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.

APPLICATION EXERCISES

Similar exercises are available on Pearson MyLab Education.

1. Kaitlyn is an undergraduate student interested in studying peer-tutoring programs for a class project. She hypothesizes that children who are tutored by older children will be more cooperative and will learn more than children who are older than their tutors. She designs an experimental study in which one group of low-achieving seventh-grade children will be tutored by ninth-graders, and another group will be tutored by high-achieving sixth-graders. She plans to give all children a pretest and a posttest to assess their content knowledge, and she plans to have teachers evaluate cooperativeness both before and after the tutoring program, which will last 4 weeks.
   i. Identify the independent and dependent variables.
   ii. Kaitlyn has worked hard to design a study that has high internal validity. How has she attempted to control for history?
   iii. How might Kaitlyn attempt to control for a possible effect of testing?
   iv. How might Kaitlyn attempt to control for a possible effect of selection?
2. Consider the following research questions, and state whether the question might be addressed with an experimental research study, and explain why or why not.
   i. Consider the following research question involving college students enrolled in a creative writing course: "Do students who receive peer comments on drafts of their papers revise their papers as well as students who receive instructor comments?"
   ii. Consider the following research question involving the cognitive skills of elementary school children: "Are there gender differences in fourth-grade learners' spatial ability as measured by a computer-delivered assessment?"
   iii. Consider the following research question: "How does positive reinforcement affect workers' attitudes toward the workplace?"
3. You are interested in the benefits of different types of technology training on nurses' technology skills and their attitudes regarding instructional technologies. For your study, you are able to study all of the nurses working in a large regional hospital.
   i. Describe how you would conduct this study as a pretest–posttest control-group design, and diagram it.
   ii. Describe how you would conduct this study as a time-series design, and diagram it.

FOR FURTHER READING

Antony, J. (2014). *Design of experiments for engineers and scientists* (2nd ed.). London, England: Elsevier.

Barlow, D. H., Nock, M. K., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior change* (3rd ed.). Hoboken, NJ: Pearson.

Bausell, R. B. (1994). *Conducting meaningful experiments: Forty steps to becoming a scientist.* Los Angeles, CA: Sage.

Cameron, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., . . . Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science, 351,* 1433–1436.

Canavos, G. C., & Koutrouvelis, I. A. (2009). *An introduction to the design and analysis of experiments.* Hoboken, NJ: Pearson.

Dugard, P., File, P., & Todman, J. (2012). *Single-case and small-n experimental designs: A practical guide to randomization tests* (2nd ed.). New York, NY: Routledge.

Glass, D. J. (2014). *Experimental design for biologists* (2nd ed.). 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.

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings* (2nd ed.). New York, NY: Oxford University Press.

Kirk, R. E. (2013). *Experimental design: Procedures for the behavioral sciences* (4th ed.). Los Angeles, CA: Sage.

Kratochwill, T. R., & Levin, J. R. (Eds.). (2014). *Single-case intervention research: Methodological and statistical advances.* Washington, DC: American Psychological Association.

Montgomery, D. C. (2012). *Design and analysis of experiments* (8th ed.). New York, NY: Wiley.

Morgan, D. L., & Morgan, R. K. (2001). Single participant research design: Bringing science to managed care. *American Psychologist, 56,* 119–127.

O'Neill, R. E., McDonnell, J. J., Billingsley, F., & Jenson, W. (2011). *Single case research designs in educational and community settings.* Hoboken, NJ: Pearson.

Phillips, D. C. (1981). Toward an evaluation of the experiment in educational contexts. *Educational Researcher, 10*(6), 13–20.

Plavnick, J. B., & Ferreri, S. F. (2013). Single-case experimental designs in educational research: A methodology for causal analyses in teaching and learning. *Educational Psychology Review, 25,* 549–569.

Ruxton, G., & Colegrave, N. (2010). *Experimental design for the life sciences* (3rd ed.). Oxford, England: Oxford 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, MA: Houghton Mifflin.

Solso, R. S., & MacLin, K. (2008). *Experimental psychology: A case approach* (8th ed.). Hoboken, NJ: Pearson.

ANSWERS TO THE CONCEPTUAL ANALYSIS EXERCISE "Identifying Quantitative Research Designs":

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-Case Intervention Research," a multiple-baseline study is described in which a single 2-year-old girl successively received a particular treatment (having mittens on her hands) in two different contexts. In this example, however, we see the approach being used with five different horses, each of which receives the treatment only once.
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 participants are not randomly assigned as *individuals* to the treatment and control groups. More specifically, the study uses 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 *factorial* design, and more specifically a *combined experimental and ex post facto* design with two independent variables. The experimental variable—in this case a within-subjects variable—is the content of the drawings (aggression vs. prosocial behavior). The ex post facto variable is the presence or absence of attention-deficit hyperactivity disorder, which the researchers do not (and *cannot*) manipulate.

This page intentionally left blank

Chapter 8 Qualitative Research Methods

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 indeed dig deep: We collect various forms of data and examine them from various angles to construct a rich and meaningful picture of a complex, multifaceted situation.

Learning Outcomes

8.1 Identify several situations in which a qualitative methodology might be especially useful.

8.2 Describe general characteristics and purposes of (a) case studies, (b) ethnographies, (c) phenomenological studies, (d) grounded theory studies, (e) narrative inquiry, and (f) content analyses. Also, describe effective strategies you might use in each of these six research methodologies.

8.3 Identify several different strategies you might use to enhance the general credibility and trustworthiness of qualitative data you collect.

8.4 Also identify effective data-collection strategies related to (a) sampling, (b) making observations, and (c) conducting interviews in a qualitative study.

8.5 Describe several general criteria that are often used in evaluating qualitative studies.

The term **qualitative research** encompasses a number of methodologies that are in some respects quite different from one another. Yet all qualitative approaches have two things in common. First, they typically focus on phenomena that are occurring or have previously occurred 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 or quantify 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.

Qualitative research can be found in many academic disciplines, including anthropology, sociology, psychology, biology, history, political science, education, and medicine. In fact, it could be argued that inquiry in *any* discipline begins in a qualitative form (e.g., Lauer & Asher, 1988). When little information exists on a topic, when variables are unknown, or 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.* For example, the field of medicine 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 in a *mixed-methods design* (details to follow in Chapter 9).

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 six kinds of qualitative studies: case studies, ethnographies, phenomenological studies, grounded theory studies, narrative inquiry, and content analyses.

As you proceed through the chapter, you will find several strategies—sampling, making observations, interviewing—that you previously encountered in the discussion of descriptive quantitative studies in Chapter 6. These are old news, you might think. On the contrary, such strategies can take on very different forms when we want them to yield qualitative data.

Qualitative research can be quite different from quantitative research in another important way as well. In discussions of quantitative designs and strategies in the preceding three chapters, we imply—intentionally—that data collection comes first, with data analysis to follow in a separate step. In qualitative research, however, the methodology often involves an *iterative and recursive process* in which the researcher moves back and forth between data collection and data analysis. For example, the researcher might (a) collect some preliminary data in a natural setting; (b) inspect the data for possible patterns; (c) return to the setting to collect additional data that might substantiate, clarify, or contradict those patterns; and (d) conduct a more thorough, detailed analysis of the data—possibly repeating Steps c and d through additional iterations. Accordingly, if you are planning a qualitative study, you should read both this chapter *and* the discussion of qualitative data analysis in Chapter 12 before beginning data collection.

RESEARCH PROBLEMS AND METHODOLOGY CHOICE IN QUALITATIVE RESEARCH

In Chapter 2 we emphasized the importance of pinning down the research problem with utmost precision. We sometimes find an exception in qualitative research. Some 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 don't remain so loosely defined, however. As a study proceeds, the qualitative researcher gains increasing understanding of the phenomenon under investigation and thus becomes increasingly able to ask more specific questions—and occasionally can begin to formulate working hypotheses as well.

When qualitative researchers ask only broad research questions at the beginning of an investigation, they may have trouble 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 therefore ask more specific questions, so, too, can they better specify what strategies they will use to answer those questions.

In some instances, then, the methodology of a qualitative study may continue to evolve over the course of the investigation. Despite this fact, we must emphasize that *qualitative research requires considerable preparation and planning.* Qualitative researchers must be well trained in observation techniques, interview strategies, and whatever other data-collection methods are likely to be necessary to learn more about and make some sense of the phenomena they are studying. They must have a firm grasp of previous research related to their research problems and questions so that they know what to look for and can separate important information from unimportant details in what they observe (some grounded theory studies are exceptions, for reasons you will discover shortly). And they 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.

Potential Advantages of a Qualitative Approach

Qualitative research studies typically serve one or more of the following purposes:

- ***Exploration.*** They can help you gain initial insights into what has previously been a little-studied topic or phenomenon.
- ***Multifaceted description.*** They can reveal the complex, possibly multilayered nature of certain situations, settings, processes, relationships, systems, or people.
- ***Verification.*** They allow you to test the validity of certain assumptions, claims, theories, or generalizations within real-world contexts.
- ***Theory development.*** They can enable you to develop new concepts or theoretical perspectives related to a phenomenon.
- ***Problem identification.*** They can help you uncover key problems, obstacles, or enigmas that exist within the phenomenon.
- ***Evaluation.*** They provide a means through which you can judge the effectiveness of particular policies, practices, or innovations.

As a general rule, however, qualitative studies do *not* allow you 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.

MyLab Education Self-Check 8.1

QUALITATIVE RESEARCH DESIGNS

In this section, we describe six commonly used 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 each methodology. Later in the chapter, we discuss data-collection strategies that are more broadly applicable to qualitative research. But our space is limited here. 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. 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.

In a typical case study, a 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 spends an extended period of time on site and regularly interacts with the person or people being studied. The researcher also records details about the context surrounding the case or cases of focus, including information about the physical environment and any historical, economic, and social factors that have a bearing on the situation. By portraying such contexts, 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.

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 events. In either circumstance, it tends to be most useful for generating or providing preliminary support for tentative explanations regarding the phenomenon being studied.

Ethnography

In a case study, a researcher looks in considerable depth at a particular person, program, or event. In contrast, in an **ethnography**, a researcher examines 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, from several months to 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, practices, 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, exclusive social cliques in adolescence, and Internet-based communities[1] (e.g., Kozinets, 2010; McGibbon, Peter, & Gallop, 2010; Mehan, 1979; Merten, 2011).

The group chosen for in-depth study should, of course, be appropriate for answering a researcher's general research problem or question. Ideally, it should also be one in which the researcher is a "stranger" and has no vested interest in the study's outcome. A group that the researcher knows well (perhaps one that involves close acquaintances) might 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, 2013).

[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.

Site-based fieldwork is the *sine qua non*—the essence—of any ethnography. Prolonged engagement in a group's natural setting gives ethnographic researchers time to observe and record processes that would be almost impossible to learn about by using any other approach. Thus, an essential first step in an ethnographic study is to gain legitimate access to the site. Often, researchers must go through a **gatekeeper**, a person who can smooth the way for their entrance into the situation. 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, researchers must establish rapport with and gain the trust of the people being studied. At the same time, they must be open about why they are 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, researchers cast a broad net, intermingling with everyone and getting an overall sense of the social and cultural context. Gradually, they identify **primary informants**—also known as **key informants**—who can provide information and insights relevant to their research question and can facilitate contacts with other helpful individuals.

In most ethnographic studies, researchers engage in **participant observation**, becoming immersed in the daily life of the people. Participant observation occurs along a continuum, from passive to active. In fact, over the course of the study, researchers' roles may gradually change from "outsiders" to "insiders" as they become more involved in daily practices and activities of setting and develop relationships with participants. The advantage here is that they might gain insights about a group and its behaviors that could not be obtained in any other way. The disadvantage is that they may become so emotionally involved as to lose the ability to assess the situation with any degree of objectivity.

Although ethnographers were once cautioned not to become so socially and emotionally involved with a group that they wholeheartedly join the group,[2] the ethnographers of today often conduct research in communities and settings in which they are already active participants. For example, in an **autoethnography**, researchers conduct research on their own perspectives, beliefs, experiences, and practices (C. Ellis, Adams, & Bochner, 2011; Muncey, 2010). Furthermore, many qualitative researchers now acknowledge that researchers' roles as insiders and outsiders are not dichotomous; researchers can occupy positions of both insider and outsider relative to the topic, setting, and participants. In addition, rather than pretend that they can remain completely objective, ethnographers are urged to be aware of and explicit about their subjective biases and assumptions.

Throughout their fieldwork, ethnographic researchers are careful observers, interviewers, and listeners. Furthermore, while in the field, they take extensive field notes that focus on describing and detailing behaviors, interactions, and activities within the setting; their field notes might also include diagrams, maps, and other graphics. These *in situ* field jottings—the ones created while actively immersed in the field—are often translated into formalized field notes after an observation has been completed. Lengthy conversations and significant events can be recorded using audiotapes and videotapes and later transcribed. Researchers may also collect or take photographs of artifacts (e.g., tools, ritualistic implements, artistic creations) and documents (e.g., site documents, personal journals, lesson plans) from the group.

In addition, many ethnographers maintain reflective journals during the course of their fieldwork. These journals might include reflections on aspects of their research process, such as gaining access to a site and building rapport and trusting relationships with members and informants. Journal entries might also address methodological challenges and ethical dilemmas, or they might discuss emerging themes or working hypotheses.

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

[2]Becoming a full member of and participant in a group midway through an ethnographic study is known as *going native.* In some instances, going native can prevent a researcher from completing the investigation with any credibility.

to obtain information, which can be either an advantage (to an experienced researcher who knows what to look for) or a disadvantage (to a 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 (Clifford & Marcus, 1986; Geertz, 1973).

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 individual. A **phenomenological study** is a study that attempts to understand people's perceptions and perspectives relative to 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 assertions regarding *what something is like* from an insider's perspective.

Phenomenological researchers depend almost exclusively on lengthy interviews (perhaps 1 to 2 hours in length) with a small, carefully selected sample of participants. A sample size can range from 5 to 25 individuals, all of whom have had direct experience with the phenomenon being studied.

One prominent phenomenological design involves the **three-interview series** developed by Irving Seidman (2013). In the first interview, which involves getting details about participants' life histories, the researcher attempts to *situate* participants' lived experiences—that is, to place those experiences within a larger context relative to the topic of study. Questions in this interview can help participants reconstruct relevant experiences in a variety of settings and contexts, including families, school, communities, and work. The second interview helps to pin down concrete details of particular experiences related to the topic under study. The first two interviews lay a foundation for the third, final interview, which encourages participants to reflect on the meanings of their experiences; in addition, it poses questions that enable both the participants and the researcher to make connections between (a) aspects and contexts of participants' personal and professional lives and (b) the research topic. During this last interview, the researcher listens closely as participants describe their everyday experiences related to the phenomenon at hand; the researcher must also be alert for subtle yet meaningful cues in participants' expressions, pauses, questions, and occasional sidetracks.

Throughout the data-collection process, phenomenological researchers try to suspend any preconceived notions or personal experiences that may unduly influence what they "hear" participants saying. Such suspension—sometimes called *bracketing* or *epoché*—can be extremely difficult for researchers who have personally experienced the phenomenon under investigation. Yet it's essential if they are to gain an understanding of people's typical experiences. The ultimate goal of a phenomenological study should be—not only for the researcher but also for readers of the final research report—to provide a sense 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 described 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 grand theory that might be generalized to other settings, groups, and processes.* 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.[3]

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 the 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).

As is true for the qualitative designs previously described, data collection in a grounded theory study is field-based, flexible, and likely to change over the course of the investigation. 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, 2014; Corbin & Strauss, 2008).

More so than in any other qualitative methodology, data analysis in a grounded theory study begins almost immediately, at which point the researcher develops *categories* to classify the data and develop theory. Researchers carefully choose participants who can represent certain attributes or experiences that can inform the development of theory (we'll say more about their sampling strategies shortly). Subsequent data collection is aimed at *saturating* the categories—in essence, learning as much about them as possible—and at finding any disconfirming evidence (a.k.a. *deviant cases*) that point to possible revisions of the categories themselves or of interrelationships among them. 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).

Virtually all experts agree that grounded theory researchers should have a firm grasp of general concepts and theoretical orientations in their discipline as a whole; hence, an in-depth literature review early in the process is essential. However, experts disagree about whether researchers 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 the data collected. 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 collected data (e.g., Hesse-Biber, 2010; Jaccard & Jacoby, 2010). As your book's authors, our own advice is to learn as much as you can about your research topic through a thorough review of the related literature. However, we strongly urge you to follow the lead of most qualitative researchers: *Actively restrain yourself from forming any* a priori *hypotheses about what you might find in your own study*.

Narrative Inquiry

Narrative inquiry is a method for studying complex, multifaceted phenomena—in qualitative lingo, very *rich* phenomena—by focusing on the recollections and stories of individuals who have had experiences related to these phenomena (Clandinin & Connelly, 2000). Rooted in a variety of disciplines, such as sociology, anthropology, and psychology, narrative inquiry methods are influenced by the theory and writings of John Dewey (1934, 1938) on

[3] Some researchers associate the term *grounded theory* with a particular method of data analysis—in particular, that of Corbin and Strauss (2008; Strauss & Corbin, 1990)—and suggest the term *emergent theory* as a broader, less prescriptive label for this approach (e.g., Jaccard & Jacoby, 2010).

the nature of experience. A central notion of narrative studies is a "three-dimensional narrative space"—comprised of the concepts of *interaction, continuity,* and *situation*—that acknowledges the ways in which personal and social experiences (a) interact with one another, (b) occur along specific time dimensions, and (c) are located within particular places and situations. Narrative studies don't just focus on events in the "here and now" but also seek to understand the past and future dimensions of events and experiences.

Narrative studies have explored a variety of topics and experiences; for example, they have been helpful in clarifying the nature of multicultural teacher education and of Chinese teachers' acculturation experiences (He, 1999; Phillion, 2002). Narrative inquirers and researchers collect a variety of data to inform their investigations, including autobiographical writings, oral histories, letters, conversations and dialogues, interviews, family stories, photographs, poetry, and artifacts. Some researchers also make use of timelines created by participants, as well as sketches and chronicles of particular life events and stages. Narrative researchers often draw on their own life experiences to help them interpret and make better sense of participants' beliefs and stories.

In their analyses of their data, narrative researchers search for common themes, threads, and tensions, often working collaboratively with participants to make sense of the data. Doing so might include examining and responding to written narratives or engaging in discussions about pivotal events. Accordingly, the process of narrative analysis is often referred to as **restorying**, because researchers gather various narratives, field texts, and other data sources, analyzing them for key themes and elements, and then reordering them—often chronologically—to craft rich narratives and stories that provide a better understanding of the setting, context, time period, and/or sequence of events (Ollerenshaw & Creswell, 2002). Some approaches to narrative analysis seek to identify syntactical features and structural attributes of the narrative that reveal underlying meanings about the speaker's purposes, intentions, and beliefs (Mishler, 1995). For example, researchers might attend to a participant's use of first-person pronouns and possessive adjectives (e.g., *I, we,* or *our*) as indicative of ownership over or connectedness to a particular topic or issue, whereas the use of third-person pronouns and possessive adjectives (e.g., *they* or *their*) might suggest a more detached perspective. Narrative studies are often written up in a narrative style, include elements of storytelling, and weave together the researcher's own experiences and reflections with a variety of data sources.

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.[4] For example, a researcher might use a content analysis to determine what religious symbols appear in works of art, how middle school science texts portray the nature of science, 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 the fine arts, education, history, psychology, journalism, and political science.

Of the six designs described in this chapter, a content analysis is apt to involve 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?").

[4] Again we refer you to Kraut and colleagues (2004) regarding ethical issues related to studying people's postings on the Internet.

Furthermore, the researcher takes measures 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 objective—for instance, 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—for instance, when the study involves categorizing discrete sections of textbooks as conveying various messages about the nature of science—two or three raters are typically involved, and a composite of their judgments is used.

Content analyses aren't 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 complex, multidimensional aspects of a descriptive or experimental study, resulting in a *mixed-methods design* with both qualitative and quantitative elements.

Even when a content analysis *is* the sole research methodology, it's apt to have a quantitative component. In many instances, quantification may involve simply counting the frequencies with which various characteristics are observed in the body of data being examined. But alternatively, a researcher might conduct one or more statistical analyses on the numbers obtained—for instance, comparing the numbers obtained from two or more distinct subsets of the materials being analyzed (e.g., see Bergman, 2010).

Table 8.1 summarizes the nature of the six designs described in the preceding sections, including a brief description of general data-analysis strategies. Keep in mind, however, that the six designs aren't necessarily as distinctly different as the table might indicate. Any particular study

TABLE 8.1 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(s)	• 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 the beliefs and practices of a cultural group	A specific field site (e.g., school, organization, community, virtual group) in which a group of people share a common culture	• Participant observation • Semistructured or unstructured interviews with "informants" • Artifact/document collection • Surveys and questionnaires	• 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, semistructured interviews • Purposeful sampling of individuals	• Search for *meaningful concepts* that reflect various aspects of the experience • Integration of those concepts into a seemingly typical experience

(continues)

TABLE 8.1 Distinguishing Characteristics of Different Qualitative Designs (*continued*)

Design	*Purpose*	*Focus*	*Methods of Data Collection*	*Methods of Data Analysis*
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 • Theoretical sampling in order to saturate theoretical categories • Any other relevant data sources	• 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
Narrative inquiry	To explore rich experiences and stories of individuals	A set of experiences shared by a few or more individuals across time periods and contexts	• In-depth unstructured and semistructured interviews • Autobiographical writings, poetry, and other narrative documents • Photographs and other visual materials	• Identification of narrative themes, threads, and tensions • Reordering of narratives into a coherent chronological or thematic story
Content analysis	To identify the specific characteristics of a body of material	Any textual, 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 address the research question

may include elements of two or more qualitative designs. Remember, much qualitative research is, by its very nature, somewhat flexible and 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.

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 a 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 six potential research projects. Identify the qualitative methodology that is probably most appropriate for each project. The answers appear after the "For Further Reading" section at the end of the chapter.

1. In an effort to learn the nature and appeal of long-standing men's social groups, a researcher plans to spend a 9-month period with a local chapter ("lodge") 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 shown 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 (e.g., body language). The researcher wants to find out what it is 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.
6. A researcher is interested in the life histories and professional trajectories of African American women in university leadership positions and how these experiences might provide insights about the nature of leadership and mentorship for African American women in academic institutions.

MyLab Education Self-Check 8.2

MyLab Education Application Exercise 8.1: Recognizing Qualitative Designs

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 questions. Potential sources of data are limited only by a researcher's open-mindedness and creativity. For example, in a school setting, a researcher might consider where various students are seated in the lunchroom, 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, 2013).

While collecting data, many qualitative researchers also begin jotting notes—sometimes called **memos**—about their initial interpretations of what they are seeing and hearing. Such memos tend to be one of three types:

1. *Reflective memos* are akin to the reflective journals previously mentioned in the section about ethnographies. They often address the researcher's **positionality**—a term that, in the lingo of qualitative research methodologies, refers to a researcher's background characteristics and experiences, past and current social roles, and personal biases as these relate to the participants, setting, and/or research topic. Qualitative researchers acknowledge that the particular positions they hold in their social environment (a) have impacted their past experiences and general paths in life and (b) are currently shaping their relationships with participants and their views of the research topic and setting.
2. *Methodological memos* can help researchers keep track of breakthroughs related to certain methodological challenges or dilemmas—for instance, they might describe effective strategies for building rapport with initially resistant participants.
3. *Analytical memos* might involve discussions of emerging themes in people's actions or statements, or they might make note of initial hunches and intuitions that a researcher should pursue through further observations or interview questions. Still other analytical memos might describe preliminary theories about possible underlying dynamics within a social group.

Many qualitative studies are characterized by an **emerging design**, in which data collected early in the investigation influence the kinds of data the researcher subsequently gathers. 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 determine how best to adjust data-collection strategies midway through a study. Thus, many experts suggest that a novice researcher set forth a definite, fairly structured plan for data collection—a strategy that can minimize the degree to which the researcher wanders off into intriguing but ultimately unproductive diversions.

A predetermined, well-thought-out plan is also essential when submitting a qualitative research proposal to an institutional review board (IRB). Most importantly, data-collection methods must be consistent with the ethical principles presented in Chapter 4. The researcher must take precautions not to expose people (or animals) to unnecessary physical or psychological harm—as could happen, say, if the researcher were to inquire about highly personal and emotionally charged topics. 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 keeping personal data *confidential* and maintaining participants' *right to privacy*). One common way of keeping personal data confidential is to (a) assign various pseudonyms to different participants and/or settings and institutions and (b) use those pseudonyms both during data collection and in the final research report. Sometimes a researcher might change some of the details or characteristics associated with a participant—for example, by describing a math teacher as a "science teacher"—in order to conceal the person's identity. This strategy can be especially important in small research settings in which participants have unique roles and characteristics.

PRACTICAL APPLICATION Ensuring That Qualitative Data Collection Yields Credible, Trustworthy Data

As you should recall, Chapter 4 includes a section called "Validity and Reliability in Assessment," which details ways to safeguard and strengthen the accuracy and consistency of quantitative assessment results. Qualitative researchers don't necessarily *measure* things—at least not in the numerical sense of the word—but they need to ensure that their research is trustworthy, that data are collected ethically and accurately, and that findings are credible, plausible, and well-substantiated (Shenton, 2004). Many qualitative researchers reject the terms *validity* and *reliability,* viewing them as inappropriate concepts for evaluating the rigor of qualitative research studies. They instead embrace notions such as **credibility** and **transferability** as criteria for judging whether the results from a study are plausible and believable from participants' perspectives and if the findings can be applied to other settings (Lincoln & Guba, 1985; Shenton, 2004). And whereas quantitative assessment strategies are judged to be *reliable* if they consistently yield similar results, many qualitative researchers prefer to speak of **dependability**—a term that accounts for the ever-changing contexts within which qualitative research studies take place—and urge one another to provide in-depth descriptions of data-collection methods.

A particular strength of qualitative methods is that a perceptive researcher might discern underlying patterns and dynamics in social interactions or cultural artifacts that a standardized, quantitative assessment instrument would never illuminate. In a sense, *the researcher is an instrument* in much the same way that an oscilloscope, questionnaire, or multiple-choice achievement test is an instrument. The potential downside of this instrument—the human mind—is that it can be biased by its preconceived theories and expectations, and such biases can adversely affect the quality of the data obtained. Accordingly, rather than strive for complete objectivity in data collection (which is probably an impossible goal), qualitative researchers adhere to the standard of **confirmability**; that is, they make a concerted effort to base their conclusions on their *actual data* as much as possible—and to describe their

data-collection and data-analysis processes in considerable detail—such that other researchers might draw similar conclusions from similarly collected and analyzed data.

Qualitative researchers use a variety of strategies to enhance the trustworthiness and credibility of the data they collect. Following are five important strategies during the *data-collection* phase of a qualitative study (we identify strategies related to *data analysis and interpretation* in Chapter 12):

- ***Reflexivity.*** Good qualitative researchers actively try to identify personal, social, political, or philosophical biases that are likely to affect their ability to collect and interpret data—this self-reflection is known as **reflexivity**—and take whatever steps they can to reduce such influences.
- ***Triangulation.*** Many qualitative researchers use a strategy they call **triangulation**:[5] They collect multiple forms of data related to the same research question, with the goal of finding consistencies or inconsistencies among the data. For example, imagine that a researcher wants to study the behaviors of an especially exclusive group of snobbish but so-called "popular" girls at a public high school. This researcher might not only interview both members and nonmembers of the group but also observe the girls in action in various locations in and around school—for instance, observing seating patterns in the cafeteria, group clusters in the hallways and schoolyard, and verbal interaction patterns during class sessions. The researcher might also scan school records regarding which students are members (and possibly officers or captains) of various extracurricular clubs and sports teams.
- ***Clearly distinguishing between data and reflections/memos.*** Right from the get-go, a qualitative researcher must keep interpretations separate from actual observations. For example, consider the ethnographic researcher who decides to take only handwritten notes in the field, perhaps as a way of blending in better with the social environment than would be possible with, say, a laptop or video camera. This researcher might draw a vertical line down the middle of each page, recording observations, interview responses, and any helpful graphics (e.g., maps, diagrams) in the left column and jotting memos about these things in the right column. Only in this way can the researcher separate descriptive data (what the researcher is actually seeing and hearing) from the researcher's interpretations and reflections (what the researcher currently thinks might be going on).
- ***Seeking of exceptions and contradictory evidence.*** By nature, human beings seem to be predisposed to look for and identify patterns and consistencies in their physical worlds (e.g., see Mandler, 2007; Rakison & Oakes, 2003). Furthermore, once they have zeroed in on their conclusions about these patterns and consistencies, they're often reluctant to revise their beliefs (recall the discussion of *confirmation bias* in Figure 1.3 in Chapter 1). A good qualitative researcher actively fights such mental predispositions, in part by continually asking the questions "Might I be wrong?" and "What disconfirming evidence can I find?" and then intentionally seeking out the answers. In many cases, discrepant data and cases might help a researcher identify more complex and nuanced patterns and thus develop more complete explanations of the phenomenon being studied (Erickson, 1986; Lawrence-Lightfoot & Davis, 1997). Remember that in qualitative studies, researchers are usually not intent on seeking an absolute "truth" or single answer related to a research question but instead are interested in uncovering multiple meanings and often divergent perspectives that can help us all better understand an issue, process, belief system, or experience.
- ***Spending considerable time on site.*** Many qualitative studies require extensive data collection in the field and immersion in the chosen setting or site; such is true for virtually any ethnography and for many case studies, phenomenological studies, and grounded theory studies. Just a brief visit to the site under investigation—popping

[5] Here we are more specifically talking about *data triangulation*—a term that some researchers prefer to use in order to contrast it with other triangulation strategies that we don't talk much about in this book.

in and popping out, as it were—is unlikely to yield the quantity and quality of data (including potentially contradictory observations) essential for drawing accurate, multifaceted understandings of any complex phenomenon.

- ***Member checking and audit trails.*** Throughout the process of data collection and analysis, qualitative researchers can incorporate a number of procedures and measures into their studies that provide additional "checks and balances" that can help enhance the credibility of the study. One of these is **member checking**, in which a researcher asks participants to review interview transcripts in order to double-check their accuracy; sometimes a researcher also seeks participants' feedback and possible validation regarding findings. Member checking doesn't necessarily mean that a researcher would completely change findings based on participants' responses and reactions, but participants might provide important insights that (a) could be included in analyses or even (b) take the research effort in new directions. Differing participant interpretations might be included in the final research report.

  A second checks-and-balances strategy is to enlist the help of a colleague to review coding and analysis and thus to furnish an additional "set of eyes" and possibly an alternative perspective. The goal of this strategy wouldn't necessarily be to achieve consensus but rather to ensure that the researcher isn't wildly off course or "making stuff up."

  Still a third strategy is an **audit trail**, which entails logging and detailing data-collection and data-analysis activities as they occur. A good audit trail provides a record that can both verify research activities and offer information to others who might be interested in undertaking a similar study.

MyLab Education Self-Check 8.3

MyLab Education Application Exercise 8.2: Enhancing Credibility and Trustworthiness

In planning for data collection, qualitative researchers must also identify one or more appropriate *samples* from which to acquire data. Furthermore, they are apt to rely heavily on *observations* and/or *interviews* as sources of data. We offer suggestions related to each of these topics in the three Practical Application sections that follow. Many of our suggestions can, in one way or another, enhance the trustworthiness of the data obtained.

PRACTICAL APPLICATION Selecting an Appropriate Sample for a Qualitative Study

Qualitative researchers might draw their data from a variety of sources—not only from people but perhaps also from objects, text materials, and audiovisual and electronic records. The particular entities they select for analysis comprise their sample.

Only rarely—for instance, when a study involves a content analysis of a small number of items—can qualitative researchers look at *everything* that has potential relevance to a research problem. More typically, they must be choosy about the data they 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. Should you need such a sample in a qualitative study, we suggest that you return to the discussion of *probability sampling* (including random sampling) in Chapter 6. For most qualitative researchers, however, using probability sampling to identify participants isn't possible, or even desirable, especially if the goal is to gain insights from people who have particular roles at a site or unique perspectives relative to the research topic. For example, a researcher who wants to study social and political processes involved in the implementation of a new school policy or program would almost certainly want to talk to the school principal, as well as to teachers who are closely involved in or affected by the policy or program.

Thus, most sampling in qualitative studies is *purposive sampling*—that is, it entails choosing those individuals or objects that will yield the most information about the topic under investigation. As you should recall from Chapter 6, some descriptive quantitative researchers also engage in purposive sampling, but qualitative researchers often use more specific terms to describe their sampling procedures. For example, grounded theory researchers tend to engage in **theoretical sampling**, choosing data sources that are most likely to help them develop a theory that can help to explain the phenomenon they are studying. Later, they may employ **discriminant sampling**, returning to particular data sources that can help them substantiate a theory. In contrast, ethnographers often rely on *primary informants* who can both (a) provide considerable information about the group being studied and (b) facilitate access to other group members and their activities. Case study researchers might focus on individuals who illustrate important features of the case and can be studied or observed more intently; this strategy of selecting focal participants is especially important when researchers are conducting observations in classrooms or other active settings, when it isn't possible to pay attention to every individual. One type of sampling used in case studies is **extreme case sampling**, in which a researcher selects individuals who have unique or deviant characteristics within a case; the researcher would then focus observations on those individuals, conduct in-depth interviews with them, or both.

Some qualitative researchers face constraints over which they have no control, such as distance or access issues (e.g., perhaps a gatekeeper denies entrance to a site). In such situations, they may employ **convenience sampling**,[6] in which they include readily available research sites and/or easily accessible individuals who can provide insights related to central research questions. In studies where researchers have exceptional difficulty recruiting participants, they might utilize **snowball sampling** (a.k.a., *chain sampling*), which involves asking existing participants to refer or recruit other individuals (e.g., acquaintances or colleagues) who also have experiences and knowledge relevant to the topic under study.

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 several suggestions to guide decision making:

- Be sure that the sample includes not only seemingly "typical" but also seemingly "nontypical" examples.
- When a power hierarchy exists—as it does in the workplace and in many clubs and communities—sample participants from various levels in the hierarchy. For example, in the workplace, you might interview both bosses and employees; in a club or

[6] Be careful that you don't confuse convenience sampling in qualitative research with the convenience sampling described in Chapter 6. In *quantitative* research, convenience sampling simply involves recruiting easily available participants; it doesn't necessarily involve choosing participants who are apt to be especially insightful.

community, you might interview not only highly active, influential members but also less involved individuals (e.g., see H. C. Becker, 1970).
- Actively look for cases that might disconfirm emerging hypotheses and theories.
- If appropriate for your research problem, sample from diverse contexts or situations.

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.

In some instances, a research problem or topic is best addressed by sampling from a large geographical area, perhaps one that includes diverse cultural groups. For example, in a dissertation project involving the experiences of White women who were raising biological children of mixed or other races,[7] doctoral student Jennifer Chandler (2014) wanted to interview mothers from diverse locations across the United States—locations that would differ in demographic makeup and possibly also in attitudes regarding multiracial families. To obtain such a sample, she created an "Invitation to Participate" letter that described the purpose of her study, the characteristics of desired participants, and the general nature of the interviews she would conduct. Many individuals across the country helped her distribute the invitation, including (a) personal friends and colleagues, (b) people she met at several professional conferences, (c) officers in parent–teacher organizations in numerous public school districts (e.g., Los Angeles, Houston, Denver, New York), and (d) people who had contributed to Internet blogs about topics related to interracial parenting. The resulting sample included 30 mothers from towns and cities in more than a dozen states across the country. It was certainly not a random sample, but it helped Chandler capture the diversity in experiences that mothers living in various geographical and cultural settings were likely to have had.

PRACTICAL APPLICATION Making Observations in a Qualitative Study

In the observation studies described in Chapter 6, observations typically have a limited, prespecified focus, and procedures are set in place in advance for quantifying the observations in some way, perhaps with a rating scale. In contrast, 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. Observations are often 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 can construct a complex yet integrated picture of how certain humans or nonhumans spend their time.

Such an approach has its drawbacks, of course. A 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

[7]More precisely, the sample included mothers who (a) *identified* themselves as being non-Hispanic White women and (b) *identified* their children as being of mixed or other races. Chandler's (2014) capitalization of "White" when referring to a racial group is consistent with American Psychological Association (APA) style.

area. And the very presence of tape recorders and video cameras may make some participants uncomfortable; it also poses challenges to confidentiality.

If you decide to conduct observations as part of a qualitative study, we offer these recommendations:

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 diverse contexts.

2. When you first enter a research site, have someone introduce you to the people you hope to observe. This is the time to briefly describe your study and get potential participants' informed consent.

3. Decide what your role will be as you observe, and discuss this role with any gatekeepers before you conduct your observations. For example, if you are observing an elementary school classroom, make sure that the teacher doesn't view you as "extra help" and expect you to assist with activities or help discipline students. However, there may be occasions when you want to take a more active role and it makes sense for you to do so, such as during a community service event. Taking a more active role can also help you build rapport with and develop relationships with participants.

Also remember a strategy alluded to earlier: *Clearly distinguish between your actual observations (data) and your interpretations (memos).* This strategy is important for two reasons. First, you need to be as objective and descriptive as you can in the records you keep, thereby minimizing the extent to which your actual data are little more than subjective impressions. And second, your interpretations of what you have seen and heard may very well change over the course of the study.

PRACTICAL APPLICATION Planning and Conducting Interviews in a Qualitative Study

Interviews can often yield a rich body of qualitative information. A researcher might 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)

Interviews in a qualitative study tend not to be as tightly prescribed and structured as the interviews conducted in a quantitative study. A second difference 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 often someone they have come to know and trust. In contrast, participants in quantitative survey research are continually aware that, yes, this is an interview and that any temporary relationship they've formed with the researcher will end after the interview is complete.

In some cases, a qualitative 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, fosters debate and discussion on pertinent issues, 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, 2013; Neuman, 2011).

Focus groups can be either homogeneous (i.e., comprised of people who share a common characteristic, such as a certain professional role, age group, or gender) or heterogeneous (i.e., comprised of people with diverse characteristics). Heterogeneous focus groups can be especially helpful as a means of getting diverse perspectives on a complex issue.

GUIDELINES Conducting a Productive Interview

Conducting an informative interview isn't as easy as it might seem. The following suggestions are based partly on our own experiences and partly on guidance offered by experts in qualitative research (Creswell, 2013; Eisner, 1998; Shank, 2006; Silverman, 1993).

1. ***Identify general interview questions and possible follow-up subquestions in advance.*** Some experienced qualitative researchers are quite skillful at conducting open-ended, unstructured interviews. As a result, they can gain intriguing information and perspectives they hadn't planned to ask for. However, a major disadvantage of an unstructured interview is that a researcher might ask different questions of different participants and thus may not be able to make cross-participant comparisons. Furthermore, the researcher must be alert to instances when a conversation is drifting in an unproductive direction and gently guide it back on course.

Novice researchers typically have greater success when they (a) prepare their general interview questions in advance, as part of an **interview guide**, along with possible follow-up questions that probe for details; and (b) make sure that all questions are addressed at some point during the interview. Such planning increases the odds that a researcher can compare the responses of different participants in the event that certain comparisons are desired. Furthermore, IRB approval of a research project may in some instances require that questions be explicitly laid out in the initial research proposal.

Obviously, interview questions should 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 (2003) examined how children with reading disabilities think 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. In planning her study, she posed five general research questions, each of which she subsequently used to identify her interview questions; she spelled out her questions in an appendix in her dissertation. Figure 8.1 presents Zambo's first three research questions (two of which she broke down into more specific subquestions) and their relevant interview questions (see Columns 1 and 2, respectively). Note, however, that her interview questions are written in the third person, using the pronoun *they.* Thus, we have translated her interview questions into second-person (*you*) questions that she might have specifically asked as she spoke with the children (see Column 3).

For any single interview, limit your list of questions to a number that can be covered in 45 to 90 minutes; the questions can be as few as 5 and as many as 20, depending on their scope. (Although Debby Zambo had many more questions than this, she spread them throughout a series of 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.

Research Questions (from Zambo, 2003)	*Interview Questions (from Zambo, 2003)*	*Translating Zambo's "They" Questions into "You" Questions*
1. What do students with reading disabilities think about reading and themselves?		
a. What are their thoughts about 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 (does a person) become good/poor at reading?	"What does it mean to *read* something?" "What kinds of things are easy for you to read?" "What kinds of things are hard for you to read?" "What things do good readers do when they read?" "How does a person *become* good at reading?"
b. What are their ideas about themselves and 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 impacted or will impact) their lives in the past, present, and future?	"What kinds of things are you reading now?" "Which of these things are easy for you to read?" "Which ones are hard for you to read?" "What goes on in your head when you read easy things?" "What goes on in your head when you read things that are hard for you to understand?" "What kinds of things would I see if I was watching you while you read?" "In what ways has reading helped you so far in your life?" "How do you think reading might help you in the 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?	"What kinds of feelings do you have when you read?" "Do you sometimes get frustrated when you read? If so, why do you think that happens?" "Does believing that you can get better at reading help you be a better reader?" "Does wishing and hoping you can get better at reading help you 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?	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?	"What does it mean to 'pay attention' to something?" "What specific things do you pay attention to when you read? Why do you focus on those things?" "Do you have any trouble keeping your mind on what you're reading? If so, what kinds of things do you do to get your mind back on your reading?"

(*continues*)

FIGURE 8.1 Example of How a Researcher Might Align Interview Questions with Research Questions

Source: Columns 1 and 2 are 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 Questions (from Zambo, 2003)	*Interview Questions (from Zambo, 2003)*	*Translating Zambo's "They" Questions into "You" Questions*
	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?	"Are there some days when you have a lot of trouble paying attention? If so, why is it so hard for you to read on those days?" "What kinds of things distract you when you read?" "Are there times when ideas or memories in your head distract you while you're reading?
b. What do they know about their memory and reading?	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?	"What is *memory*? How would you describe it to someone else?" "What do you do when you want to put something into your memory?" "What kinds of things help you *keep* stuff in your memory?" "Do you do anything special to help you understand and remember what you're reading?" "What kinds of things do you do in your head when you're trying to remember something you read a few weeks ago?"

FIGURE 8.1 Example of How a Researcher Might Align Interview Questions with Research Questions *(continued)*

Ideally, interview questions encourage people to talk about a topic *without* hinting that they should give a *particular* answer. In other words, avoid leading questions. Questions such as "What's going on now?" "What's 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 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, some cultural groups may avoid bragging about their personal accomplishments. A naive researcher might erroneously conclude that such individuals are less productive than others, when in reality they are merely less *boastful*. Also, some groups may be more tentative in expressing their opinions or unlikely to reveal their emotions during conversations (Morelli & Rothbaum, 2007; Ward, Bochner, & Furnham, 2001).

Various cultural groups differ, too, in their general verbal interaction patterns—for instance, in how talkative and assertive they are and in how much physical distance (*personal space*) they prefer when conversing with another person (e.g., see Tyler et al., 2008; Ward et al., 2001). Given such diversity across cultural groups—and often among certain subgroups of a cultural group—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 sample includes people who will give you the kinds of information you are seeking.*** You should, of course, 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 who might give you unique insights related to your research problem. When you do so, however, 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. You should also be mindful of confidentiality issues when selecting an appropriate interview locale.

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, if required for the study. You might 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. Ideally, you should try to motivate people to *want* to participate in your study. For example, in a dissertation research project with Chinese mothers who had immigrated to the United States, doctoral student Christy Leung and her research team began interviews by following this general script:

> As you know, we are doing this project because we really want to help Chinese families succeed in the U.S., and to have the children be happy and successful. So, we are trying to learn from the different experiences of the families, in order to find patterns of things that can help the transition, and also things that can be negative for the transition and ways to avoid those negative things. That is why we really appreciate your participation and sharing.
>
> After we started this project, we also found that there are aspects of the families' experiences that we were not able to understand in detail through some of the questionnaires that we are using. Many parents really wanted to talk about and share their experiences regarding how they and their children are doing in the U.S. We feel that these rich experiences (both positive and negative) can add so much to our understanding of the issues that Chinese families in the U.S. are facing. (Leung, 2012, p. 295)

Because interviews in qualitative studies tend to be rather informal, they might 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 concrete and 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*—that is, if you ask about actual behaviors. For example, if you're 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 & Brinkmann, 2009).

9. ***Record responses verbatim.*** Whether you use handwritten notes, shorthand, a tape recorder, smartphone application, or laptop computer, capture everything the person says, especially if interview questions are fairly open-ended. 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, opinions*, and *constructions* of events, rather than as actual facts and "truth."

12. ***When conducting a focus group, take group dynamics into account.*** Whenever you gather two or more people 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're shy or feel uncertain about the validity of their perspective. 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.

An Example in International Relations

A student researcher wanted to interview certain United Nations personnel to get their opinions concerning issues related to his study. He planned to travel 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 his trip, the student wrote to 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 created a master chart of list 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 along with another copy of the interview questions. He also enclosed his full interview schedule—*without* revealing other participants' names—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 $3^1/_2$ 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. Later, he sent relevant sections of the semifinal draft of his final report to—once again—get each interviewee's approval. 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.

Using Technology to Facilitate Collection of Interview Data

With appropriate software, most laptops and many smartphones can serve as audio recorders. And, of course, videos recorded on a camcorder can be easily downloaded to a personal computer. Meanwhile, transcription software (e.g., 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 voice-recognition software programs (e.g., Dragon Naturally Speaking) can help with transcriptions but will need to be trained to your voice, and final transcripts must be reviewed for accuracy.

In some cases, you can conduct long-distance qualitative interviews through various Internet mechanisms, including e-mail, Skype, or video conferencing. Focus groups might also be conducted online, perhaps through Internet-based chat rooms or bulletin boards (e.g., see Krueger and Casey [2009] for suggestions). Keep in mind, however, that ethical standards don't fly out the window simply because you're 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.

MyLab Education Self-Check 8.4

MyLab Education Application Exercise 8.3: Collecting Data for a Qualitative Study

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, 2013; 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 a 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 ultimately gives readers an in-depth, multifaceted picture of the phenomenon (i.e., *thick description*; see Chapter 4).
6. ***Coherence.*** The data yield consistent findings, such that the researcher can present a portrait that "hangs together." Multiple data sources converge on consistent conclusions (*triangulation*), and any contradictions within the data are discussed and interpreted.
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 a better understanding of the phenomenon, enable more accurate predictions about future events, or lead to interventions that enhance the quality of life.

In this chapter we have addressed issues related to only some of these criteria—especially issues related to purposefulness, rigor, and open-mindedness. We address issues related to other criteria in discussions of data analysis in Chapter 12 and report writing in Chapter 13.

PRACTICAL APPLICATION Planning the Logistics of a Qualitative Study

As should be clear by now, a qualitative research project is not something to be entered into casually. One key consideration is that, 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, sketches, photographs, audio recordings, videos, or some combination of these. And as you will discover in Chapter 12, data organization and analysis must be equally meticulous and time-intensive.

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 relative 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 listed in the "For Further Reading" section 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 Methodology 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, narrative inquiry, content analysis, a combination of two or more of these, or none of these?

What Data Are Needed?

_____ 5. Will you need to gain access to one or more sites in the field? If so, how will you do it?

_____ 6. How much time will you need?

_____ 7. What special 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 confidentiality for the participants?

_____ 12. What procedures will you follow, and in what order?

_____ 13. What technological tools can assist you in data collection?

MyLab Education Self-Check 8.5

A SAMPLE DISSERTATION

As an example of a qualitative research study, 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 also looked 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 5

DATA COLLECTION PROCESSES

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 . . . 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. . . .

Comments

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.

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.

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.

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!

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 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.

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.

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.

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.

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.

The author conducted an earlier pilot study*—something we urge any beginning researcher to do, particularly when planning a qualitative 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

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.

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.

The ability to contrast one situation with another is a key advantage of a multiple case study.

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 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". . . .

Here we see open coding, *the first step in data analysis in a grounded theory study (see Chapter 12).*

*NUD*IST was an early, groundbreaking computer database program especially suited for data collection in qualitative research.*

(The author continues the discussion of data coding and other issues and then turns to the subject of values.)

Here data analysis has moved on to axial coding, *where the author is refining her categories and their interconnections.*

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. . . .

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.

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. . . .

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.

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.

Here the author is looking for disconfirming evidence, one effective strategy for minimizing the influence of a researcher's biases on data interpretation.

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 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.

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.*

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.

Note: 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.

APPLICATION EXERCISES

Similar exercises are available on Pearson MyLab Education.

1. For each of the following sample research purpose statements, determine if the study is qualitative. If it is qualitative, briefly describe the context and focus of the study and the type of qualitative study that is most likely.
   - **i.** Research has consistently indicated that the transition to college life is difficult. The purpose of this study is to describe the typical first-year experience for women at a small liberal arts college.
   - **ii.** As children age and gain experiences, their ideas of future occupations change as well. The purpose of this study is to discuss occupational interests with children of varying ages.
   - **iii.** Professor Garcia is developing a new kindergarten curriculum. She approaches a private school and asks the faculty to implement the program in one class for 1 year. During that year, she visits the school as a participant observer, interviews the teachers, and collects examples of children's work.
   - **iv.** A researcher is interested in the sorts of messages around discipline contained within parenting manuals and how these reflect changing attitudes about and perspectives on child development and parenting over the last several decades (1950s–present).
2. A researcher is interested in designing a qualitative case study in an elementary school, examining fifth- and sixth-grade students' use of technology for instructional and social purposes. Thus far, she is planning on interviewing students, teachers, and staff to gain insight into technology use. She wants to make sure her study is credible and the data are trustworthy.
   - **i.** As she puts together her proposal, what other forms of data might she gather?
   - **ii.** As she collects data, how can she further enhance the credibility and trustworthiness of the study? Identify notions related to enhancing credibility that might be relevant to this study and specific ways that these can be addressed within the study.
3. Consider each of the following scenarios, and answer the accompanying questions.
   - **i.** Imagine that you want to conduct a content analysis to study the nature of student interactions in an inclusionary classroom—that is, one in which students with disabilities are educated along with their nondisabled peers. Qualitative researchers who use content analysis as a research design sometimes use documents and artifacts as data. Give one concrete example of each of these types

of records, and explain how you could use it to address the nature of student interactions in an inclusionary classroom.

ii. Emily is an undergraduate college student who is interested in conducting fieldwork for an ethnographic study of social interactions among high school girls. Give at least three specific examples of the types of data Emily might collect for this type of study.

iii. Should Emily approach her study as a participant observer? Explain your answer, paying attention to the benefits and pitfalls of this method.

FOR FURTHER READING

Barlow, D. H., & Nock, M. K. (2009). Why can't we be more idiographic in our research? *Perspectives on Psychological Science, 4,* 19–21.

Birks, M., & Mills, J. (2015). *Grounded theory: A practical guide* (2nd ed.) Los Angeles, CA: Sage.

Bloomberg, L. D., & Volpe, M. F. (2016). *Completing your qualitative dissertation: A roadmap from beginning to end* (3rd ed.). Los Angeles, CA: Sage.

Bogdan, R., & Biklen, S. K. (2007*). Qualitative research for education: An introduction to theory and methods* (5th ed.). Columbus, OH: Pearson.

Brinkmann, S., & Kvale, S. (2014). *InterViews: Learning the craft of qualitative research interviewing* (3rd ed.). Los Angeles, CA: Sage.

Butler-Kisber, L. (2010). *Qualitative inquiry: Thematic, narrative and arts-informed perspectives.* London, England: Sage.

Charmaz, K. (2014). *Constructing grounded theory: A practical guide through qualitative analysis* (2nd ed.). Los Angeles, CA: Sage.

Clandinin, D. J., & Connelly, F. M. (2004). *Narrative inquiry: Experience and story in qualitative research*. San Francisco, CA: Jossey Bass.

Clifford, J., & Marcus, G. E. (1986) *Writing culture: The poetics and politics of ethnography*. Berkeley, CA: University of California Press.

Corbin, J., & Strauss, A. C. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Los Angeles, CA: Sage.

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Columbus, OH: Pearson.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Los Angeles, CA: Sage.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). Los Angeles, CA: Sage.

Denzin, N. K., & Lincoln, Y. S. (2013). *Strategies of qualitative inquiry* (4th ed.). Los Angeles, CA: Sage.

Dyson, A.H., & Genishi, C. (2005). *On the case: Approaches to language and literacy research*. New York, NY: Teachers College Press.

Eisner, E. W. (1998). *The enlightened eye: Qualitative inquiry and the enhancement of educational practice.* Upper Saddle River, NJ: Prentice Hall.

Ellis, C., Adams, T. E., & Bochner, A. P. (2011). Autoethnography: An overview. *Forum: Qualitative Social Research, 12*(1), Article 10.

Emerson, R. M., Fretz, R. I., & Shaw, L. L. (2011). *Writing ethnographic field notes* (2nd ed.). Chicago, IL: University of Chicago Press.

Erickson, F. (1986). Qualitative methods in research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 119–161). New York, NY: MacMillan.

Fetterman, D. M. (2010). *Ethnography: Step by step* (3rd ed.). Los Angeles, CA: Sage.

Gast, D. L., & Ledford, J. R. (2014). *Single case research methodology: Applications in special education and behavioral sciences* (2nd ed.). New York, NY: Routledge.

Geertz, C. (1973). *The interpretation of cultures*. New York, NY: Basic Books.

Gibson, B., & Hartman, J. (2014). *Rediscovering grounded theory.* London, England: Sage.

Guest, G., Namey, E. E., & Mitchell, M. L. (2013). *Collecting qualitative data: A field manual for applied research.* Los Angeles, CA: Sage.

Hammersley, M. (2008). *Questioning qualitative inquiry: Critical essays.* Los Angeles, CA: Sage.

Hammersley, M., & Atkinson, P. (2007). *Ethnography: Principles in practice* (3rd ed.). New York, NY: Routledge.

Hatch, A. J. (2006). *Early childhood qualitative research.* New York, NY: Routledge.

Hays, D. G., & Singh, A. A. (2012). *Qualitative inquiry in clinical and educational settings.* New York, NY: Guilford Press.

Heidegger, M. (2005). *Introduction to phenomenological research* (D. O. Dahlstrom, Trans.). Bloomington, IN: Indiana University Press.

Hesse-Biber, S. N., & Leavy, P. (Eds.). (2011). *Handbook of emergent methods.* New York, NY: Guilford Press.

Jaccard, J., & Jacoby, J. (2010). *Theory construction and model-building skills.* New York, NY: Guilford Press. (See Chapter 10, "Grounded and Emergent Theory.")

James, N., & Busher, H. (2009). *Online interviewing.* London, England: Sage.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14–26.

Jorgensen, D. L. (2007). *Participant observation: A methodology for human studies.* Los Angeles, CA: Sage.

Josselson, R. (2013). *Interviewing for qualitative inquiry: A relational approach.* New York, NY: Guilford Press.

Kamberelis, G., & Dimitriadis, G. (2013). *Focus groups: From structured interviews to collective conversations.* New York, NY: Routledge.

King, N. (2010). *Interviews in qualitative research.* Los Angeles, CA: Sage.

Krueger, R. A., & Casey, M. A. (2015). *Focus groups: A practical guide for applied research* (5th ed.). Los Angeles, CA: Sage.

Latimer, J. (Ed.) (2003). *Advanced qualitative research for nursing.* New York, NY: Wiley.

Liamputtong, P. (2011). *Focus group methodology: Principles and practice*. Thousand Oaks, CA: Sage.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry.* Los Angeles, CA: Sage.

Mishler, E.G. (1995). Models of narrative analysis: A typology. *Journal of Narrative and Life History, 5*(2), 87–123.

Muncey, T. (2010). *Creating autoethnographies*. Thousand Oaks, CA: Sage.

Richards, L., & Morse, J. M. (2013). *Read me first for a user's guide to qualitative methods* (3rd ed.). Los Angeles, CA: Sage.

Rubin, H. J., & Rubin, I. S. (2012). *Qualitative interviewing: The art of hearing data.* Los Angeles, CA: Sage.

Salmons, J. S. (2014). *Qualitative online interviews* (2nd ed.). Los Angeles, CA: Sage.

Savin-Baden, M., & Major, C. H. (2013). *Qualitative research: The essential guide to theory and practice.* New York, NY: Routledge.

Schram, T. H. (2006). *Conceptualizing and proposing qualitative research* (2nd ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Seidman, I. (2013). *Interviewing as qualitative research: A guide for researchers in education and the social sciences*. New York, NY: Teachers College Press.

Smith, J. A., Flowers, P., & Larkin, M. (2009). *Interpretative phenomenological analysis: Theory, method and research.* London, England: Sage.

Stake, R. E. (2006). *Multiple case study analysis.* New York, NY: Guilford Press.

Stake, R. E. (2010). *Qualitative research: Studying how things work.* New York, NY: Guilford Press.

Strauss, A., & Corbin, J. (1994). Grounded theory methodology: An overview. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 273–285). Los Angeles, CA: Sage.

Tashakkori, A., & Teddlie, C. (Eds.). (2010). *Mixed methods in social & behavioral research.* Los Angeles, CA: Sage.

Yin, R. K. (2011). *Qualitative research from start to finish.* New York, NY: Guilford Press.

Yin, R. K. (2013). *Case study research: Design and methods* (5th ed.). Los Angeles, CA: Sage.

ANSWERS TO THE CONCEPTUAL ANALYSIS EXERCISE "Choosing a Qualitative Research Design":

1. The researcher wants to learn about the general *culture* of an Elks group; 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.
6. The researcher is interested in gaining insights into the past experiences and stories of a particular group, which would be well suited to a *narrative inquiry.*

This page intentionally left blank

Chapter 9 Mixed-Methods Designs

Many research problems and questions have both quantitative and qualitative dimensions. To fully address such problems and questions, then, the researcher must use both quantitative and qualitative techniques. Thus, quantitative and qualitative methodologies aren't necessarily a case of either–or; they're better described as a case of more–or–less.

Learning Outcomes

9.1 Identify situations in which mixed-methods designs are especially useful.

9.2 Describe general characteristics and purposes of (a) convergent designs, (b) exploratory sequential designs, (c) explanatory sequential designs, (d) longitudinal mixed-methods designs, and (e) multiphase iterative designs.

9.3 Describe effective strategies for planning and conducting mixed-methods research, including strategies related to (a) identifying research questions and hypotheses, (b) conducting a literature review, (c) choosing one or more samples, and (d) ensuring a study's overall credibility.

9.4 Explain how some mixed-methods research projects may require two or more separate proposal submissions to and subsequent approvals from an institutional review board; also explain how some mixed-methods designs may require extra precautions to ensure informed consent and protect participants' right to privacy.

Some research problems practically scream for both quantitative and qualitative data. These problems call for **mixed-methods research**. Such research involves not only collecting, analyzing, and interpreting both quantitative and qualitative data but also *integrating* findings from the two kinds of data into a cohesive whole.

As you may have noticed, this chapter is one of the shorter ones in the book. It's short not because mixed-methods studies are quick and easy—they definitely are *not* quick and easy—but because mixed-methods research draws largely on quantitative and qualitative research strategies addressed in previous chapters. For instance, 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 assessment 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 underlying the data, etc.)
- Calculating and drawing inferences from descriptive and inferential statistics
- Drawing and persuasively arguing for reasonable conclusions from qualitative data (through triangulation, negative case analysis, thick description, etc.)

Because a mixed-methods study requires both quantitative and qualitative research skills, it can be an especially challenging undertaking for a novice researcher. The trickiest part of mixed-methods research is in *combining* the two methodological traditions and *integrating* the two kinds of data 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."[1]

If you have the requisite skills and experience, however, we urge you to take on the challenge, because you might very well obtain a more complete, comprehensive answer to your research problem or question than would be possible with only quantitative or only qualitative methodologies.

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 on 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 only quantitative data, some call for only qualitative data, and some are best addressed with *both* kinds of data.

On average, a mixed-methods research study requires more of a 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 quantitative and qualitative 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 quantitative and qualitative 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 to a larger population) can possibly be replicated by administering a questionnaire to a larger, more representative sample of the population.
- ***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 the 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).

[1]Teddlie and Tashakkori (2010) have suggested that a study that *doesn't* effectively integrate findings and conclusions from quantitative and qualitative data might better be called a *quasi-mixed* study.

- ***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.
- ***Triangulation.*** A researcher can make a more convincing case for particular conclusions if both quantitative and qualitative data lead to those conclusions.

MyLab Education Self-Check 9.1

MyLab Education Application Exercise 9.1: Planning Mixed-Methods Research

GENERAL CATEGORIES OF MIXED-METHODS RESEARCH

In previous chapters we have touched on a few ways in which quantitative and qualitative elements might be combined into a single study. For example, in a large-scale survey, questionnaires might include both (a) rating scales or checklists that yield numerical data and (b) open-ended questions that yield verbal responses requiring qualitative analysis (see Chapter 6). 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 7).

Mixed-methods researchers tend to use certain notational conventions to describe how the quantitative and qualitative aspects of a study contribute to its overall design. In general, they use uppercase "QUAN" and "QUAL" to indicate that quantitative and qualitative methods both play major roles; they use lowercase "quan" or "qual" to indicate that one of these methods plays only a minor, supplemental role. For example, the notation "QUAN+QUAL" means that both quantitative and qualitative data contribute substantially and equally to a study or to one of its phases. The notation "QUAL+quan" indicates that a study or phase is predominantly qualitative but has a minor quantitative element; if you instead see "QUAL(quan)," a researcher might be communicating that a small quantitative element is in some way *embedded* within a larger qualitative element. The notation "QUAL→QUAN→QUAL" might be used to summarize a three-phase study in which a researcher first (a) conducts qualitative interviews to identify various aspects and nuances of people's attitudes regarding a topic, then (b) uses the results to create and administer a quantitative survey to gather data from a much larger sample, and finally (c) interviews a subset of the survey participants to get in-depth qualitative information that can enhance the researcher's ability to interpret the survey responses.

The specific ways in which a researcher might combine quantitative and qualitative methods are almost limitless, restricted only by the researcher's imagination and creativity, as well as by the nature of the research problem or question. To help you envision some of the possibilities, we describe five general types of mixed-methods designs: convergent, exploratory sequential, explanatory sequential, longitudinal mixed-methods, and multiphase iterative.[2]

Convergent Designs

In a **convergent design**, a researcher collects both quantitative and qualitative data within the same general time frame and with respect to the same general research problem or question. Some convergent designs take a QUAN+QUAL form, in that they give similar or equal weight to the two types of data. A researcher using such a design might strive for

[2]Four of these general design types reflect categories suggested by Creswell (2014, 2015). We have added longitudinal mixed-methods designs to the list because of its increasing popularity in recent years.

triangulation, hoping that analyses of both data sets lead to similar conclusions about the issue being studied. Alternatively, the researcher might collect quantitative and qualitative data to address different subproblems or subquestions couched with the major research problem or question, in which case the goal would be to reveal various, complementary facets of the phenomenon under investigation.

A study by Czechowski, Miranda, and Sylvestre (2016) provides a good example of a convergent, QUAN+QUAL design. Apparently being big fans of American musician and songwriter Bob Dylan, Czechowski and colleagues wanted to determine how Dylan's song lyrics might have changed over a 50-year time span (1962–2012) in ways that reflected and possibly also influenced social movements during that period. (Dylan's subsequent receipt of the Nobel Prize for Literature in 2016 lends credence to the notion that his work has indeed been influential.) But rather than conduct a strictly qualitative content analysis (as described in Chapter 6), the researchers conducted a *mixed-methods* content analysis of Dylan's lyrics. To obtain quantitative information, the researchers used a software package called Linguistic Inquiry and Word Count (LIWC) to count words in the lyrics that fell into various categories—for instance, words conveying positive emotions (e.g., *love, sweet*), negative emotions (e.g., *hate, cruel*), religion (e.g., *pray, sacred*), and social referents (e.g., *friend, family*). To gain qualitative information, they used a multistep coding process in which they developed and then refined general themes within the lyrics (we describe such a multistep process in Chapter 12). For instance, one general theme they found within the qualitative data was *oppression* (e.g., "How many years can some people exist before they're allowed to be free?"); another was *daily hardships* (e.g., "A lot of people don't have much food on their table"). Among the researchers' findings—many of which were supported by both quantitative and qualitative data—were (a) a small increase in both positive and negative emotions, (b) a more noticeable increase in religious content, and (c) a gradual (albeit inconsistent) decrease in social referents over the 50-year period.

In other convergent designs, one general approach dominates—perhaps a qualitative approach, but more often a quantitative one—with the other approach serving in a secondary, supplementary, and possibly supportive role to one or more of the quantitative variables. If the secondary form of data is simply *in addition to* the primary form of data, the design might be symbolized as either QUAN+qual or QUAL+quan. But if, instead, the secondary form of data in some way *clarifies the nature of* one or more variables within the primary data, the design is apt to be symbolized as either QUAN(qual) or QUAL(quan). For example, when planning a large-scale survey of people's attitudes or opinions about a controversial topic, a researcher might create a series of statements with which participants either "agree" or "disagree" at various points along a rating-scale continuum. Within this generally quantitative instrument, however, the researcher could embed several open-ended items in which participants explain their ratings. The qualitative data in this QUAN(qual) design could help the researcher make better sense of the numerical findings.

The quantitative component of a convergent design might even be an experiment with two or more treatment groups. For example, one of us authors looks way back to her own doctoral dissertation (J. O. Ellis, 1975). At the time, an ongoing debate in the experimental psychology literature concerned the kinds of cognitive processes adults use in their efforts to solve logical, *three-term series* problems such as the following:

John is taller than Bill.
Bill is taller than Fred.
Who is tallest?

One well-known theorist argued that people engage in strictly language-based strategies to solve these kinds of problems (e.g., verbally memorizing and then comparing parts of the problem), whereas another prominent theorist proposed that, instead, people use mental imagery (e.g., picturing a tall "John," a shorter "Bill," and then an even shorter "Fred").

In an effort to resolve the issue, Ellis hypothesized that at least three factors could affect the kinds of cognitive processes that people use to solve the problems: the ease with which problem elements could be mentally visualized, the modality of problem presentation (auditory versus visual), and the extent to which presentation format imposed significant demands on people's memories. She further hypothesized that because visual imagery can potentially help people synthesize multiple pieces of information into an integrated whole, an imagery-based problem-solving strategy might be more effective in solving especially memory-challenging problems.[3]

Using today's mixed-methods classifications, we can best classify Ellis's study as having a QUAN+qual design. In the dominant, quantitative aspect of her study, undergraduate students were asked to solve 40 three-term series problems that varied in terms of the three potentially influential factors she had identified; these factors were her independent variables. The experiment also included three dependent variables—percentage of correct problem solutions, response times (i.e., how long it took a participant to solve each problem), and participants' subjective ratings of problem difficulty on a 7-point scale.[4]

Immediately after solving and rating all of the problems, participants moved on to the shorter, qualitative part of the study, in which they responded in writing to the following open-ended essay task:

> We would like you to tell us how you think you solved these problems. "Look inside your head," and, with as much detail as possible, describe the process by which you think you solved these problems. If you used different methods for different problems, please describe each method and the problems you used it for. (J. O. Ellis, 1975, p. 51)

Ellis and a fellow graduate student independently classified participants' responses as describing (a) verbal processes, (b) visual imagery, (c) differential use of strategies depending on whether problems were easy or difficult to visualize, or (d) miscellaneous other strategies. A third graduate student weighed in whenever the first two raters disagreed on how best to classify a response, with the response ultimately being assigned the category on which at least two raters agreed. The raters didn't know which respondents had been in which treatment groups, nor, of course, were the participants aware of how their own problem-solving experiences might have been different from those of other participants. Hence, the entire study took the form of a *double-blind experiment* (see Chapter 7).

When Ellis statistically analyzed the quantitative data, she found that, as hypothesized, participants performed most poorly on problems that posed substantial demands on memory and yet were difficult to think about in terms of visual images. However, the numerical data alone did *not* specifically indicate that, yes, some participants were using visual imagery when trying to solve some or all of the problems. Here is where participants' introspective responses to the essay task were helpful: Many participants described strategies in which they "pictured" or mentally "arranged" the three items in the problems, and of those participants who said that they used different kinds of strategies for different kinds of problems, the vast majority reported that they used imagery-based strategies more frequently with easy-to-visualize problems than with hard-to-visualize ones.

Exploratory Sequential Designs

An **exploratory sequential design** typically encompasses two phases. In Phase 1, a researcher uses one or more qualitative methods to get a general sense of characteristics, phenomena, or other issues related to the topic of study. The qualitative data—perhaps

[3]Readers who would like more information about the theoretical bases for Ellis's hypotheses can find it in her subsequent journal article about the study (J. E. Ormrod, 1979).

[4]For readers who would like more information about the nature of the experimental treatments, we again refer you to Ellis's subsequent journal article (J. E. Ormrod, 1979).

from observations, interviews, or both—provide a basis for a more systematic, quantitative study in Phase 2. For example, qualitative observations of a phenomenon in a real-world setting might help a researcher develop hypotheses to be systematically tested in an experimental study. Alternatively, the results of a few unstructured interviews might help a researcher develop appropriate items for a questionnaire to be administered to a much larger sample—and possibly also to determine important subgroup differences to keep in mind when identifying the Phase 2 sample. For instance, interviews with a small group of people might help a researcher create a questionnaire about issues that divorced fathers face when their former wives have primary custody of their children (Winchester, 1999). Similarly, a few interviews might help a researcher construct an online survey regarding the roles that people's self-identified "bookmarks" or "favorites" play in people's subsequent searches for information on the Internet (Thakor, Borsuk, & Kalamas, 2004). Depending on how much the Phase 1 qualitative data and Phase 2 quantitative data play a role in a researcher's overall findings and conclusions, such studies might be classified as QUAL → QUAN, QUAL → quan, or qual → QUAN.

Explanatory Sequential Designs

Like an exploratory sequential design, an **explanatory sequential design** is usually a two-phase process, but in this case, the quantitative phase comes first. More specifically, Phase 1 involves collecting considerable numerical data, perhaps in an experiment, ex post facto study, or survey. Collecting qualitative data in a Phase 2 follow-up—for instance, asking a subsample of Phase 1 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. Often these studies place greater focus on the quantitative data; accordingly, they can usually best be described as QUAN → qual in nature.

A study by Jacobson, Hanson, and Zhou (2015) provides a good example of an explanatory sequential design. These three researchers' general area of concern was the effectiveness of graduate and postgraduate training for licensed psychologists who were regularly administering psychological tests in their clinical practices and often had to interpret the test results for clients. More specifically, the researchers wanted to learn whether practicing psychologists had been trained in specific strategies for providing feedback to clients about the results of, say, personality tests or neuropsychological assessments and how effective any such training had been. They posed four research questions, one of which will suffice for our purposes here: "How does graduate and postgraduate training prepare psychologists to provide TFB [i.e., test feedback]?" (p. 395).

Jacobson and colleagues (2015) used the membership directory of the Canadian Psychological Association to recruit a nationwide sample via e-mail invitations. Despite the low response rate (14.4%), the researchers obtained a sample of almost 400 practicing psychologists for their study. The first, quantitative phase of the study involved a 40-item online questionnaire that, among other things, requested information about (a) participants' use of various kinds of psychological tests, (b) the extent to which they provided test-result feedback to their clients, and (c) the existence and usefulness of prior graduate and postgraduate training in test-feedback strategies. The questionnaire also asked respondents whether they were willing to participate in a follow-up interview; if so, a second question asked for contact information.

Among the interview volunteers, six participants whose prior training had *inadequately* prepared them for giving test feedback were chosen for the second, qualitative phase of the study. These participants were divided into two groups: (a) three individuals who regularly provided test feedback to their clients despite insufficient training and (b) three individuals who rarely, inconsistently, or never provided such feedback. Depending on their group membership, they were asked such questions as "Could you tell me more about how you

learned to provide test feedback to clients?" and "Are there circumstances in which you would provide test feedback?" (p. 396), and their responses were audiotaped, transcribed, and later coded and classified for their content.

Although a majority of the questionnaire respondents indicated that their graduate or postgraduate training had included strategies for providing feedback about clients' test results, most respondents indicated that, to some degree, they taught *themselves* about how best to give feedback. The interview data supplemented these quantitative findings in several ways—for example, by shedding light on psychologists' self-instruction methods (e.g., a trial-and-error approach to identifying effective practices) and by pinning down reasons why psychologists didn't always provide feedback (e.g., existence of certain legal restrictions).

Longitudinal Mixed-Methods Designs

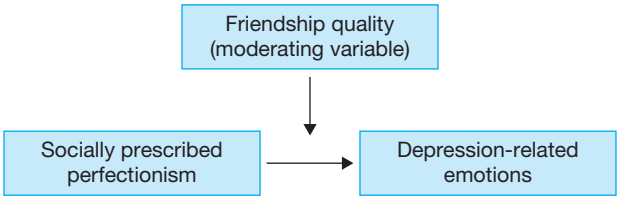
One of the quantitative descriptive designs introduced in Chapter 6 is the *longitudinal study,* in which a single group of participants is followed on two or more occasions over a lengthy time period, perhaps over the course of several months or years. Yet some longitudinal studies include qualitative as well as quantitative data, thus constituting a **longitudinal mixed-methods design**. As is true for a strictly quantitative longitudinal study, a mixed-methods version enables a researcher to determine the extent to which assessments conducted in any early phase correlate with assessments conducted at subsequent phases; thus, it also enables the researcher to find out how well variables assessed early on can *predict* the values of one or more variables assessed at a later time. Depending on whether the quantitative data and qualitative data are weighted equally or differently, a longitudinal mixed-methods study might be symbolized in a variety of ways. For example, one study might best be symbolized as QUAN+QUAL → QUAN+QUAL, another study might best be symbolized as QUAL(quan) → QUAL+QUAN, and still another—one with three phases instead of two—might best be symbolized as QUAN+QUAL → QUAN → QUAL(quan).

As an illustration of a longitudinal mixed-methods approach, we describe a QUAN+qual → QUAN+qual study by Mackinnon, Sherry, Pratt, and Smith (2014) that focused on perfectionism, friendship, and depression in college students. These researchers cited prior research supporting the idea that people who set very high, almost error-free standards for their own performance (i.e., perfectionists) are more likely to experience stress in demanding situations and, as a result, are more likely to suffer from chronic depression. They were especially interested in the possible effects of *socially prescribed perfectionism,* in which people believe that other individuals hold unrealistically high standards for their performance and judge them harshly when those standards aren't met. The researchers hypothesized that such perfectionists might benefit from the support of friends, who could help alleviate the social-emotional stress of failing to meet any excessively high standards and thus lessen any depressive symptoms. Friendships, then, could serve as a *moderating variable* that, through their impact on people's stress levels, influence any predictive (and possibly causal) relationship between socially prescribed perfectionism and depression. If you need a memory refresher on the nature of moderating variables, refer back to Chapter 2.

To address their hypothesis, Mackinnon and colleagues (2014) conducted a two-phase longitudinal study with first-semester undergraduate students, most of whom had graduated from high school only a few months before. Early in the first semester (Time 1), the research team used questionnaires to get quantitative estimates of participants' perfectionistic tendencies and emotional aspects of depression (e.g., sadness and/or irritability). In the same session, members of the team conducted and audiotaped one-on-one interviews in which students were asked to describe a "friendship-defining moment" with a same-sex friend during the preceding summer vacation. Students' responses during the interviews were subsequently coded on a 0–3 scale that indicated the degree to which their recollections revealed: (0) a conflictual relationship; (1) a task-oriented cooperative event in which interpersonal relationships weren't really described; (2) an emerging supportive friendship;

or (3) a very supportive, socially intimate friendship. Near the very end of the semester (Time 2), the team collected the same kinds of data, the only major difference being that students were asked to describe a friendship-defining moment that had occurred during the semester rather than during the preceding summer (i.e., before beginning college).

In some of their data analyses, Mackinnon and his colleagues (2014) discovered that a high level of socially prescribed perfectionism at Time 1 was associated with an increase in depression-related emotions from Time 1 to Time 2, but *only* for students who didn't appear to have close, supportive friendships at Time 1.[5] In other words, friendship quality did moderate the perfectionism–depression relationship, as follows:



Be careful that you don't interpret the arrows in our diagram as indicating clear-cut cause-and-effect relationships, however. Notice how we have just said that perfectionism *was associated with* an increase in depression; we didn't say that perfectionism specifically *led to* or *caused* an increase in depression. Also, we've said that the association occurred only when students didn't seem to have close friends; we never said that close friends specifically *helped* the students deal with any stressful events that might cause depression. Words such as those we have just italicized imply causal, some-things-affect-other-things relationships. Mackinnon and colleagues' correlational data were certainly consistent with the cause-and-effect relationships about which they hypothesized, but we remind you yet again that, by itself, *correlation doesn't necessarily indicate causation.* Furthermore, although stress might very well be involved in the perfectionism–friendship–depression relationships, the researchers never assessed this variable in their study.

Although longitudinal mixed-methods designs can be important tools in almost any researcher's toolbox, you should be aware of a potential drawback: In the process of collecting both quantitative and qualitative data at two or more different times, the collection of qualitative data at one time might influence the nature of quantitative data collected at a later time, and vice versa. For example, participants' introspective reflections on how they have mentally tackled certain kinds of logical problems in Phase 1 might give them insights into how best to solve similar problems in a subsequent Phase 2. Likewise, participants' self-perceived performance on an objectively scoreable questionnaire assessing, say, their religious beliefs or political ideologies might lead them to respond to a subsequent unstructured (qualitatively coded) interview in ways they might not otherwise have done. We authors aren't trying to discourage you from conducting a longitudinal mixed-methods study; we are only alerting you to an issue that may in some instances adversely affect the credibility of your findings.

Multiphase Iterative Designs

A **multiphase iterative design** includes three or more phases, with early ones providing foundational data on which later phases can build. The design is *iterative* in that the researcher moves back and forth among quantitative and qualitative methods, with each new body of data informing the conceptualization and implementation of subsequent phases. Whereas a

[5]The friendship-quality interview data collected at the end of the semester didn't appear to have a moderating effect, for reasons about which the researchers could only speculate.

longitudinal mixed-methods study can be planned in its entirety ahead of time and cleared with the appropriate institutional review board (IRB) in one fell swoop, a multiphase iterative design, by its continually evolving nature, typically requires a researcher—or more often, a team of researchers—to get IRB approval only one or two phases at a time.

Multiphase iterative designs have become increasingly common in program evaluation research—that is, in determining the efficacy of certain intervention programs. For example, imagine that a team of researchers wants to evaluate the effectiveness of a program created to reduce drug abuse in teenagers who live in a particular city. The team might take steps such as the following, with Steps 3a and 3b being two complementary parts of a single data-collection step:

1. Collect baseline quantitative data regarding the current prevalence of illegal drug use in the city's teenage population.
2. Implement an intervention program based on current theories and previous research studies.

3a. Collect subsequent quantitative data regarding illicit drug use in teenagers who have participated in the intervention; compare the data to those for a control group of nonparticipants.

3b. Conduct qualitative interviews of both program staff members and program participants to discover their views about the strengths and weaknesses of the initial intervention.

4. Make modifications to the intervention program in an effort to improve its outcomes.
5. Collect follow-up quantitative and qualitative data (e.g., prevalence of drug use by program participants; staff members' and participants' opinions about the revised program).
6. If deemed necessary, repeat Steps 4 and 5 through additional iterations.

Another arena in which multiphase iterative designs are common is *design-based research* (also known as *design experiments*), in which researchers in educational technology and other learning sciences apply existing knowledge and theories about human learning processes to create effective instructional programs or curricula. You will learn more about this type of research in the discussion of action research in Chapter 10.

When mixed-methods designs are lopsided in favor of one form of data over the other, they tend to lean in the quantitative direction (Hesse-Biber, 2010). Unless there's a compelling reason to do otherwise, we authors 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 experienced mixed-methods researcher 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).

CONCEPTUAL ANALYSIS EXERCISE Identifying Mixed-Methods Research Designs

The following five scenarios describe hypothetical mixed-methods research projects. For each one, identify the project's general design, choosing from one of these five options: convergent, exploratory sequential, explanatory sequential, longitudinal mixed-methods, and multiphase iterative. Also, depict the design using the symbolic notations "QUAN," "QUAL," "quan," and/or "qual." The answers are provided after the "For Further Reading" section at the end of the chapter.

1. A researcher conducts in-depth case studies of three average-intelligence middle school boys who have dyslexia; that is, these boys have reading abilities substantially lower than would be predicted from measures of their general cognitive ability levels. In the case studies, the researcher especially focuses on (a) the challenges and frustrations the boys face in their daily academic work and (b) strategies they have developed to help them cope with these challenges and frustrations. The researcher then uses the case study data to develop checklist questions for interviews to be conducted with a sample of 45 middle school students who have been identified as having dyslexia.
2. A researcher recruits 500 adult volunteers to complete an online questionnaire regarding their beliefs about several politically charged topics (e.g., climate change, legalization of marijuana, capital punishment). The questionnaire includes 40 statements (e.g., "Over the past several decades, the Earth's atmosphere has gradually been getting warmer") to which participants respond "strongly agree," "agree," "disagree," or "strongly disagree." Also, each of the last 10 rating-scale items is immediately followed by an open-ended question that asks participants to explain why they rated the preceding statement as they did. Based on the rating-scale responses, the researcher identifies some participants as being politically "conservative" and others as being politically "liberal"; the researcher then statistically contrasts the political leanings of people in various age groups and geographical regions. Also, the researcher and a second rater jointly code the answers to the 10 open-ended questions as reflecting the use of either (a) empirical evidence or (b) ideological thinking; the researcher then looks for possible differences in how "conservatives" and "liberals" reason about certain issues.
3. Based on currently existing theories and research evidence regarding effective study skills, a team of researchers creates a 6-month after-school program for low-achieving 10th-graders. In a quasi-experimental study, the team compares the pre-program and post-program GPAs of program participants with those of control-group students who do not attend the program. Later, at the end of the school year, the team conducts unstructured interviews of program participants to discover which aspects of the program they did and did not find helpful for improving their study skills. Over the summer, the team revises the program's curriculum in line with Year 1 participants' interview responses and then, the following year, conducts the program again with the school's incoming 10th-graders. Once again, pre-program and post-program GPAs are compared for program participants and nonparticipants.
4. At the request of the chief executive officer (CEO) of a large paper-products company, a researcher wants to determine why an 18-employee department within the company is especially dysfunctional. In a short paper–pencil questionnaire, the researcher asks each person in the department to identify (a) one to three individuals within the department whom the person likes best, (b) one to three individuals within the department whom the person especially *dis*likes, and (c) one to three individuals within the department for whom the person has no strong feelings. Based on the data collected, the researcher creates a sociogram similar to the one depicted in Figure 4.3 in Chapter 4; this sociogram enables the researcher to identify "leaders" and "isolates" within the group. The following week, the researcher interviews each department member, asking the individual simply to "Describe what you like and dislike about _____," with the blank being filled by the names of the apparent leaders and isolates within the department.
5. A researcher recruits a large group of 70-year-old women to participate in three-times-a-week exercise classes for a 2-year-period. The researcher randomly assigns each woman to one of two kinds of classes; one class focuses on strength-building exercises, and the other focuses on increasing aerobic fitness. At the beginning of the study and at regular 6-month intervals thereafter, the researcher obtains several quantitative measures of participants' overall health and fitness (e.g., blood pressure,

balance, and core strength). At each of these five times, the researcher also asks participants to write a one-page report describing their current activity levels and the kinds of illnesses and/or accidents that have afflicted them in the interim.

MyLab Education Self-Check 9.2

MyLab Education Application Exercise 9.2: Selecting a Mixed-Methods Research Design

PLANNING A MIXED-METHODS STUDY

As you have undoubtedly noticed, some mixed-methods designs are, in advance, planned 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 8). 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 many aspects of planning in earlier chapters (especially in Chapter 4), so here we alert you only to a few additional things that you should 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. In general, however, such questions should all be related to a single, overarching research problem or question.

As an example, we return to a study described in Chapter 7, 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 general research question was essentially this one:

> Do geographers study and remember maps differently than nongeographers do and, if so, in what ways?

Both quantitative data and qualitative data were potentially relevant to the problem: 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 and interpret new maps 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 landforms and human constructions are arranged on the Earth's surface. The answer to the second (qualitative) question was also *yes:* Geographers think about a new map differently than nongeographers do. In particular, geographers try to make sense of the arrangement of features on a map, whereas nongeographers are likely to engage in rote memorization as they study it. 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 several separate subproblems or research questions to guide you in your investigation. Furthermore, if you conduct a study with two or more phases, additional subproblems or questions may emerge as being important ones to address. 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 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 that might be supported or disconfirmed by quantitative data?

The final question in the preceding list 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 related literature. In other cases—for instance, in the exploratory sequential and multiphase iterative designs described previously—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, researchers using a mixed-methods design should conduct much of their literature review at the very beginning of the project. (Grounded theory studies are occasionally exceptions to this rule; see Chapter 8.) A review of related literature can help the researcher pin down appropriate questions and hypotheses, suggest possible methodological strategies, and reveal potentially helpful assessment instruments.

Especially in studies with two or more phases, additional visits to the library or its online databases may be useful midway through a research project. For example, in the first, qualitative phase of a two-phase exploratory sequential study, you might 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 sequential study, your initial quantitative findings might necessitate a search for new literature that will assist you as you conduct the subsequent, qualitative phase.

Choosing One or More Appropriate Samples

In virtually any kind of research—whether it be quantitative, qualitative, or mixed methods—the quality of the data obtained can be only as good as the quality of the sample(s) used. When the goal is to estimate what a relatively large population does, thinks, or feels, some form of *probability sampling* is called for—perhaps stratified random sampling or cluster sampling (see Chapter 6). Alternatively, a researcher might engage in *purposive sampling,*

choosing particular participants who can provide certain desired perspectives on a topic or issues. Following are examples of the many forms that purposive sampling might take in a mixed-methods study:

- Choosing participants who will represent diverse attitudes or opinions, including some who hold extreme views
- Intentionally choosing seemingly "average" or "typical" individuals
- Choosing participants on the basis of certain prespecified characteristics relevant to the research problem (e.g., parents who have children with significant physical disabilities, or psychologists who report that they have had inadequate training in strategies for giving clients feedback about test results; recall the study by Jacobson et al. [2015], described earlier)
- In a later phase of a multiphase study, choosing participants who can best help to either support or cast doubt on conclusions drawn in an earlier phase

In some studies with two or more phases, the samples used in various phases should be connected, or *linked,* in some meaningful way. For example, a study might involve one of these strategies:

- Using the same participants throughout a study (as is done in a longitudinal mixed-methods design)
- Choosing a large sample for a first, quantitative phase of a study and then selecting a subset of the sample for a subsequent, qualitative phase
- Choosing two samples that are related by biology, marriage, or some other connection relevant to the research problem (e.g., collecting data from parents in one phase and then collecting data from their children in a second phase)

Even when the focus of study is on nonhuman animal species, plants, or nonliving objects—rather than on people—a researcher should identify samples that can truly yield the information the researcher seeks. And regardless of whom or what is being studied, a researcher should provide a reasonable rationale for the sampling technique(s) being used.

Ensuring That the Study Will Have Reasonable Credibility

Whether conducting a qualitative, quantitative, or mixed-methods study, any researcher must choose or develop assessment techniques that will have adequate validity for their purposes (see the section "Validity of Assessment Strategies" in Chapter 4). Even more importantly, the researcher must take steps to ensure that the overall research effort is a credible one—that other individuals will judge its design and methods as being appropriate and its findings as being plausible and defensible.

Many strategies previously described in earlier chapters are potentially helpful in ensuring the credibility of a mixed-methods research study; we especially refer you back to Chapter 4 ("Planning Your Research Project") and the three chapters on specific research designs (Chapter 6, Chapter 7, and Chapter 8). But additional credibility issues arise 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 suggestions by Creswell (2010; Creswell & Plano Clark, 2011) and O'Cathain (2010), 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 else 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)?
- What personal or methodological biases might have differentially affected collection and interpretation of the quantitative data and of the qualitative data?

- Are quantitative sample sizes significantly larger than qualitative sample sizes? If so, what impact might the differing sample sizes have on the validity of conclusions?
- Can specific statements or artifacts from the qualitative element of the study be used to support or illustrate some of the quantitative results?
- Do the quantitative and qualitative data lead to the same or similar conclusions? If not, can discrepancies be reasonably resolved?

Analyzing and Interpreting the Data

The final activity in any mixed-methods study is to analyze and then interpret the collected data—a fact that should hardly be news at this point in the book. Unfortunately, there are no fixed procedures—no prescriptions, no recipes—for analyzing data in a mixed-methods study. Specific data analysis procedures depend on the design of the study, the kinds of data collected, and, of course, the nature of the research problem or question and any subproblems. Typically, one or more of the statistical analyses to be described in Chapter 11 are appropriate for quantitative data. Analyzing qualitative data often involves a multistep coding process using two or more raters; we provide more details about qualitative coding strategies in Chapter 12. But in addition, a researcher conducting a mixed-methods study must also *integrate* the quantitative and qualitative data in some meaningful and defensible way. We offer suggestions for doing so near the end of Chapter 12.

And remember, you can't stop with an integration of the quantitative and qualitative data; you must also *interpret* those data. Not only must you draw inferences from each form of data, but you must also draw inferences—meta-inferences, if you will—from the entire data set (Onwuegbuzie & Combs, 2010). Your final step must be to determine *what the data mean* relative to your research problem.

MyLab Education Self-Check 9.3

MyLab Education Application Exercise 9.3: Conducting Mixed-Methods Research

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 and confidentiality 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 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 exploratory sequential, explanatory sequential, and multiphase iterative designs. First, when the results of one phase in some way guide the implementation of a subsequent phase, the researcher may have to submit two or more complete proposals to the IRB or IACUC, one for each phase of the study. Alternatively, the researcher's initial proposal might describe Phase 1 in detail and give a general overview of the intended procedure for later phases, followed by updated, more specific (but perhaps briefer) proposals for the later phases. Either way, securing permission for the entire study will require more effort than would be the case for a one-shot research study or a repeated-measures longitudinal one. Don't let this fact discourage you: A project with two or more distinct phases might make a more substantial and enlightening contribution to your field than a shorter, more limited investigation could possibly do.

A second ethical issue arises in explanatory sequential designs and some multiphase iterative designs, especially if the researcher wants to use the results of Phase 1 quantitative data to choose a subsample to interview in a later phase (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 a subsequent phase. You must plan for this eventuality before you implement Phase 1 and must 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 e-mail address so that you can contact them (e.g., once again, recall the study of psychologists' test-feedback practices conducted by Jacobson et al. [2015]).

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 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 sequential, explanatory sequential, longitudinal mixed-methods, or multiphase iterative 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 credibility in the eyes of other scholars?

_____

_____

_____

_____ 4. Given your design, how much of your 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 might 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 or question? Alternatively, how might you revise your problem or question so that you can reasonably address it, given the time, skills, and resources you have?

MyLab Education Self-Check 9.4

A SAMPLE DISSERTATION

We conclude the chapter with an example of a mixed-methods study that used a two-phase, explanatory sequential design. Laura Lara (who shortly thereafter became 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 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 quantitative and qualitative results in interpreting her data.

DISSERTATION ANALYSIS 6

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?

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; Tashakkori & Teddlier, 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

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.

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.*

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.

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 4-year colleges and universities are quite accustomed to and comfortable with using e-mail for many school-related tasks. Before the turn of the century, however, such a procedure might have led to considerable bias in a researcher's sample.

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.

Notice the precautions that the author takes to ensure confidentiality and participants' right to privacy.

(In two subsequent paragraphs, the author describes how she statistically analyzed responses to the questionnaires.)

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.

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. 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.

Epoché *is an important strategy for maximizing a researcher's objectivity in a phenomenological study (see Chapter 8).*

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 (uses) 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.

The author uses what we have, in this book, called purposive sampling; *she also engages in a qualitative form of the* stratified sampling *we describe in Chapter 6. In particular, the author intentionally selects students of Mexican American heritage (both genders), including some with high GPAs and some with low 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 8 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 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).

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.

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

The author makes good use of technology to facilitate interview transcription and content analysis.

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 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.

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.

Note: 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.

APPLICATION EXERCISES

Similar exercises are available on Pearson MyLab Education.

1. Kevin is a researcher interested in the physical activity of elementary school students. He wants to describe the trends in teachers' practices in his local school district and to explore teachers' perceptions related to physical activity practices.
   - **i.** Why would a mixed-methods approach be appropriate for this study topic?
   - **ii.** What type(s) of data do you expect Kevin to collect in his mixed-methods study?
   - **iii.** Kevin's advisor is not familiar with mixed-methods research and suggests that he simply pick one method and stick to it. How might Kevin explain his choice of a mixed-method design to his advisor?
2. Debbie is interested in adjustment for college freshman students coming from small rural high schools to a large university. She plans to use mixed-methods research and is considering different possible purposes for her study. Help

her to identify the best design to use in each of the following scenarios.

i. Debbie wants to measure the adjustment scores of first-year students from rural high schools and to identify the overall trends. She then wants to explain the trends by describing the experiences of three students with high adjustment, three with medium adjustment, and three with low adjustment. What design she is thinking of using here, and how do you know?
ii. Upon further reflection, Debbie realizes that most of the research on college adjustment to date has involved students from urban settings, so she is unsure if existing instruments are appropriate for her rural participants. She decides to first qualitatively generate a theory that describes rural students' adjustment, then to quantitatively test that theory. What design is she thinking of using?
iii. As another option, Debbie decides that she would like to focus on getting an in-depth understanding of rural students' adjustment during the first 2 weeks of college. She decides to gather quantitative scores on various variables hypothesized to be related to adjustment and, at the same time, to ask students to tell her a story about their adjustment during the first 2 weeks. What design is she thinking of using?
iv. Finally, Debbie considers conducting an experimental test of an academic skills course for helping students adjust to college. In addition to assessing the experimental outcomes, she aims to explore the process of implementing the academic skills course. What design is she thinking of using?

3. Consider the following question: Do first-year physicians study and remember medications differently than their counterparts with 10 or more years of experience?

i. Explain how you would break this question into quantitative and qualitative subproblems.
ii. What kind(s) of data would you collect to address the quantitative and qualitative subproblems you identified?
iii. How could a researcher sufficiently address the issue of validity in a mixed-methods study of this question?

FOR FURTHER READING

Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research, 6*(1), 97–113.

Creswell, J. W. (2015). *A concise introduction to mixed methods research.* Los Angeles, CA: Sage.

Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). Los Angeles, CA: Sage.

Edmonds, W. A., & Kennedy, T. D. (2016). *An applied guide to research designs: Quantitative, qualitative, and mixed methods* (2nd ed.). Los Angeles, CA: Sage.

Greene, J. C. (2007). *Mixed methods in social inquiry.* San Francisco, CA: Jossey-Bass.

Hesse-Biber, S. N. (2010). *Mixed methods research: Merging theory with practice.* New York, NY: Guilford Press.

Hulme, D., & Toye, J. (2006). The case for cross-disciplinary social science research on poverty, inequality and well-being. *The Journal of Development Studies, 42,* 1085–1107.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14–26.

Morse, J. M., & Niehaus, L. (2009). *Mixed method design: Principles and procedures.* New York, NY: Taylor & Francis.

Onwuegbuzie, A. J., & Johnson, R. B. (2006). The validity issue in mixed research. *Research in the Schools, 13*(1), 48–63.

Plano Clark, V. L., & Ivankova, N. V. (2016). *Mixed methods research: A guide to the field.* Los Angeles, CA: Sage.

Plowright, D. (2011). *Using mixed methods: Frameworks for an integrated methodology.* London, England: Sage.

Tashakkori, A., & Teddlie, C. (Eds.). (2010). *Handbook of mixed methods in social and behavioral research* (2nd ed.). Los Angeles, CA: Sage.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences.* Los Angeles, CA: Sage.

Teddlie, C., & Yu, F. (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed-Methods Research, 1,* 77–100.

ANSWERS TO THE CONCEPTUAL ANALYSIS EXERCISE "Identifying Mixed-Methods Research Designs":

1. This is an *exploratory sequential design:* Qualitative data from the case studies (Phase 1) is used to create quantitative checklist questions in Phase 2. The design might be depicted as QUAL → QUAN.
2. This can best be categorized as a *convergent design:* The rating-scale responses yield quantitative data; the responses to the open-ended questions yield qualitative data

that help to explain why participants responded to particular rating-scale items as they did. Because qualitative data are collected for only 25% of the rating-scale items, are specific to those items, and appear to carry less weight than the quantitative data, this design might be symbolized as QUAN(qual).

3. This is a *multiphase iterative design* with three phases. In Phase 1, the program is administered, and quantitative data (GPAs) are collected for program participants and nonparticipants. In Phase 2, qualitative data are collected from unstructured interviews to gain insights about how the program might be improved. In Phase 3, the revised program is offered, and once again, quantitative data are collected to determine the program's effectiveness. The design might be depicted as QUAN → QUAL → QUAN.
4. This is an *explanatory sequential design:* The initial questionnaire yields quantitative data regarding apparently popular and unpopular individuals, and the subsequent interviews yield qualitative data that might shed light on why certain individuals are well-liked and disliked. The design might be depicted as QUAN → QUAL.
5. The design of this study is more complex than it might initially appear to be. It encompasses a two-group, random-assignment experiment with repeated measures (see Chapter 7). However, its incorporation of both quantitative and qualitative assessments conducted over a lengthy period of time leads us authors to suggest that it is best classified as a *longitudinal mixed-methods design.* The quantitative data (measures of blood pressure, balance, etc.) and qualitative data (self-reports about activity level and frequency of illnesses/accidents) are presumably being used to address different (albeit related) research questions and thus are presumably being weighed equally. Hence, the design's five phases might be depicted as QUAN + QUAL → QUAN + QUAL → QUAN + QUAL → QUAN + QUAL → QUAN + QUAL.

This page intentionally left blank

Chapter 10 Action Research and Participatory Designs

In some instances, we might be interested in conducting research that doesn't simply aim to identify, describe, explain, or understand an issue, problem, or phenomenon but that also aims to use research to inform practices, enact change, or transform conditions in some way. There are a number of research designs under the broader umbrella of action research.

Learning Outcomes

10.1 Identify situations in which action-oriented research might be a useful approach for investigating a research problem, issue, or topic.

10.2 Describe characteristics and purposes of a number of action-oriented research designs, including (a) teacher research, (b) design-based research, (c) participatory action research/youth participatory action research, and (d) community-based research.

10.3 Describe various strategies and approaches for designing and implementing action-oriented research studies, including (a) development of research questions and purposes, (b) preparation and implementation of data-collection activities, and (c) collaboration with various entities and stakeholders, such as school district leaders, community partners, and youth co-researchers.

10.4 Identify a number of approaches and venues for disseminating findings from action-oriented research studies that go beyond merely reporting findings and seek to enact change.

Thus far in the book, we have learned quite a bit about various quantitative and qualitative designs for conducting empirical research; we have also spent some time examining mixed-methods models that combine quantitative and qualitative methods and approaches. Regardless of the divergent methods employed in these models and designs, most of the paradigms and types of studies we have previously explored share one thing in common: They are designed and implemented by one or more academic researchers to learn more about an issue or problem. Although the researchers may articulate implications for practice and policy from their results or findings, the explicit *goal* of most research studies usually isn't to seek change in practices and policies through the research process. Quantitative studies might describe or define a problem using statistical data, identify a correlation between variables, or provide a causal explanation for a particular issue, but they usually don't then use these results to actually create or change programming. Qualitative researchers examine complex issues and phenomenon and aim to increase our understanding of a particular topic or process, but rarely do they use findings to initiate the implementation of a new policy, although others, such as policy groups, activists, and lobbyists, might use a research study to help create or shape policy.

Furthermore, in most of the aforementioned traditional research models, the researchers develop the research without consulting participants and sites beforehand; they may check in with certain organizations and institutions at the outset to gain approval and make sure they can actually conduct their study, but research questions would have already been crafted at this point. Some tweaks or modifications in a study's design might be made as a result of conversations or feedback from sites and participants, and qualitative designs, in particular, often take unexpected directions within the data-collection process, but the design and implementation of the study are largely the purview of the researcher. In most empirical studies, be they quantitative, qualitative, or mixed methods, there are two distinct entities: (a) the academic researcher or group of researchers who comprise the research team and (b) the researched, usually referred to as research participants. For the most part, these entities have divergent roles within the study: Researchers develop instruments and data-collection tools and then collect and analyze data, whereas the researched *provide* the data by completing surveys, answering researchers' questions in interviews, participating in focus groups, or serving as subjects for observations. After data collection is completed, the researchers generally analyze data with little or no subsequent involvement of the research participants[1] and write a research report, journal article, or book on their own; participants might be anonymously acknowledged but would not be involved as authors.

The subject of this chapter is a type of research that upends many of the conventions of traditional research just described. The terms **action research** and **participatory action research** refer to types of inquiry that aim to use research not just to describe or understand an issue or problem but also to take action and make concrete changes and transformations in conditions, resources, practices, and/or policy. These action-oriented models are also highly collaborative and challenge traditional divisions between "the researcher" and "the researched." In action and participatory action studies, researchers are often members of the group being researched; for instance, they might be teachers who are interested in better understanding their students' learning and developing more effective teaching practices, or they might be young people attending a community center who want to address violence in their community. People who might be only the subjects of investigation in traditional studies—students, teachers, administrators, community members—instead play a key role in action and participatory studies. More specifically, they are responsible for the design and development of a study and tasked with essential data-collection activities, such as conducting interviews and facilitating focus groups; they also play a major role in the analysis of data and in the write-up and reporting of findings.

The ways in which results and findings from action-oriented studies are disseminated to a larger audience also set them apart from more traditional forms of research. Whereas the primary vehicles for sharing findings of quantitative, qualitative, and mixed-methods empirical studies are articles in peer-reviewed journals, academic monographs, or book chapters, action and participatory action studies are apt to be "reported" in many innovative ways. For example, the research team might use findings to develop curricula, prepare grants for funding new programs, and/or organize a community forum devoted to a particular issue or concern (more about such dissemination strategies later in the chapter).

MyLab Education Self-Check 10.1

[1]As discussed in Chapter 8, many qualitative studies employ member checking to give participants the opportunity to review transcripts for accuracy or offer feedback on findings. However, those approaches don't involve participants as co-researchers in the *analysis of data,* as is the case for the types of designs described in this chapter.

ACTION-ORIENTED MODELS AND DESIGNS

There are a variety of designs and models under the broader umbrella of action-oriented research. This chapter will focus on a few especially prominent ones:

- Action research conducted by teachers (a.k.a., *teacher research*)
- Design-based research (DBR)
- Participatory action research (PAR) and youth participatory action research (YPAR)
- Community-based research (CBR)

Although each of these models possesses distinct characteristics, they have two features in common: (a) a focus on using research to make changes in practice and/or policy and (b) a flexible, iterative process that often involves collaboration among academic researchers and practitioners.

Action Research/Teacher Research

Action research is probably most commonly conducted by practicing teachers and others interested in educational outcomes, hence its alternative name *teacher research.* More specifically, as a specific category of research, it is a form of action-oriented research wherein teachers and other educational stakeholders systematically investigate an issue or topic of interest to them, with the intent of informing their own practices and enhancing student learning and outcomes. Rather than designs in which research is done *on* or *about* teachers, in action research, teachers conduct research *for themselves*, often investigating their own practices. Mills (2018) recommends a four-step process in action research, which we can summarize as follows:

1. Determine a research question that, when answered, can yield concrete strategies in the here and now.
2. Collect data that might help to answer that question.
3. Analyze and interpret the data relative to the research question.
4. Develop a plan for action.

The first three steps should look familiar, as you have seen them often addressed in previous chapters. Step 4 is largely a new one: A key component of action research is finishing up with a *concrete plan* regarding how to solve the problem at hand.

Action research is rooted in Kurt Lewin's (1948) seminal theory and research in applied/action research and experiential learning, Lawrence Stenhouse's (1981) work on teachers' systematic practice and reflective inquiry, and the writings of John Dewey (1904), who exhorted educators to not just consume educational research but also serve as creators of knowledge about teaching. Teacher research acknowledges teachers' expertise, knowledge, and capacity for theorizing about their own practices and contexts; in contrast to research designs where teachers are merely receptors of knowledge, in action and teacher research, teachers are viewed as "architects of study and generators of knowledge" (Cochran-Smith, & Lytle, 1993, p. 2).

Action and teacher research have been developed and gained popularity largely in response to views of most academic research in education as irrelevant and not useful or applicable to many practitioners and teachers. Thus, action research focuses on research that can be applied and is useful in real-life contexts. Research questions for action and teacher research emerge from relevant problems and salient issues within the setting. Data-collection instruments and approaches used within action research studies are similar to what might be collected in traditional quantitative and qualitative studies; interventions, test scores, surveys, observation protocols, interviews, focus groups, and/or textual and visual documents (e.g., school materials

or student work) might be employed to shed light on a particular issue or topic. Teachers' own reflective journals and narrative writings might also be used as data sources and can provide insights into teachers' reflective practices and unique ways of thinking about what good teaching involves (Cochran-Smith & Lytle, 1993; Nieto, 2003). Data collection and analysis within action research studies are emergent and often iterative processes, meaning that data-collection activities are informed by an ongoing process of analysis. As working hypotheses are identified, teacher researchers may decide to collect new data to provide additional evidence and insights. Although action research might be undertaken by a sole researcher/practitioner, many action and teacher research studies are collaborative, possibly involving a group of teachers, practitioners, and administrators intent on designing and implementing action research to better understand a program, policy, reform efforts, or student learning and achievement at their site.

Reports of results and findings from action research studies are also very practitioner-oriented. Most, but not all, teacher research articles tend to be shorter than those aimed at an academic research audience, and they usually also include a substantial section on recommendations for practice, maybe even specific teaching tips or curricular suggestions. Many professional organizations publish journals that are specifically geared toward teachers and practitioners. For example, the International Literacy Association is known for its balance of high-quality journals, some of which are written primarily for academic scholars and others of which are written primarily for teachers and other practitioners. For example, *Reading Research Quarterly* highlights scholarly academic research and theoretical perspectives, *The Reading Teacher* focuses on teachers' needs and concerns and includes concrete suggestions for everyday classroom practices, and *The Journal of Adolescent and Adult Literacy* (also geared toward teachers) features articles with "classroom-tested ideas" and reflective pieces on current trends and issues. Many professional organizations dedicated to educational research, such as the American Educational Research Association (AERA), also have special units or groups on action or teacher research.

Teacher research can explore a variety of topics and issues, from examining programs that promote preschoolers' transition to kindergarten through social networking and child-led literacy activities (Tolentino & Lawson, 2017) to investigating the use of critical pedagogy among vulnerable students and the development of educators' critical awareness of their students' individual needs (Jones, 2017). Some action researchers have sought to better understand their adolescent students' reading and writing and provide other teachers with instructional frameworks and effective teaching strategies (Rief, 2014), whereas still others have looked at the intellectual strengths of and challenges faced by the low-income immigrant students they teach (Ballenger, 2009).

Although teacher/action research is a form of inquiry that can yield rich insights into topics, issues, and problems in schools and classrooms, it is not without its challenges. First and foremost, as is true for much of qualitative research, dealing with authentic, real-life contexts is often very messy. Typically, many conditions and variables cannot be controlled (essentially becoming confounding variables), and unexpected events might interfere with the execution of the study or muddy its results. In teacher research, this messiness is often heightened due to the inherent unpredictability of many schools and classrooms: Children whom a teacher researcher may have chosen as the focus of observations may end up transferring to another school, and budget cuts might mean that an innovative program that a teacher researcher had chosen to examine is no longer being offered. Thus, individuals engaging in action-oriented research studies need to be extremely flexible in the face of such unforeseen circumstances.

Another challenge associated with teacher research involves the unavoidable fact that the people conducting the research are often also those responsible for delivering instruction and/or implementing activities, therefore posing difficulties in terms of negotiating and managing roles as practitioner and researcher. Teachers who are conducting research related to their own classrooms don't have the luxury of neutrally observing activities from a perch

in the corner, and they may not have much time during the school day to write field notes, conduct interviews, or lead focus groups.

To address challenges such as these, action researchers and teacher researchers need to establish certain procedures and strategies at the outset of a project—at the *planning* stage—as means of maximizing the odds that they can carry out research activities efficiently and in a timely manner. Reduced teaching loads and teacher interns who assist with instruction are two possible options through which teachers can gain the time they need to plan a solid design, identify or create appropriate assessment procedures, collect needed data, and defensibly analyze them—all essential elements of a successful study.

Design-Based Research (DBR)

Emerging from the field of the educational technology and other learning sciences, **design-based research (DBR)** aims to produce theories and knowledge about learning and teaching that are grounded in real-life contexts, and it involves manipulating these contexts in particular ways and conducting and refining experiments and interventions so that researchers can make evidence-based claims about learning (Barab & Squire, 2004; A. Brown, 1992; Collins, 1992).

As an example, imagine that a researcher wants to develop instructional software that can help middle school students learn how to *control for confounding variables* when conducting simple physics experiments—say, when determining what one or more variables (e.g., length, weight) influence the rate at which a pendulum swings back and forth. In developing a preliminary version of the software, the researcher draws on existing knowledge regarding educational strategies that have been empirically shown to enhance children's scientific reasoning skills. The researcher has a small sample of middle schoolers use the software to swing a virtual "pendulum" while controlling for such factors as the length of the pendulum, the weight of the object at its bottom, and the force with which the pendulum is initially set in motion. The researcher then quantitatively and/or qualitatively assesses the students' subsequent ability to control for confounding variables in other computer-based "experiments"—for instance, experiments in which students must determine (a) what environmental conditions influence how fast sunflowers grow and (b) what factors affect an automobile's fuel efficiency. After paying particular attention to parts of the software that seemed to have worked well and to those that have not, the researcher modifies the program, has a second sample of students use it, tweaks the program again, and so on, until—ideally—a truly effective software program emerges.

As illustrated in the preceding example, DBR is an iterative process that entails conducting and then continually revising experiments based on an ongoing analysis and evaluation of evidence collected during interventions. It's a pragmatic approach to conducting research that can generate evidence-based, locally usable knowledge in complex environmental settings while simultaneously creating a "bridge" between that knowledge and evidence that has previously been acquired from more basic laboratory studies (Sandoval & Bell, 2004).

Most design-based research is multidisciplinary, in that it draws perspectives and methodologies from a variety of fields in education and the social sciences. The methodological approaches used in DBR are typically a blending of positivist and interpretivist orientations; you might want to return to Chapter 1 for a refresher on these philosophical orientations. For the most part, DBR tends to be largely quantitative and experimental in nature, although with a greater emphasis on understanding certain phenomena or processes within complex, real-world contexts than is true in more traditional, laboratory-based research (Barab & Squire, 2004; Design-Based Research Collective, 2003). Most DBR studies involve collaboration with teachers or other practitioners, and participants in DBR studies are viewed as co-participants who play integral roles in the design and analysis of the research, as these individuals are apt to have essential knowledge and understandings regarding the problem at hand and the setting in which the research is to be conducted.

Design-based research studies have examined a wide range of topics, such as the possible effectiveness of an inquiry-learning approach in biology (Reiser et al., 2001) and second-language learning in an online environment (Pardo-Ballester & Rodriguez, 2009). Much DBR has focused on traditional educational contexts (e.g., elementary, secondary, and university classrooms). However, there have also been DBR studies conducted in community and workplace/vocational settings, such as research on how fisherwomen in Indonesia learn the ins and outs of the fishing business (YuliEfiyante, Nirbito, Wahjoedi, & Sunaryanto, 2016) and work that has investigated how 3-D gaming and learning spaces can be used to support vocational learning (Hämäläinen & Oksanen, 2012).

A primary goal of DBR studies is to produce new theories and articulate practices that can have a bearing on learning and teaching in actual educational settings. However, DBR researchers often face two challenges when striving to achieve these goals: replicating their studies and making generalizations from their findings to other settings. Although replication and generalization aren't always concerns for other action-oriented research designs, DBR is specifically intended to generate theoretical knowledge that can be useful in other contexts, making replication in other contexts and generalization of findings to other settings highly desirable. But remember, because DBR focuses on better understanding the complex conditions and factors involved in teaching and learning processes, typically through a series of interventions, replication (and hence generalization) can be tricky, if not impossible.

Other challenges in DBR designs should be noted as well. In messy, real-life settings in which multiple variables and factors are at play in shaping and impacting learning processes, determining actual cause-and-effect relationships within the data can be extremely difficult. Also, because of the iterative nature of data collection—involving multiple stages and interventions, often carried out in a single setting over a prolonged time period—DBR studies can be very time consuming and require a sustained and consistent partnership between researchers and teachers.

Participatory Action Research (PAR) and Youth Participatory Action Research (YPAR)

Participatory action research (PAR) is an action-oriented research approach rooted in the theory of Paulo Freire (1970), a Brazilian educator who developed the notion of problem-posing education, wherein the generative themes of students' lives were used as the basis for instruction, and education was viewed as a liberatory and transformative act. PAR researchers disagree with positivist views of research as being objective, impartial, and value-neutral. Instead, they conduct research projects that involve reciprocal collaboration with individuals typically viewed as the "subjects" of the research; this collaboration occurs at all stages of the research, from the initial design of the project to final dissemination of its findings (Lather, 1986). PAR studies are designed to engage participants in investigating issues of importance and relevance to their lives and, furthermore, to challenge status-quo explanations of complex issues and problems. PAR researchers often make a particular effort to engage individuals who have a history of being marginalized in a particular social group or in society at large. They pay special attention to characteristics such as race, gender, ethnicity, class, age, sexual orientation, and other identity categories, with a focus on how these characteristics differentially affect people's access to resources and experiences relative to the phenomenon or issue under study.

During the course of a PAR study, members of the research group share leadership responsibilities and authority; also, as they continually engage in dialogue with one another, they share their unique knowledge, skills, and experiences. Another key concept underlying PAR studies is *praxis,* a term that refers to practical knowledge and integrates reflection and action, with the end goal of making transformative changes in social conditions. Within PAR studies, researchers usually align themselves with particular social groups, employing

research to address inequities and make substantial changes in conditions and resources. The products of PAR studies are meant to enact change and might take the form of a community forum devoted to the issue, a social media campaign addressing a particular concern, or a grant proposal to create or expand programming.

Youth participatory action research (YPAR) is a type of PAR that involves young people as co-researchers. The pedagogical aspects of PAR make this approach especially useful as an instructional tool in classrooms or community settings, where it can facilitate youth investigation of meaningful issues and topics. YPAR projects are often led by an academic researcher working with a group of students to design and implement a study and then identify and carry out a plan for disseminating findings; often, a goal of these projects is to utilize the research process to help youth acquire and develop important academic skills (Irizarry, 2011). One notable YPAR project examined the educational experiences of Latino/a youth, using these experiences to challenge negative stereotypes about the young people's possibilities for success (Irizarry, 2011). Another YPAR study investigated stigmatizing gender views and portrayals of women in a particular community and designed an alternative advertisement campaign to promote more positive images and perspectives (Cahill, Rios-Moore, & Threatts, 2008).

Participation action research (PAR) and YPAR studies often begin with a general topic of interest—say, the quality of the environment or food insecurity, which refers to lack of access to food resources—which members of the research team refine and develop into a research project through discussion and dialogue. If an academic researcher from a university is undertaking a PAR study in concert with members of a community-based organization, research questions are developed collaboratively rather than determined beforehand. Genuine PAR and YPAR studies aren't merely opportunities for stakeholders and insiders to "help out" with an academic researcher's project but are authentic collaborative endeavors in which all participants are valued for the expertise they bring to the project and wherein knowledge and perspectives are co-constructed. Data collection in PAR/YPAR studies entails bidirectional and reciprocal processes of training and teaching. For example, workshops might be offered to prepare participants to conduct interviews or focus groups, community members could provide panels to educate others on pertinent issues and concerns, or young people might train adult researchers in the use of youth-oriented social media. Data-collection approaches used in a PAR or YPAR study often include qualitative approaches, such as in-depth individual interviews, focus groups, participant observation, and collection of textual and visual documents. Members of the research team might also design and administer surveys or perform a content analysis of program materials. The research group might complete a community mapping activity as a way of gaining insight into the distribution and location of community resources.

Dissemination of the findings of PAR/YPAR studies is usually action-oriented and geared toward making substantive changes related to the issue of concern. These reporting and dissemination efforts should also be collaborative. Research team members might coauthor an article or book chapter; for example, Irizarry (2011) co-wrote an entire award-winning book with Latino/a youth who served as co-researchers in a YPAR study on Latino/a education. Likewise, in addition to writing a book on her PAR dissertation work on the role of a group of Latina mothers in the creation and development of a small school, Dyrness (2011) and her research team—called *Las Madres Unidas*/Mothers United—produced a video and organized a school forum to address issues and problems of parent voice and involvement at the school, resulting in the establishment of a parent center at the school. And in studies by Morell (2004, 2008), young people participated in specially designed summer seminars and conducted research projects on topics of importance to them—for example, the use of hip-hop to transform secondary literacy curricula, teachers' attitudes concerning students' home languages, or students' resistance to conditions and practices (e.g., inferior instruction or unfair behavior management strategies) in urban schools. Seminar participants produced

a variety of reports and materials, including a survey instrument for secondary students, research papers, documentary videos, and public presentations. These were disseminated through a number of social media and public contexts and largely aimed at parents, teachers, and public officials.

Although PAR bears significant potential for using research to bring about transformations and change, it is not without challenges. First of all, it can be difficult for academic researchers who have historically had sole control of their research projects to, instead, cede expertise and authority to community members, youth, and other stakeholders. When academic researchers engage in PAR/YPAR research, it is of utmost importance that they ensure that *all* co-researchers have a voice and that any inequities and uneven power relationships existing in larger society—for example, as related to age, class, educational level, or racial differences—are not replicated within the study (e.g., see Nygreen, 2013). Research teams need to establish clear roles for various group members and maintain constant vigilance to ensure that multiple perspectives and voices are elicited, perhaps by including ongoing assessments of the project's progress as means of ensuring regular feedback from group members. Another issue that can pose challenges for PAR and YPAR researchers involves the time and resources needed for such projects; planning and finding time for group meetings can be difficult, as members may have busy schedules and multiple demands on their time. For academic researchers who undertake a PAR or YPAR project, designing and planning the project requires a significant outlay of time long before data collection begins. For example, if a YPAR project is offered as part of a class at a public school or other educational institution, there may be institutional constraints and regulations (e.g., required curricular content or rules against taking students off-site) that pose challenges to successfully implementing the project. Above all, PAR and YPAR researchers need to be flexible and able to adapt a study in order to address unforeseen events and problems.

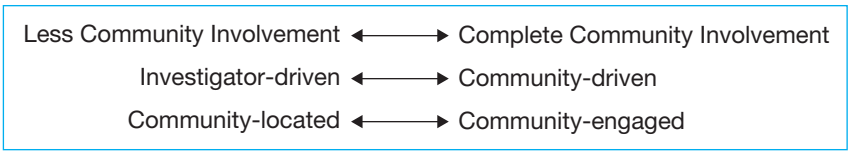
Furthermore, because many elements of the project, such as research questions and interview guides, are developed *in situ*—that is, on site—and after the project has begun, not all elements can be specified in advance for the institutional review boards (IRBs) that must grant approval for the project. As PAR becomes a more widely used model, IRBs are becoming more informed about the design of these studies; many are willing to approve the projects and allow for submission of amendments for research questions and instruments, but it may fall on an academic researcher to provide background on and explanation of the model. Many university-based PAR and YPAR researchers also confront questions and doubts from colleagues and tenure committees regarding the rigor of this type of research, and because authority and expertise are shared with participants and co-researchers, the university researchers may have to defend their work as "real research" to those who view it merely as community service or outreach (Fine, 2008).

Community-Based Research (CBR) Models

Another group of action-oriented designs consists of those that are part of **community-based research (CBR)** models and approaches. As the name suggests, these models are usually bound in a particular community setting and involve developing partnerships with community organizations and groups. Some research projects might investigate health issues and disparities in a particular community, working with local public health organizations and hospitals to develop action plans and health programming that address the cultural and social needs of community members (A. B. Becker, Christoffel, Morales, Lopez, & Longjohn, 2011). Other projects could explore the role of mentorships in the lives of young mothers living in a specific community, collaborating with a local alternative school to design a program that builds on the experiences and knowledge of mothers attending the school (L. R. Johnson, 2017).

Some of the community-based research models described in this chapter emerged out of university initiatives to build partnerships with communities and provide engaged learning

FIGURE 10.1 The Three Closely Related Continua Along Which Various Community-Based Research Projects Lie



experiences for students. Many of the models are premised on the beliefs that (a) communities are storehouses of resources and knowledge and (b) community members possess significant expertise and skills that can provide insights on key community issues and problems. These may seem like fairly obvious points, but for many years, communities—particularly low-income ones—were regarded only as sites and subjects of research to be conducted by external academic researchers. Rarely were residents and key stakeholders consulted as to what sort of research might be helpful, nor were such individuals engaged as authentic research partners. In some cases, communities have been exploited or damaged as a result of research because findings reinforced negative stereotypes and pejorative views of certain groups.

Researchers at Virginia Commonwealth University's Center for Clinical and Translational Research have identified a community-engaged research continuum, from (a) less community involvement to complete community involvement, (b) investigator driven to community driven, and (c) research that simply takes place in a community setting to research that fully engages community members in research (see Hacker, 2013). This continuum—which actually consists of three closely correlated continua—is illustrated in Figure 10.1.

Many different models exist under the broader umbrella of *community-based research.* Following are a few of the more prominent ones.

Course-Based Action Research

Course-based action research (CBAR) is a type of research that integrates community-based research and participatory action research as a way of bringing together undergraduate students, faculty, and local community partners to undertake collaborative research projects that yield benefits for community partners. Within their coursework, students are introduced to various community organizations and local politics, and they conduct fieldwork that physically, cognitively, socially, and politically immerses them in community settings and problems; in the process, they become better able to understand and "critically evaluate their sociopolitical surroundings" (Hofman & Rosing, 2007, p. viii). CBAR is rooted in a view of *praxis* wherein theory and practice are inextricably intertwined within one or more social, economic, and political contexts. Two key goals are to (a) foster a commitment to social justice on the part of both students and faculty and (b) strengthen connections between the university and the community.

In one situation, a CBAR study was conducted to explore the role of local corner stores in a low-income African American community in Chicago and the relationship between poverty and community conditions related to access to food and nutrition (Rosing, 2007); students used their projects to support the work of a local organization addressing food issues and working in the area of urban agriculture. In another CBAR project, university students in an urban geography course examined physical, social, and economic development that (thanks in part to dramatic increases in home values and rents) was displacing long-time residents in a historically Mexican neighborhood in Chicago (Curran, Hague, & Gill, 2007). More specifically, the students engaged in community mapping activities (e.g., surveying properties in certain districts within the area to assess building conditions, collecting information on property taxes and values) to gather and organize data that could be useful to a local community-based organization that was trying to preserve the neighborhood's Mexican heritage and affordability.[2]

[2]Readers in such disciplines as geography and urban development might recognize that the issue of concern in Curren et al.'s (2007) CBAR study was the *gentrification* of the neighborhood. In particular, gentrification is a process of renovating a neighborhood and in other ways changing its character to enhance its appeal to middle-income and upper-middle-income individuals and families.

Community-Based Participatory Research

An approach that is largely used in the field of community health research, **community-based participatory research (CBPR)** employs a collaborative, partnership approach to conducting research. Communities are acknowledged for their strengths and resources, which are used as the foundation for research projects. CBPR is motivated by the acknowledgment that evidence-based innovations and practices are often slow to be integrated into community practice, further illustrating the gap between academic researchers and community-based practitioners (Hacker, 2013). CBPR researchers believe that studies should collaboratively involve community members and practitioners in the design and implementation of projects. Such an approach can engender a sense of *shared ownership* that should yield relevant and useful findings that are likely to be integrated into community health programs and practices and lead to improvements in health outcomes. Another goal of CBPR is to empower the research partners and enhance the capability and willingness among community-based organizations to address local problems and issues. As you can reasonably guess, effective CBPR programs require a long-term commitment from their partners.

As examples of CBPR in action, one project examined barriers to immigrant health in a particular community, working with local immigrant advocacy groups to make changes in policy that addressed these obstacles (Hacker et al., 2011), and another project aimed to confront and ameliorate high obesity and diabetes rates in a Puerto Rican community in Chicago, working with various community groups and businesses to design and implement holistic interventions and programming (A. B. Becker et al., 2011).

Community-Based Qualitative Research

A model that is often called **community-based qualitative research (CBQR)** builds on the tradition of PAR, utilizing qualitative methodology to investigate and address issues and problems in particular community settings (L. R. Johnson, 2017). CBQR projects are collaborative, in that leadership in the project is shared with community partners and groups, and projects are developed, implemented, and disseminated collaboratively. CBQR projects take a critical stance on local issues, placing them in their social, cultural, economic, political, and/or historical contexts and connecting community problems to broader inequities and disparities. Finally, CBQR intends to be a transformative process for academic researchers and community co-researchers, in that it can expand group members' notions of community and research and possibly shift their thinking on a number of issues and topics. CBQR projects also aim to make substantial improvements and changes in community conditions and resources. As examples, CBQR studies have examined young people's civic engagement (L. R. Johnson, 2017) and the role of mentorship among teen mothers (L. R. Johnson, Sokacich, Rutledge, & Rosario-Ramos, 2016), in both cases using the findings to create new programming.

Regardless of the specific design used, many community-based research studies begin with a set of **orienting experiences** that (a) help academic researchers learn about the community and pertinent issues and (b) provide opportunities for dialogue between academic researchers and community partners (L. R. Johnson, 2017). These orienting experiences might include community tours or panels comprised of community leaders and stakeholders with experiences and expertise related to the issue or topic. Student co-researchers might be assigned readings and articles that provide insights about the community; alternatively, students might complete reflective writing assignments to discuss their own backgrounds and identify their existing assumptions and biases related to the community and research topic. Before visiting the community and meeting with community partners, academic researchers should learn as much as they can about the community's history, demographics, cultural resources, and key concerns and challenges; census data and organizational websites are often great sources of information about a community.

In conducting any community-based research project, it's essential that academic researchers continually keep in mind the importance of community members' knowledge and perspectives, which are essential to the credibility, integrity, and usefulness of the project. Yet researchers who are about to undertake a community-based research study can't just "drop in" and expect community members to voluntarily and immediately educate them about the community's characteristics, cultural foundations, and needs. Rather, researchers must be planful in their efforts not only to recruit community members' participation but also to build a solid, shared body of community knowledge and shared understandings.

The challenges faced by community-based researchers are similar to those faced by PAR researchers. As they work with community partners, academic researchers need to be willing to cede control and share authority and leadership, and they must remain mindful of existing unequal power relationships, both between the academic researchers and community members and within the larger community.

Also, when working with community-based organizations, it's important to remember that a single community-based organization doesn't necessarily represent "the community." Community settings are diverse and dynamic spaces, rather than homogeneous groups that all share the same experiences and viewpoints, and community-based researchers should attend to and elicit people's divergent and multiple experiences and perspectives related to the topic or issue under investigation. When studying controversial issues and topics, community-based researchers can find themselves in the middle of divisive struggles and contentious debates, and they should be aware of power hierarchies and existing tensions. In some cases, however, aligning oneself with a particular group within a community might be intentional as a way of advancing a particular cause. For example, a researcher might want to learn more about and help youth who are working in the area of violence in the community, or the researcher might work on behalf of low-income residents who have been displaced because of urban redevelopment.

As is true for other action-oriented models, community-based projects require a significant time commitment on the part of research group members in order to (a) build rapport and reciprocal relationships with community members, (b) design and implement the study, and (c) develop and carry out a plan for dissemination. It's essential that academic researchers who work with community partners follow through with all aspects of projects and make sure that the research findings are used to support the work of the community, rather than just lying dormant and gathering dust somewhere on the researchers' shelves. Academic researchers must also set and communicate realistic goals and outcomes of a study. If they make promises they either don't or can't keep, the entire research experience is likely to make community members feel exploited and misused and hence either create or intensify tensions and divisions between a university and the community it is intended to serve.

An additional consideration is the fact that, similar to other action-oriented designs such as PAR, community-based models can struggle to earn legitimacy as research and scholarship. Researchers engaged in this sort of activity may have to educate tenure and promotion committees that view community-based work as simple community outreach and service rather than as rigorous research.

CONCEPTUAL ANALYSIS EXERCISE Choosing an Action-Oriented Design

Following are brief summaries of five potential research projects. Identify the action-oriented or participatory design that is probably most appropriate for each project. The answers appear after the "For Further Reading" section at the end of the chapter.

1. An academic researcher has recently learned about an environmental justice project in her local community. She wants her undergraduate students to join the project for a semester, ideally to conduct a data-gathering study that can help identify effective strategies for getting young people more interested and engaged in envi-

ronmental justice efforts. Her ultimate goal is to create a curriculum that can be used in an after-school program for high school students.

2. The members of a group of researchers and providers in public health are troubled by a recent report about the lack of prenatal care for adolescent mothers in a specific region. They would like to learn more about the roots of this problem and address it with increased educational programming and support services.
3. A fifth-grade teacher is struggling with the issue of how best to implement a particular reading program in his classroom of students with widely varying reading abilities. He would like to learn more about the reading skills, needs, and interests of his students so that he can better help them all become successful readers.
4. A university professor wants to help students enrolled in her master's level sociology course learn more about action-oriented approaches in research. She has already developed a good relationship with a local community-based organization that is working to address the needs of recent immigrants to the area.
5. A team of researchers is beginning a project investigating the use of technology-enhanced learning environments in high school classrooms. The team is interested in conducting a study that can both (a) contribute to theoretical understandings of adolescent learning processes and (b) guide the design and implementation of a new instructional initiative at a local secondary school.

PRACTICAL APPLICATION Deciding Whether to Use Action Research and/or a Participatory Design

Action research and participatory designs comprise a unique category of research and are certainly not for everyone. For researchers who are committed to conducting activist research and/or hope to use research findings to enact changes, there is no question that action-oriented inquiry is quite suitable for those purposes. Similarly, practicing teachers are well advised to employ an action research design when they're interested in learning more about their students and specific classroom strategies in order to improve their students' academic and social success at school (and we authors hope that teachers often are interested in doing these things!). However, some researchers may be sitting on the fence regarding the use of an action or participatory design and wondering if the advantages outweigh the drawbacks. The questions in the following checklist can help you decide if an action or participatory design is appropriate for your study. It can also increase your chances of staying true to the key elements of action-oriented designs.

CHECKLIST

Pinning Down the Logistics and Feasibility of Action Research, a Participatory Design, or a Combination of the Two

_____ 1. What do you hope to accomplish with your study? Is your research issue or question one that is best answered by conducting research in a real-life, authentic setting?

_____

_____

_____ 2. In what ways will your study involve practitioners, stakeholders, and/or community members? Is it possible to involve them in the way(s) you hope to?

_____

_____

_____ 3. How will collaboration between academic researchers and community partners be integrated into the following stages of the research process?

- Design (e.g., writing research questions)

  _____

  _____

- Data collection (e.g., conducting interviews)

  _____

  _____

- Data analysis (e.g., coding data)

  _____

  _____

- Dissemination (e.g., planning events, writing up results)

  _____

  _____

_____ 4. What resources, materials, and activities will you use to help yourself and any academic co-researchers learn about the topic, issue, and/or community (e.g., readings, websites, orienting activities)?

_____

_____

_____ 5. What training in data collection and analysis will be required of you and/or your research partners? Who will provide the training?

_____

_____

_____ 6. Given your proposed design, how much of your time is the study likely to take?

_____

_____

_____ 7. What challenges might you face related to the implementation of the study (e.g., time, ensuring collaboration)? How will you address each of these challenges?

_____

_____

_____ 8. What is your dissemination plan?

_____

_____

- What traditional, scholarly venues will you use to report research findings?

  _____

  _____

- What sorts of innovative approaches and methods will you use to reach a wider audience?

  _____

  _____

_____ 9. What specific ideas do you have for using the research findings to enact changes related to your topic of concern?

_____

_____

MyLab Education Self-Check 10.2

MyLab Education Application Exercise 10.1: Selecting and Designing an Action-Oriented Design

DATA COLLECTION AND ANALYSIS IN ACTION RESEARCH

The types of data collected in action-oriented studies are not that different from the data collected in other designs. Most of the designs described in this chapter make use of both quantitative and qualitative approaches, including experiments/interventions, surveys, collection of demographic data, interviews, focus groups, observations, and inspection of textual documents, visual images, and social media sources. Obviously, community-based qualitative research focuses on the collection of qualitative data for analysis, although some quantitative demographic data might be used to provide a description of the community setting. Design-based research emphasizes interventions and the collection of quantitative data as a way of making claims about the effectiveness of particular practices. Although the forms of data collected might vary from model to model and from one study to the next, certain common features of data collection and analysis are noteworthy in action-oriented studies.

1. ***Data-collection strategies are emergent.*** Data collection in action-oriented studies isn't guided by any prescriptive or otherwise fixed and rigid view of data collection. Although researchers usually generate an initial list of data that should provide insights into their research questions, other data sources might be added after the study is under way. Ideas for additional data-collection approaches and sources might be generated from discussions with community partners who possess knowledge regarding the sorts of evidence that could shed light on the topic and help researchers address key issues.

2. ***Data collection is iterative.*** In action-oriented studies, data are collected in iterative cycles that are informed by data analyses; for example, an initial analysis of focus-group data could reveal issues and concerns that would be explored in more depth in individual interviews with key stakeholders. Participants might spontaneously share personal writings or documents with researchers, which could lead to requests for similar documents from other participants.

The emergent and iterative nature of action-oriented designs means that researchers might be continually responding to events and occurrences in the research setting that, in some instances, take the research effort in unexpected directions. For example, research purposes and questions may be revised, leading to the need to collect new and different types of data. Unanticipated new demands placed on teachers, practitioners, and/or community partners—perhaps by school districts, state agencies, external funding agencies, and/or constituents—might also require new and different forms of data. And as researchers immerse themselves in data collection, some issues might emerge that are more salient and relevant to the setting than the ones identified at the outset of the study, again calling for a shift in the focus of data collection. For such reasons, researchers must be flexible in responding to changing circumstances and conditions and willing to tinker with and revise their data-collection and data-analysis procedures.

3. ***Data collection is collaborative and ultimately the result of a coordinated group effort.*** As we have stressed throughout the chapter, action-oriented research designs are more collaborative than other paradigms, in that they involve building research partnerships with schools, teachers, practitioners, youth, community-based organizations and groups, and/or community members. Research group members share responsibility for collecting and analyzing data and might work in pairs to administer surveys, conduct interviews, and/or meet with focus groups.

Such collaborative data-collection activities have the additional benefit of enhancing research partners' individual and collective knowledge and capabilities. For example, academic

researchers might share methodological knowledge with partners who have had less research experience, possibly providing training and offering support for the development and refinement of skills (e.g., through practice sessions or mock interviews). Teachers, practitioners, and other stakeholders can provide invaluable knowledge about the setting and predominant practices and beliefs, all of which can be invaluable in the identification of data-collection strategies. Furthermore, this local knowledge might be essential to interpreting and making sense of complex data and concepts. For instance, group members might engage in **collective coding**, in which members individually code pieces of data and then come together to generate and decide on a list of common codes that can be used to code the other data (L. R. Johnson, 2017).

It's important to note here that members of a research team won't necessarily all be involved in data collection and analysis to the same extent. Ultimately, the team will want to make the best use of each group member's expertise and skills; also, it shouldn't burden any individual with an unreasonable number of tasks and responsibilities. Community partners, in particular, can have many demands on their time. When engaging in collaborative research endeavors, then, it's important to clarify roles and responsibilities in advance. And especially in large and complex research projects, it may be necessary to establish memoranda of understanding and formal agreements regarding the scope and details of collaboration, as well as regarding who ultimately owns the data and how the data will be used (Hacker, 2013).

MyLab Education Self-Check 10.3

MyLab Education Application Exercise 10.2: Conducting Action-Oriented Research

DISSEMINATING THE FINDINGS OF ACTION RESEARCH PROJECTS

In this chapter, you've learned about various action-oriented research designs and how they might be used to engage in collaborative and reciprocal research projects with teachers, practitioners, and community members, with the ultimate purpose of obtaining findings that can lead to substantial changes and transformations in policy, programming, and/or practices. Such a goal is quite different from that of most traditional research studies, which certainly aim to have an impact on a field of study but usually don't have an explicit agenda for action. In our preceding discussions of particular research models, we have often described potential dissemination activities for achieving *change* goals.

In this section, we summarize some of the primary ways in which action-oriented researchers report and share their findings. You might think of dissemination strategies for action research as lying along a continuum going from "traditional" at one end (e.g., common for design-based research) to very "action-oriented" at the other end (e.g., common for PAR/YPAR designs). The activities in the following subsections are roughly organized along this continuum, beginning with fairly traditional dissemination strategies and moving on to strategies that are less conventional and quite action-oriented.

Journals and Books

The most commonly used and traditional dissemination strategy in academia is the publication of articles in peer-reviewed journals, chapters in edited books, and for some especially productive scholars, entire scholarly books on specific research topics. Publications of this nature are a necessity for academic faculty members who are intent on earning tenure and promotion. Even though action-oriented researchers tend to be more interested in enacting change than might be true for some of their institutional colleagues, the age-old axiom of "publish or perish" holds true for them as well. For example, given that design-based researchers are interested in advancing theory, they are quite likely to publish their work in traditional peer-reviewed, research-oriented journals. In contrast, because action and teacher researchers are primarily interested in informing other teachers and practitioners about their work, they are apt to publish some or all of their work in

practitioner-oriented journals (for examples, we refer you back to the chapter's earlier section on teacher research). Numerous journals devoted to service learning and community-based research are appropriate venues for the publication of CBAR and CBQR studies. The results of many PAR/YPAR studies have been reported as part of sole or coauthored books or edited volumes dedicated to participatory research studies (e.g., T. M. Brown & Rodriguez, 2009; Cammarota & Fine, 2008; Dyrness, 2011; Irizarry, 2011; Nygreen, 2013).

Conferences and Symposia Another way that action-oriented researchers disseminate their findings is by presenting papers, posters, or roundtable discussions at local, national, and international conferences. In contrast to journals and books, which may have a rather lengthy revision and production process—sometimes more than 2 years—conference presentations usually require less time to prepare than an article manuscript, and the submission and acceptance process is usually less onerous. Many conferences are practitioner-focused, and some are specifically dedicated to action-oriented and participatory approaches; also, many broader research conferences have strands or groups for particular action-oriented designs, such as design-based research and participatory action research.

Conference presentations and panels can also be a good fit for research teams that include practitioners, youth, and/or community members because they allow for wider participation of group members, including those who (perhaps because of time constraints or skill level) haven't been involved in the preparation of articles or book manuscripts. Furthermore, presentations at conferences and at small, topic-specific symposia provide the opportunity for dialogue with other scholars and practitioners who are interested in the topic that was the focus of investigation, general strategies for conducting action-oriented studies, or both.

Public Forums and Events Because action-oriented researchers are often interested in engaging a fairly broad public in the research process, public forums and events can be an effective means for sharing findings with a wide audience. And because community-based researchers, in particular, are interested in using research findings to guide community decision making and programming, community events can be a great way to get feedback from community members and further develop and refine ideas for action. For example, teacher researchers might make presentations at a school forum for administrators, parents, and/or other teachers in order to (a) report their study's findings related to a particular practice and (b) discuss how these findings might be used to enhance instruction. Effective forums require considerable advance planning, however.

Program Resources and Materials One goal of a research study might be to design and produce certain resources and materials—for instance, innovative instructional materials or grants for funding program extensions—that can help strengthen or expand programming. For example, teacher researchers might develop curricula, lesson plans, or classroom management strategies based on their research findings; design-based researchers might create teaching guides that can be used to improve instructional practices. Members of a community-based research team might prepare a grant for funding a new program that addresses certain needs identified in their research study. Researchers might also create a website that includes information about the study and allows for ongoing updates of findings and links to related resources.

Advocacy and Activism At the least traditional, most radical end of the action-oriented dissemination spectrum are efforts aimed at changing broader policy and taking direct action on particular issues or concerns. Members of a research team might engage in lobbying efforts or organize letter-writing campaigns to legislators on behalf of a certain issue. The research group might also organize protests or demonstrations to promote awareness of a community concern raised within the research. For studies involving youth, the use of social media, such as hashtag campaigns and Twitter posts, can be effective ways to disseminate research findings and reach a wide audience related to a particular issue.

PRACTICAL APPLICATION Using Community Forums as a Means of Disseminating the Results of Action Research and Participatory Research Projects

Community forums can be a great way for those conducting community-based and participatory research to share findings from research studies with a broader public and, on the flip side, to gain insights and feedback from community members and stakeholders. If implemented strategically, forums also provide opportunities for researchers to engage in dialogue with the community related to the research topic and further develop and refine ideas for action. If not well organized, however, such events may not be terribly successful in achieving a research study's goals for action and transformation.

GUIDELINES Planning and Conducting an Effective Community Forum

Following are several tried-and-true suggestions for organizing and developing a successful community forum.

1. ***Meet with co-researchers and community partners in advance to co-plan the event.*** For example, draft an invitation list that includes community leaders and stakeholders from local organizations. Also invite individuals from local academic institutions, especially administrators and faculty working in the area of community–university partnerships and especially those with which one or more members of the research team are affiliated.

2. ***Identify an appropriate location.*** If possible, choose a venue that is easy to get to via public transportation; alternatively, consider offering shuttle service. If attendees will need to drive to the location, there should be free or inexpensive parking on-site or nearby. And perhaps most importantly, the venue should be one that is accessible for people with disabilities.

3. ***Widely publicize the event.*** For example, eye-catching leaflets and brochures can be created and distributed to local organizations and businesses (e.g., cafés, grocery stores, beauty salons, and barber shops). Announcements can be put in local newspapers and newsletters. Announcements can also be made at town meetings or other local events. Social media, such as Facebook and special-interest blogs, have the potential to reach a great many people and, indirectly through them, an even larger audience. And personalized invitations—perhaps in the form of written letters, e-mail, telephone calls, and face-to-face encounters—can be especially effective in recruiting a broad range of participants.

4. ***Create a visual presentation.*** With a little training and practice, you can easily create a polished presentation through PowerPoint, Prezi, or similar presentation software. It should include information about the background and context for the study, key aspects of the problem under investigation, the study's design and methodology, a summary of the study's findings, and concrete ideas for action. Be sure to give co-researchers and community partners ample opportunity to review and critique—and *strengthen*—the presentation beforehand.

5. ***Prepare handouts.*** Well-designed handouts can help audience members follow along with a presentation and remember key findings and recommendations. Handouts should also include the contact information of research team members and participating organizations to enable follow-ups by interested and potentially helpful parties. Also consider whether handout

materials should be available in two or more languages to accommodate the linguistic diversity of participants.

6. ***Provide refreshments, such as beverages and light snacks.*** If possible, make use of local restaurants, delicatessens, catering services, or similar businesses. Relying on local businesses is just one more way—and potentially a very productive way—of reaching out to other community members in your efforts to bring about change.

7. ***Provide any appropriate resources for attendees.*** Depending on the topic and likely attendees of the forum, helpful resources could take such forms as interpreters, on-site childcare, or free shuttle service. For example, on-site childcare services might be essential if the topic of a forum is community resources for young parents; they are much less likely to be necessary for a research project investigating how to maintain safe and affordable housing for senior citizens.

8. ***Allow sufficient time for questions and dialogue.*** Generally speaking, two-way conversations and exchanges of information are far more effective and useful than strictly one-way, didactic presentations. You should probably assign a moderator—ideally, someone who is widely respected throughout the community—to introduce the event and facilitate question-and-answer sessions. Alternatively, you want to have *two* moderators, one representing the university and the other representing the community. In addition, you could distribute comment cards that audience members could use to submit their ideas and feedback.

9. ***Consider making the event available on the Internet, either through live-streaming or a video recording.*** Rarely can all interested individuals attend an event at a single, specific time and place. For those individuals who cannot be there in person, live-streaming the event (e.g., via Facebook, a university website, or webinar software) gives them an electronic way of both attending and, if the electronic platform allows, actively *participating in* the event. And in cases in which the timing is bad, video recordings can provide a viable and effective alternative.

Identifying realistic plans for dissemination often requires a balancing act between addressing academic researchers' publication requirements at their respective institutions, on the one hand, and the legitimate demands of practice-oriented research partners, on the other (Hacker, 2013). An easy way of making everyone happy, and simultaneously having a major impact on policies and practices, is to describe your findings not only in print and in person but also through the multimedia and the Internet— in other words, in as many ways as possible.

MyLab Education Self-Check 10.4

MyLab Education Application Exercise 10.3: Disseminating Findings from Action-Oriented Studies

A SAMPLE RESEARCH REPORT

We end this chapter with excerpts from a hypothetical research report about an action research study. This research focused on the increasing violence and aggression among children of middle school (grades sixth to eighth) and explored the intervention strategy of mindfulness meditation over a five-week course to see whether it curtailed aggression by students in the school and improved their emotion regulation strategies.

As you will discover when you read the excerpts, this study was carried out by two teachers with the support of their school administration. The researchers collected data through multiple methods and at multiple points in time to determine if mindfulness meditation could be used as a tool in schools to improve school climate and to reduce student violence and aggression. Although they couldn't conclusively determine that their intervention strategies were the *cause* of the decreases in student aggression that they observed—they would have needed an experimental design to draw such a conclusion—their findings were sufficiently compelling that they concluded the following:

> We found sufficient evidence, through the means of the verbal reports of the students, the verbal reports of the teachers, our own observations and reflections, to conclude that mindfulness meditation can serve in integral and important ways to curb violent and aggressive behaviour among students and enhance their capacity for reflection and co-operation . . .
>
> This study will have to be replicated and adapted to the specific needs of each set of students or schools . . . (*A citation would be included here if this was a real, rather than a hypothetical, study.*)

DISSERTATION ANALYSIS 7

Middle school is a time which involves surging changes for students, with the rising academic workload, the concerns brought about by the onset of adolescence and the beginning of the striving to find one's own identity and carve a niche. Amidst this chaos, students often struggle to manage all the issues surrounding them. This is especially true because schools often do not have any organized redressal process for students. Counseling services are not provided at many schools and even if they are, they are rarely assessed for their ability to provide support to middle school students. There is also lack of awareness and lack of initiative on part of the students as well as the teachers and the parents to make use of such services. The only support system available to many of the students are their peers who many not be in a position to guide them effectively.

This study came about as a result of our observations as we went about doing our daily duties and routine tasks at school. We were concerned about the upward surge in the cases of bullying and violent and aggressive behaviour by the students in our classroom. A lot of middle school students from our classrooms we talked to mentioned that they often engaged in violent behavior and aggression because their friends talked them into it, or because they were provoked by other students. Importantly, many of the students we talked to displayed little insight about the impact of their (violent) behavior on others and little empathy about the feelings of recipients of the aggressive act. These students felt that the education they received at school didn't address student relationships and did little to alleviate tensions and aggressive behavior.

Comments

As can be seen, action research seeks to set the specific context of the community with which the research is being carried (in this case are students middle school) and underline the need for working with them.

Teacher research may involve the use of more than one researcher as seen here by the use of "our". It also relies on support from other stakeholders in the education system.

As is the case in action research, the topic of study is located in the participants' real-life concerns. Herein, the teachers have reflected upon the serious concerns being faced by the students as the motivation behind carrying out the study.

As a result, we, in the capacity of both teachers and researchers, focused on using effective strategies to provide students with resources that might help them manage their aggressive behaviors and develop attributes that foster positive student relationships and, thereby, curbing the prevalence of aggressive and violent behavior among them. We chose mindfulness meditation as our intervention strategy and oriented its incorporation through various practical and theoretical exercises over a five-week period using the following questions as our guide:

Designing interventions to improve pedagogical outcomes, which in this study includes improving the overall well-being of students and enhancing their learning outcomes, is central to teacher research.

1. What are the primary concerns and issues of middle school students with respect to school violence and aggression?
2. In which ways do the students think that the school can help them in nurturing their selves and managing aggressive behavior?
3. Which coping resources serve as long term anchors to help students work towards a sense of reconciliation, empathetic understanding, and positive student relationships?

Participants' voices and opinions are important in action research, which is what led to the researchers embracing them from the very beginning, as seen in the guiding questions of the research.

(At this point, the authors have included a 'Background' section which provides the theoretical concepts and literature review related to their chosen topic of study. It covers aspects related to the well-being of middle school students and the use of meditation programmes as interventions in the field of education.)

Methodology

Participants

The study focused on 35 students in a single section of the seventh grade of a private school. This included 19 boys and 16 girls. The average age of the students was 12 years. Out of these, 5 (all boys) had been reported by teachers for violence in the school while 2 of the girls had been reported by a teacher to have engaged in bullying of a few other girls in another section of the same grade. One girl had sought counseling services of the school counselor in the recent past. All the children came from middle class and upper middle class families.

The nature of action research is such that it usually involves a small sample. Here too, the intervention is focusing on children of one section at a time. The researchers give additional details about the students in order to provide more than just cursory information to the reader about the participants.

Intervention

We designed our intervention based on our review of previous studies about the relationship of mindfulness traits and education with aggressive/violent behavior and psychological well-being in young people in educational settings. We further adapted the module in terms of the exercises, discussions, and assignments included in it on the basis of a pre-assessment survey (Appendix A). We included the responses of the students across all the sections of the seventh grade in the pre-assessment data. This yielded a total of 280 responses. In the pre-assessment survey which included a total of six open-ended questions, we asked students about their understanding of other bullying victims' feelings. We also asked them for their opinions on mindfulness and whether they thought that they would be curious to explore it further as a way of regulating their emotions and improving their relationships with other students. We further asked them about topics that they would want included in case such a practice-based curriculum was to be set up. We enquired from them other ways in which they thought teachers could render further support to them.

Action research uses interventions customized for the community under study. Hence, it is no wonder that the researchers have adapted the intervention module after conducting a pre-assessment survey to better understand what the students are seeking and what their concerns are.

(The authors then discuss the findings from the pre-assessment survey in detail.)

The intervention consisted of a five-week module covered as a part of two forty minute sessions each week. Each session included a 15-minute formal meditation practice focusing on regulating anger and negative emotions. Most of these short meditations were developed by us and validated by experienced mindfulness practitioners. Some of them were taken from other public domain sources and mindfulness practice websites as well as from published studies using mindfulness-based techniques for similar purposes. Homework exercises included daily formal and informal mindfulness practice sessions. Reflective exercises focusing on managing emotions and relations with other students were a vital part of the intervention and included brief journal entries each week focusing on the students' emotional states and the positive and negative experiences of the week. These journal entries were meant for only the students to be able to pay attention to their inner feelings and inner world. Discussions were often initiated in the sessions which focused on the benefits of cultivating positive attributes such as those of kindness and compassion, and how these attributes could foster positive student relationships.

Interventions in action research are comprehensive, and involve active engagement of the participants, as can be seen here as well.

As seen in the activities which are a part of the intervention module, instructional exercises are often a part of teacher research. The act of teaching cannot be separated from the process of a teacher research.

The first week focused on acquainting the students with mindfulness practice. The first session included a beginner's meditation on paying attention to the breath (Appendix B) and a discussion on how being able to regulate our attention can be useful for us. The homework assignment (Appendix C) included going to a nearby park and paying attention to the surrounding sounds for 15 minutes. A report was to be submitted about what ordinary and unordinary sounds the students encountered, including a brief reflection on what the students felt during this experience.

The researchers have provided extensive details about the intervention module and have also provided copies of the instructional materials in the appendices. This is to keep the reader invested in the central piece of an action research, i.e., an intervention which has practical, real-life value. In a way, it also makes materials accessible to other researchers who wish to replicate the study.

(The authors continue to describe the intervention module as it proceeded along the different sessions, to the very last tenth session at the end of the fifth week.)

The end of module assignment included a reflective piece by the students regarding whether and which changes they had observed in themselves over the course of the past five weeks. They also had to mention whether and how the changes in their own thoughts, emotions and actions influenced their external surroundings, including friends, classmates, and other people they interacted with.

In an action research, data collection tool may also serve as an instructional tool. The end-of-module assignment reflects that instruction and intervention are often intertwined.

Eventually, a post assessment survey (Appendix E) was given to the 35 students who were a part of the intervention. The survey included questions on how effective the students found the module in terms of its efficacy in improving their emotion regulation capacities and on how likely they would be to be a part of a similar module again, to be rated on a five- point Likert scale. Two open-ended questions enquired about the most prominent benefits of the intervention experienced by the students and what their reasons would be for continuing or not continuing the meditation practice and the other informal exercises that were a part of the module.

Action research seeks to involve participants at all stages of the research process. Their voices are integral to drawing any kind of conclusions. A post-assessment survey and verbal reports of the students are two such ways to value their perspectives even after the formal intervention module is over.

Data collection tools

Multiple data collection tools were utilized to obtain a comprehensive picture of the impact of the intervention on the students. Pre- and post-assessment surveys were also used to get an idea about the students' own observations about the changes in their internal and external world.

The study makes use of multiple forms of qualitative and quantitative data, which should help the researcher better substantiate assertions.

We also interacted personally with each of the 35 students to obtain feedback from them regarding the module and whether and what differences they think

they have achieved in their lives as a result of the intervention module. We interacted with the other teachers who teach the students on a daily basis in order to get an external perspective on perceptible changes in the students' behaviour. This helped us gain information with respect to the attention pattern of the students, their engagement in the classes and attributes displayed by them in their interpersonal and social interactions.

Note: This is a hypothetical study.

As you can see here, the researcher made use of multiple forms of data in order to triangulate findings—a feature of many rigorous qualitative, quantitative, and mixed-methods studies.

APPLICATION EXERCISES

Similar exercises are available on Pearson MyLab Education.

1. Answer the following questions on selecting and designing an action-oriented study.
   i. A sixth-grade teacher is interested in conducting an action research study to better understand her students' mathematics learning and to inform her instruction in the area of mathematics. What steps should she take to undertake the process of conducting research in her classroom? What challenges might she face as she attempts to collect data? How might she address these challenges?
   ii. A university researcher who studies and teaches environmental issues has been wanting to provide her graduate students with opportunities to conduct research in community settings. How could she do this? How should she proceed? What will be involved in the planning process?
2. Answer the following questions on conducting action-oriented research.
   i. An academic researcher is conducting a Youth Participatory Action Research (YPAR) study that involves youth in investigating violence in their community and devising solutions for addressing this violence. What strategies for data collection and analysis might be useful to investigate this issue/topic? What should the researcher be mindful of throughout the data-collection process?
   ii. A design-based researcher is working with a group of reading specialists at an elementary school to investigate particular innovative instructional strategies for teaching reading in the early grades. He is meeting with teachers to discuss ideas for the study. What types of data-collection approaches would be appropriate? What would the researcher want to know more about from the reading specialists in order to make decisions about the study?
3. Answer the following questions on disseminating findings from action-oriented studies.
   i. The members of a research team have just completed a YPAR study examining youth civic engagement at their school and in the broader community. How might they disseminate their findings to various audiences?
   ii. A group of community-based researchers is interested in hosting a community forum to share research findings and engage community members in dialogue related to the topic of the study. What steps should the group members undertake as they plan the forum?

FOR FURTHER READING

Ballenger, C. (2009). *Puzzling moments, teachable moments: Practicing teacher research in urban classrooms*. New York, NY: Teachers College Press.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences, 13*(1), 1–14.

Brown, A. (1992). Design experiments: Theoretical and methodological challenges in creating interventions in classroom settings. *The Journal of the Learning Sciences, 2*(2), 141–178.

Cammarota, J., & Fine, M. (Eds.). *Revolutionizing education: Youth participatory action research in motion*. New York, NY: Routledge.

Collins, A. (1992). Towards a design science of education. In E. Scanlon & T. O'Shea (Eds.), *New directions in educational technology* (pp. 15–22). New York, NY: Springer.

Dyrness, A. (2008). Research for change versus research as change: Lessons from a *mujerista* participatory research team. *Anthropology & Education Quarterly*, *39*(1), 23–44.

Dyrness, A. (2011). *Mothers United: An immigrant struggle for socially just education*. Minneapolis, MN: University of Minnesota.

Fishman, S. M., & McCarthy, L. (2000). *Unplayed tapes: A personal history of collaborative teacher research*. New York, NY: Teachers College Press.

Ginwright, S., Noguera, P., & Cammarota, J. (Eds.). (2006). *Beyond resistance: Youth activism and community change*. New York, NY: Routledge.

Hacker, K. (2013). *Community-based participatory research*. Thousand Oaks, CA: Sage.

Hall, B. L. (1992). From margins to center? The development and purpose of participatory research. *The American Sociologist, 23*(4), 15–28.

Hofman, N. G., & Rosing, H. (Eds.). (2007). *Pedagogies of praxis: Course-based action research in the social sciences*. Boston, MA: Anker.

Irizarry, J. G. (2011). *The Latinization of U.S. schools: Successful teaching and learning in shifting cultural contexts*. Boulder, CO: Paradigm Press.

Johnson, L. R. (2017). *Community-based qualitative research: Approaches for education and the social sciences*. Thousand Oaks, CA: Sage.

Jones, S. (Ed.). (2017). *Writing and teaching to change the world: Connecting with our most vulnerable students*. New York, NY: Teachers College Press.

Kirshner, B., O'Donoghue, J., & McLaughlin, M. (2005). Youth-adult research collaborations: Bringing youth voice to the research process. In J. L. Mahoney, R. W. Larson, & J. S. Eccles (Eds.), *Organized activities as contexts of development: Extracurricular activities, after-school and community programs* (pp. 131–156). Mahwah, NJ: Lawrence Erlbaum.

Lather, P. (1986). Research as praxis. *Harvard Educational Review, 56*(3), 257–277.

McIntyre, A. (2008). *Participatory action research* (Qualitative Research Methods Series, 52). Thousand Oaks, CA: Sage.

Morrell, E. (2004). *Becoming critical researchers: Literacy and empowerment for urban youth*. New York, NY: Peter Lang.

Mills, G. E. (2018). *Action research: A guide for the teacher researcher* (6th ed.). New York, NY: Pearson.

Nygreen, K. (2013). *These kids: Identity, agency, and social justice at a last chance high school*. Chicago, IL: University of Chicago Press.

Pardo-Ballester, C., & Rodriguez, J. C. (2009). Using design-based research to guide the development of on-line instructional materials. In C. A. Chapelle, H. G. Jun, & I. Katz (Eds.), *Developing and evaluating language learning materials* (pp. 86–102). Ames, IA: Iowa State University.

Reason, P., & Bradbury, H. (Eds.). (2006). *Handbook of action research*. London, England: Sage.

Rief, L. (2014). *ReadWriteTeach: Choice and challenge in the reading-writing workshop*. Portsmouth, NJ: Heinemann.

Stoeker, R. (2013). *Research methods for community change: A project-based approach*. Thousand Oaks, CA: Sage.

ANSWERS TO THE CONCEPTUAL ANALYSIS EXERCISE "Choosing an Action-Oriented Design":

1. This project calls for *youth participatory action research* (YPAR), in which young people participate as co-researchers. In instructional settings, a YPAR study can be quite useful as a means of facilitating students' involvement in and understanding of contemporary issues.
2. Researchers and practitioners typically conduct *community-based research* when they want to jointly learn more about a local problem and then develop some sort of intervention or other program that can help to address the problem.
3. This is a situation in which *action research/teacher research* is appropriate. His research question is specific to his own classroom, and he doesn't necessarily need one or more co-researchers or a community organization to assist him in his efforts.
4. In this case, *community-based research* is called for because she wants her students to become temporary partners in addressing a local problem. Probably the best model for this situation is *course-based action research* (CBAR).
5. The goals of both contributing to a larger theoretical knowledge base and designing a new instructional intervention (perhaps through multiple iterations and modifications) are best served with *design-based research.*

Chapter 11 Analyzing Quantitative Data

Numbers are meaningless unless we can discern patterns that lie within them. With statistics, we can summarize large numerical data sets, make predictions about future trends, and determine whether various experimental treatments have led to significantly different outcomes. Thus, statistical procedures are among the most powerful tools in a researcher's toolbox.

Learning Outcomes

11.1 Describe several strategies for organizing a data set in order to reveal possible patterns within the data.

11.2 Explain how the nature of the data may limit the kinds of statistical analyses that can be conducted.

11.3 Distinguish between descriptive statistics and inferential statistics, and describe various purposes that each of these two categories of statistics can serve, including (a) estimating central tendency, variability, and correlation; (b) estimating population parameters; (c) testing hypotheses; and (d) conducting a meta-analysis.

11.4 Describe several ways in which you might use the results of statistical analyses to help you interpret and impose meaning on a data set.

In quantitative research, we try to make better sense of the world through measurement and numbers. Sometimes the numbers represent aspects of the observable, physical world, such as the weights of concrete objects, the growth rates of invasive species, or the number of people engaging in particular activities. We may also use numbers to represent non-physical, intangible phenomena, such as how much students learn in the classroom, what beliefs people have about controversial topics, or how much influence various news media have on political decision making. We can then summarize and interpret the numbers with 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. Statistics help us answer a critical question: *What do the data mean?* In other words, what message do they communicate?

EXPLORING AND ORGANIZING A DATA SET

Before using any statistical procedure—before making a single computation—look closely at your data and consider potentially productive 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—although admittedly a very artificial one—to illustrate the point. Following are the scores on a reading achievement test for 11 children, arranged alphabetically:

Adam	76	Mary	92
Alice	80	Ralph	64
Bill	72	Robert	60
Chuck	68	Ruth	96
Kathy	84	Tom	56
Margaret	88		

What do you see? Perhaps you've noticed that the highest score was earned by a girl and the lowest score was earned by a boy. Hmm, might gender be an important dynamic in the data set? Let's arrange the scores horizontally across the page, attach gender designations to each one, and see what happens:

76	80	72	68	84	88	92	64	60	96	56
♂	♀	♂	♂	♀	♀	♀	♂	♂	♀	♂

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.

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 11.1, the trends are clear: The girls' scores increase as we proceed through the alphabet, and the boys' scores decrease. Furthermore, the scores are equidistant from one another: Each score is 4 points either above or below the preceding one.

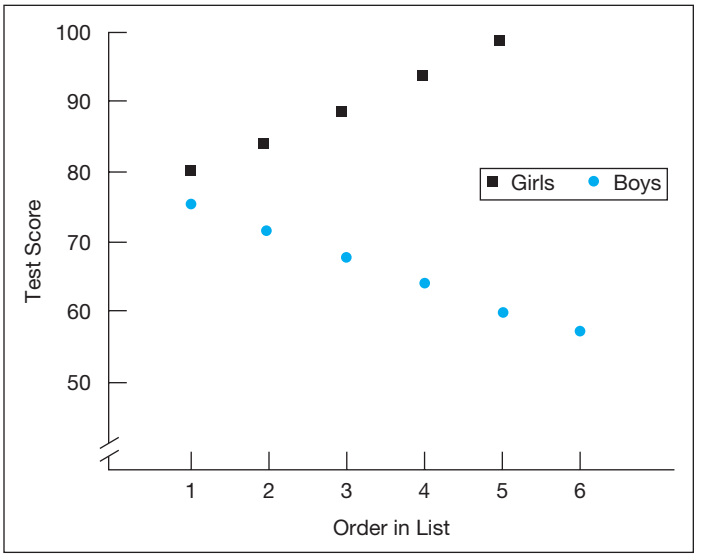


FIGURE 11.1 A Visual Representation of the Reading Achievement Test Scores

Whatever we've observed may have no relevance whatsoever for our project, but because it represents *dynamics within the data,* it's important that we see it. In the preceding example, 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 can reveal important new insights. 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 contemporary scientific theory.

In the 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 surnames—which would have been equally illogical—we would still have seen that the girls had higher scores than the boys, but we wouldn't necessarily have seen the diverging trends depicted in Figure 11.1.

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 previously 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 all at once. In fact, unless a 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 various 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—obtained directly from the situation. Listed in chronological order, they don't seem to say very much, except that Joe's performance has been inconsistent.

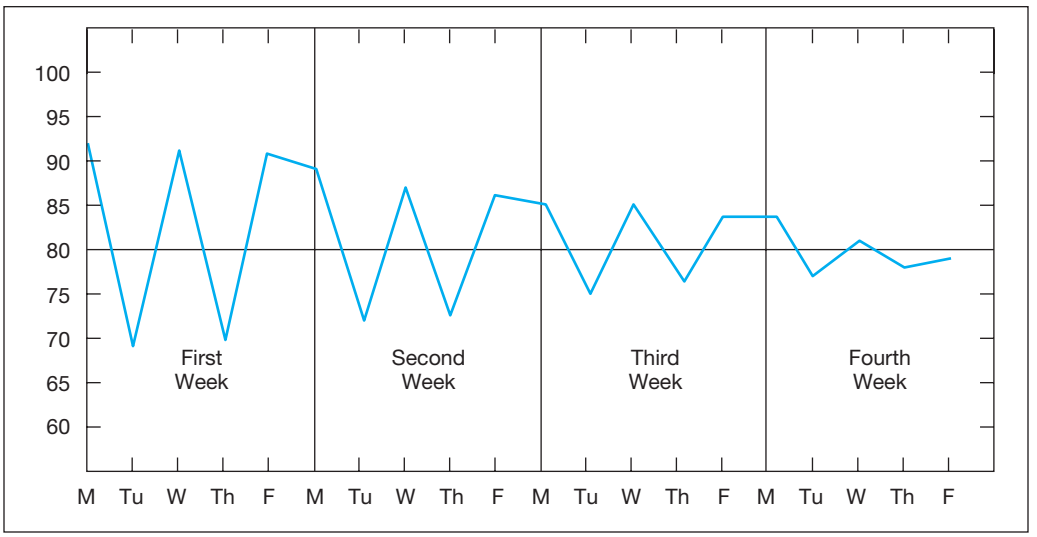
Let's put Joe's grades in a *two-dimensional table* organized by weeks and days, as shown in Figure 11.2. The table reveals some patterns in Joe's grades. If we compare the five columns, we notice that the grades on Mondays, Wednesdays, and Fridays are considerably higher than those on Tuesdays and Thursdays. 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.

FIGURE 11.2 The Reading Achievement Test Scores in Table Form

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

FIGURE 11.3 Line Graph of Joe's Daily Grades



Now let's represent Joe's grades as a *line graph,* shown in Figure 11.3. In this graph, we see phenomena that weren't 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 an alert researcher to explore the data further in an attempt to explain the pattern that the graph reveals.

Graphing data is often quite useful for revealing patterns in a data set—that is, for making its overall nature more transparent and obvious (e.g., Tay, Parrigon, Huang, & LeBreton, 2016). For example, let's return to a study first described in a Practical Application exercise near the end of Chapter 7:

> 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 7 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 created 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.

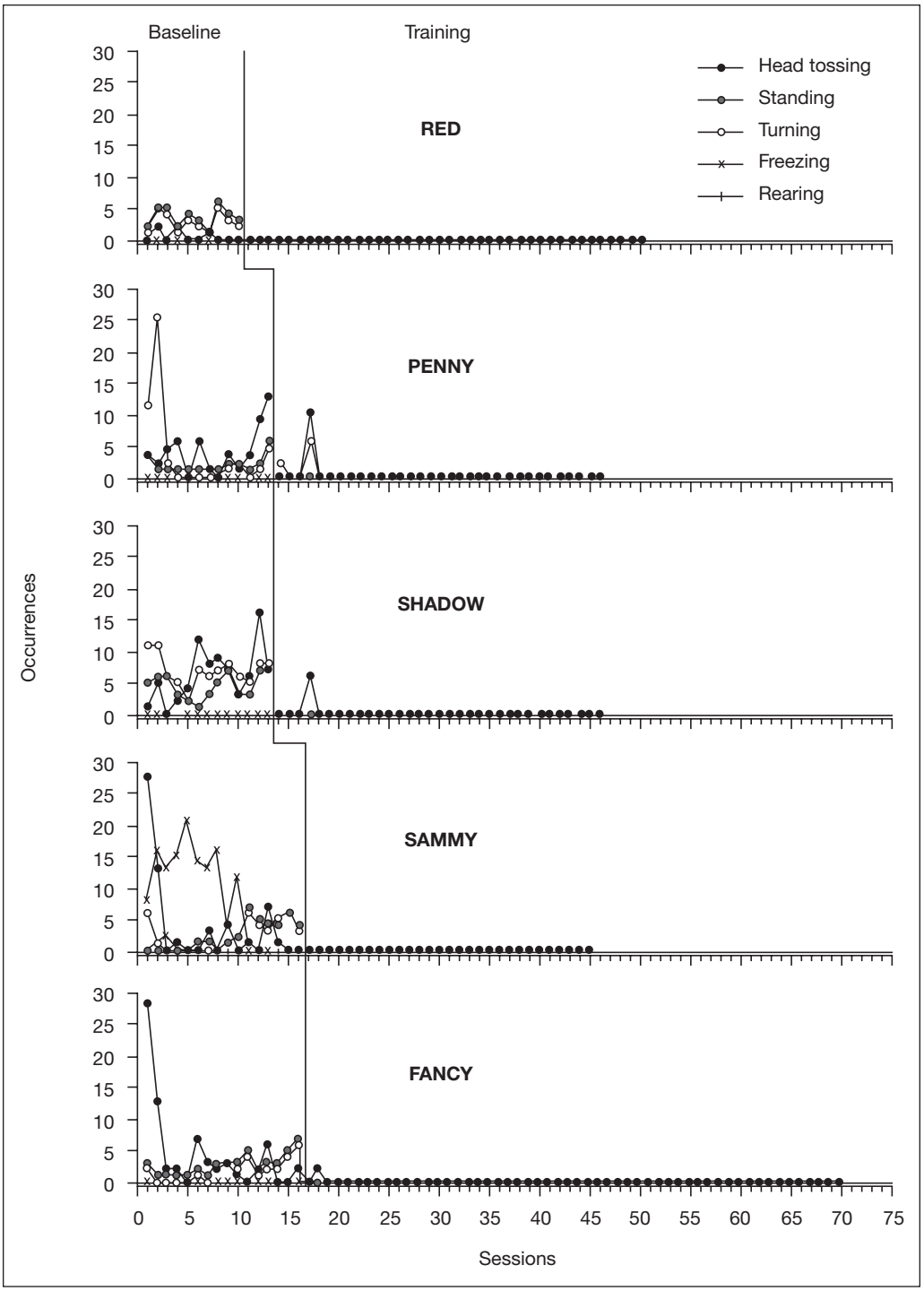


FIGURE 11.4 Undesirable Trailer-Loading Behaviors for Five Horses Before and After Training

Source: 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.

Time-series studies sometimes yield data that show clear hit-you-between-the-eyes patterns; for another example, return to Figure 7.3 in Chapter 7.[1] 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.

[1]Some statistical techniques have been proposed for time-series results that *don't* have a clear-cut, hit-you-between-the-eyes quality. For our readers who are comfortable with complex statistical techniques, we refer you to Ferron and Levin (2014), Shadish (2014), and Shadish et al. (2014).

Using Computer Spreadsheets to Organize and Analyze Data

The process of organizing large amounts of data was once a tedious, time-consuming task. Fortunately, computers have made the process much simpler and more efficient. One important tool is an **electronic spreadsheet**, a software program that enables a researcher to enter and then manipulate data in a two-dimensional table. Probably 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., Gnumeric, Kingsoft Office, LibreOffice, OpenOffice Calc, Spread32).

The beauty of an electronic spreadsheet is that once you enter data into it, the software can quickly and easily help you organize the data and perform 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 an entry and so must correct it—any calculations that are based on that entry are automatically updated. The software typically also lets you copy (*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's possible to reorganize them in any way you wish. For example, suppose you have assessment scores for a large number of people 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 participants' ages. In a matter of seconds, an electronic spreadsheet can sort the scores by age and list them from youngest to oldest participant, 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 assessment scores for children 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 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 appropriate parts of the data, the program can automatically generate a graph from those data—perhaps a line graph, bar graph, or pie chart. Users can select how the axes are labeled, how the legend is created, and how 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're 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?"

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.

MyLab Education Self-Check 11.1

MyLab Education Application Exercise 11.1: Organizing Data in Tables and Graphs

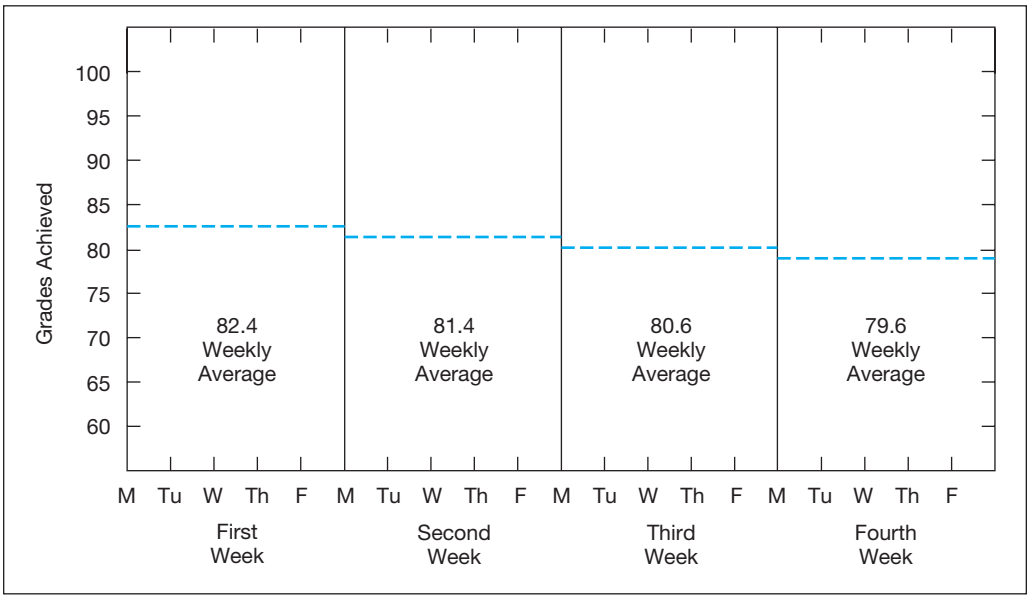
CHOOSING APPROPRIATE STATISTICS

In a single chapter we cannot thoroughly describe the many statistical procedures that researchers might use. 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 authors 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.

Earlier 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're 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 4 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 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

FIGURE 11.5 Line Graph of Joe's Weekly Average Grades



a *standard deviation*—to summarize the variability we see each week. (The nature of such statistics will be described shortly.)

Thus far we have discovered an important point: *Looking at data in only one way yields an incomplete view of those data and their underlying meanings.* For this reason, we have many statistical techniques, each of which is suitable for a different purpose and can answer a different question for a particular set of data.

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 limit 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 associated with one another, and the like. 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're an immigration officer. Although you've never been to Egypt, you've met numerous Egyptians as they disembark from incoming planes and ships. Perhaps you've even become well acquainted with a small number of Egyptian people. 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 the people you've met aren't necessarily representative of the entire population of Egypt. However, that's a sampling problem, not a statistical one.

More generally, inferential statistics involve using one or more small samples and then *estimating* the characteristics of the population from which each 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 might be 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, in that it's always 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, whereas small circles have short radii. Thinking of a parameter in this way, we know that any 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 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.

TABLE 11.1 Examples of Conventional Statistical Notations for Population Parameters and Sample Statistics

	The Symbol Used to Designate the Characteristic	
The Characteristic in Question	*Population Parameter*	*Sample Statistic*
The mean	μ	M or $\bar{X}$
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*.

Considering the Nature of the Data

As you begin to think about statistical procedures that might be most appropriate for your research problem or question, 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.

Continuous Versus Discrete Variables

In Chapter 2 we have defined 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. An example is chronological age. Some participants in a research study 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-olds—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. An example is a student's high school grade level. At a 4-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. A student can't be somewhere between two grade levels, such as in the "9.25th grade."

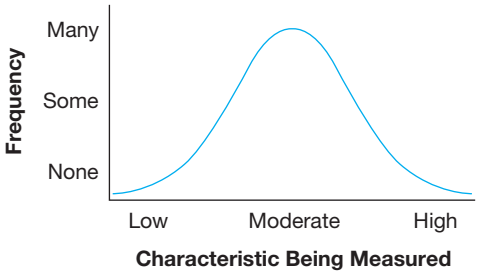
Nominal, Ordinal, Interval, and Ratio Data

In Chapter 4 we distinguished among four different scales of measurement; these scales, in turn, dictate how numerical data can be statistically analyzed 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 don't 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 party affiliation with numbers, perhaps using 1 for Republicans, 2 for Democrats, 3 for "Other affiliation," and 4 for "No affiliation." In neither case do the numbers indicate that participants have more or less of something; girls don't have more "gender" than boys, and Democrats don't have more "political party affiliation" than Republicans.
- *Ordinal data* are those for which the assigned numbers reflect an order or sequence. They indicate 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 doesn't 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're 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 doesn't mean that a person has no intelligence whatsoever.
- *Ratio data* are similar to interval data, in that they reflect equal intervals between values for the characteristic being measured. However, they also have a true zero point: A value of 0 tells us that there's a complete absence of the characteristic. An example is 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

Many theorists have proposed that characteristics of living populations (e.g., populations of maple trees, platypuses, human beings, or a certain subgroup of human beings) often 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:

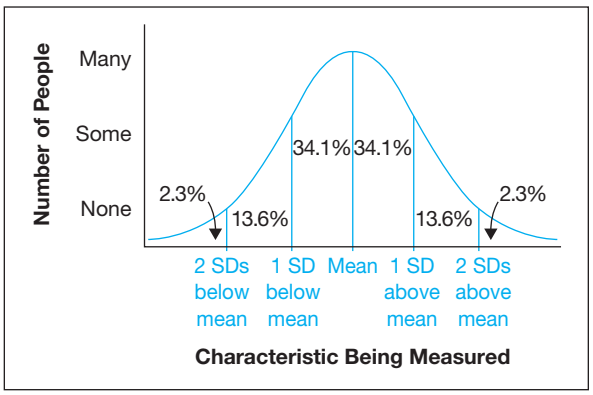
- ***It is horizontally symmetrical.*** One side is the mirror image of the other.
- ***Its highest point is at its midpoint.*** More people (or whatever other entities are the focus of investigation) are located in the exact middle than at any other point along the curve. In statistical terms, three 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 keywords "normal distribution table" in a search engine such as Google, Bing, or Yahoo!

To better understand the normal distribution, take any fortuitous happening and analyze its distribution pattern. For example, let's take corn production in the state of Iowa, which produces more corn than any of the other 49 U.S. states. If we could survey all Iowan corn farmers about their per-acre corn yield in a given year—the total population, or universe, of Iowa's corn farmers and cornfields that year—we would probably find that a few farmers had an unusually small corn crop per acre for no discernible reason except that "that's the way it happened." A few other farmers, for an equally unknown reason, may have had especially bountiful yields from their fields. However, most farmers would have had middle-of-the-road yields—some a little bit more than average, others a little bit less than average. The normal curve might describe Iowa's corn production. No one planned it this way; it's 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 reaches a peak. Gradually, with the passing of the storm, the number of flashes subsides. The normal curve might be 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's always bell-shaped. In any one situation, the *values* within it vary. The number at the peak of the curve doesn't always have the same value, and the overall shape may be more broadly spread or more compressed, depending on the situation.

FIGURE 11.6 Percentages Within Each Portion of the Normal Distribution



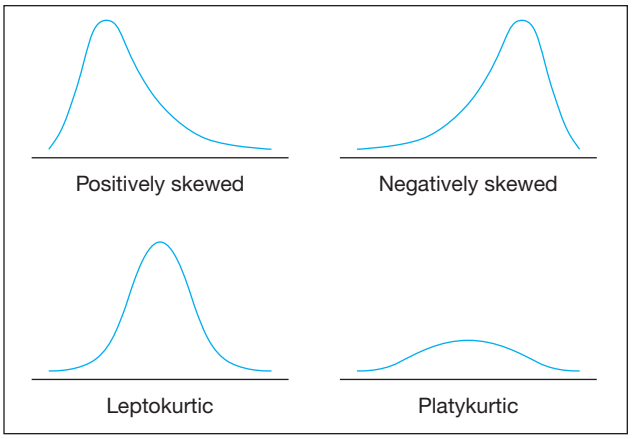


FIGURE 11.7 Common Departures from the Normal Distribution

Sometimes, however, data don't reflect a normal distribution. For instance, their distribution might be lopsided, or **skewed**; the "skew" is the part of the distribution that stretches out a bit to one side. If the peak lies to the left of the midpoint, the distribution is **positively skewed**; if the peak lies to the right of the 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 distribution** and an unusually flat one being a **platykurtic distribution** (see Figure 11.7).

Some data sets don't remotely resemble a normal distribution, not even a lopsided, pointy, or overly flattened variation of one. Ordinal data, by virtue of how they are created, *never* fall in a normal distribution. For instance, 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 also form a flat distribution rather than a bell-shaped curve. Percentile ranks—sometimes simply called **percentiles**—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 items correctly answered or number of points accumulated—that each person in the sample earns on an assessment 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're calculated, percentile ranks spread people evenly over the number of possible ranks one might get; for example, 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 don't 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 can be flawed.

In contrast, **nonparametric statistics** aren't 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 example, 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 or other credible source 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 then make appropriate adjustments for non-normal distributions.

MyLab Education Self-Check 11.2

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

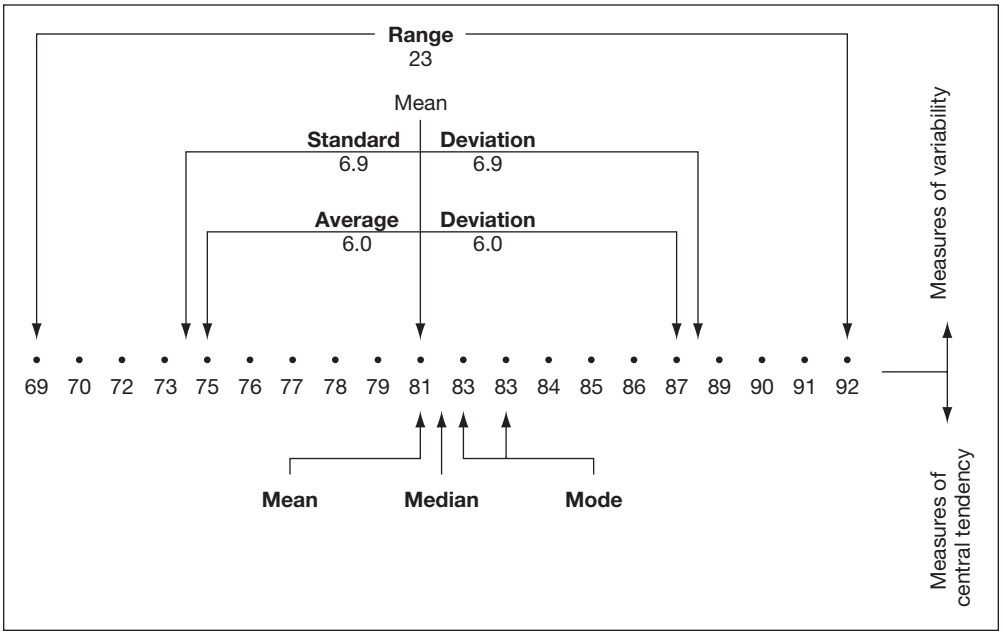
A *point of central tendency* is a point around which the data revolve, a middle number around which the data regarding a particular variable seem to hover. In statistical language, we use the term *measures of central tendency* to refer to techniques for identifying such a point. Three commonly used 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 previous list of Joe's grades for February (see Figure 11.2), we see that only one grade (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.

FIGURE 11.8
Measures of Central Tendency and Variability for Joe's Grades



The **median** is the numerical center of a set of scores; to simplify our discussion here, we will call each piece of data in the set a "score." The median is the number in the very middle of the scores, with exactly as many scores above it as below it. Recall that Joe's record has 20 grades for February. 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).

Meanwhile, the **mean** requires a mathematical calculation. More specifically, the mean is the *arithmetic average*[2] of scores for a certain variable. 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 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 find the mean as follows:

$$M = \frac{X_1 + X_2 + X_3 + \ldots + X_N}{N}$$

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:

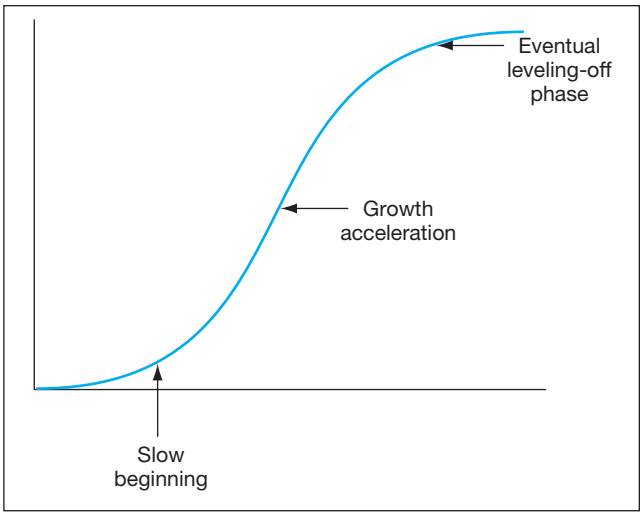
$$M = \frac{\Sigma X}{N}$$

Using the formula, we find that the mean for Joe's grades is 81, as shown in Figure 11.8. (The variation in Joe's grades, depicted in the figure as *measures of variability,* is discussed shortly.)

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 mathematically "balance." The mean is also 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.

[2]As previously noted in Chapter 1, when the word *arithmetic* is used as an adjective, it's pronounced with emphasis on the third syllable ("ar-ith-MET-ic").

FIGURE 11.9 Typical Growth Curve



The median is more appropriate for dealing with ordinal data. The median is also frequently used 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. Accordingly, 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[3]—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, as shown in 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,* Malthus (1826/1963) 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's an upper limit to what Mother Earth can produce in the way of food to sustain the population. Thus, many growth curves resemble a stretched-out *S,* as the curve in Figure 11.9 does.

If we are recording the growth of beanstalks in an agronomy laboratory, we don't find the average growth by assuming a normal distribution and calculating the arithmetic mean. The statistical technique doesn't fit the data. 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:

$$M_g = \sqrt[N]{(X_1)\,(X_2)\,(X_3)\,\ldots\,(X_N)}$$

[3]Again the emphasis in *arithmetic* is on the third syllable ("ar-ith-MET-ic").

TABLE 11.2 Using Measures of Central Tendency for Different Types of Data

Measure of Central Tendency	*How It Is Determined (N = number of scores)*	*Data for Which It Is Appropriate*
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 *N* is an odd number) or the midpoint between the two middle scores (when *N* 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 (*N*) 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 *N*th root of their product is computed.	• Data on ratio scales • Data that approximate an ogive curve (e.g., growth data)

For growth phenomena, we use the geometric mean because that's the way in which things grow and the way cells divide—geometrically.

Demographers, 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:

- Population growth
- Biological 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, they call for one measure of central tendency. If they have an ogive-curve nature (characteristic of a growth situation), they demand another measure. A *polymodal distribution*—one with several 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 statistic.

Thus, we must emphasize an essential rule for researchers who use statistics in their data analysis: *The nature of the data 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.

Measures of Central Tendency as Predictors

Sometimes a researcher uses a measure of central tendency as a rough estimate of the most likely outcome—essentially trying to answer the question *What is the best prediction?* As an example, imagine that you're walking down a relatively busy 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 might be an accident, a street fight, or a person giving away free chocolate 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're referring to those citizens, students, and wage earners who are huddled around the point of central tendency. In the broad spectrum of possibilities, we're 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 the data cluster around the point of central tendency, the greater the probability of making a correct guess about where any particular data point lies. As illustrated in Figure 11.10, data are more similar if they cluster close to 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."

Consider, for example, the case of Stephen Jay Gould, a renowned evolutionary biologist who in 1982 learned that he had abdominal mesothelioma, an especially lethal form of cancer. His prognosis was grim: Medical literature at the time reported a median life expectancy of 8 months following the diagnosis. Two years later, Gould—still very much alive—wrote a classic essay, "The Median Isn't the Message" (Gould, 1985), which you can easily find and read on the Internet. In fact, Gould lived for 20 years after his diagnosis, eventually succumbing to a very different form of cancer. Clearly the median was *not* a good predictor in his case.[4]

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 statistics.

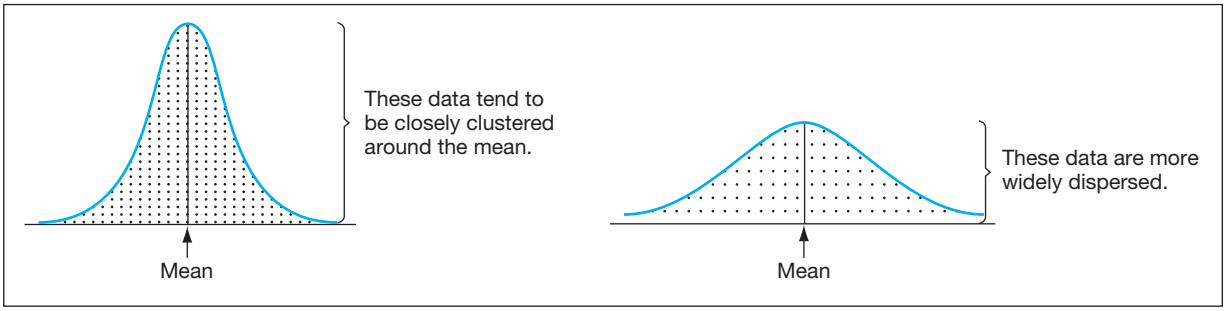


FIGURE 11.10 Distributions that Differ in Variability

[4]Here you can see an example of a positively skewed distribution. At most, people could die only 8 months earlier than the 8-month median time period—that is, at the very moment of diagnosis—but people might survive much, much longer than the 8-month median, as Gould did.

How Great Is the Spread?

The simplest, most straightforward indicator 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 can 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 10 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 something such as "Eighty percent of the families have from three to six children."

Other measures of variability use less extreme values as starting points. 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.[5] 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 using the median as a measure of central tendency should also consider the quartile deviation as a possible statistical measure for variability. When working with either ordinal data or highly skewed data, some researchers report what is called a **five-number summary**, which consists of the lowest and highest numbers (depicting the range) plus Quartile 1, the median (which is also Quartile 2), and Quartile 3.

Now, instead, let's use the mean as a starting point. Imagine that we determine how far away each score is from the mean. That is, we calculate the *difference* between each score and the mean (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 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 might be acceptable when no further statistical procedure is contemplated. Its primary weakness is in its use of absolute values. Imagine what would happen if, instead, we kept the plus and minus signs in the score–mean differences we calculated. 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 the

[5]If, instead of dividing the data into four 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.*

positive and negative deviations together, they would entirely counterbalance one another, essentially "canceling out" one another and yielding an overall sum of zero. The average deviation eliminates this problem by ignoring the pluses and minuses. Yet this is a rather dubious procedure. It is neither sound mathematics nor sound research practice to ignore what we don't like.

Unlike the formula for an average deviation, the formula for a **standard deviation** changes negatives to positives in a perfectly acceptable mathematical manner. More specifically, the formula requires *squaring* each of the score–mean differences, like so:

$$(X - M)^2$$

In arithmetic, whenever we multiply a negative number by itself, the result is a positive number. The same is true for squaring a positive number, of course; hence, there are no negative score–mean differences to counterbalance the positive differences.

To calculate a standard deviation (symbolized as either σ or s for a population parameter and sample statistic, respectively; see Table 11.1), we follow a procedure similar to that for calculating an average deviation. However, rather than getting absolute values of the score–mean differences, we *square* each of those differences. We then calculate the sum (Σ) of all the differences, divide that sum by the number of scores, and finally find the square root of the quotient just obtained. Thus, the formula for a standard deviation is as follows:

$$s = \sqrt{\frac{\Sigma (X - M)^2}{N}}$$

It's 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 (again look at Figure 11.8).

The standard deviation is the measure of variability most commonly used in statistical procedures. Many statistical procedures use a second, related measure of variability either in addition to or instead of the standard deviation. This statistic is known as the **variance**, which is simply the standard deviation squared:

$$s^2 = \frac{\Sigma (X - M)^2}{N}$$

Four of the measures of variability just described are summarized in Table 11.3. The table omits the five-number summary, which isn't a single statistic, and the average deviation, which is rarely used.

Using the Mean and Standard Deviation to Calculate Standard Scores

Earlier in the chapter we introduced the term *raw score,* the number of correct answers or points that a person gets on a test or other quantitative assessment instrument. Such scores typically need some context to 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 introverted or extroverted Mary is.

Sometimes researchers provide context by tying scores to a specific rating scale or rubric (see Chapter 6). On other occasions they convert raw scores to **norm-referenced scores**, scores that reflect where each person is positioned in the distribution relative to a group

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	$s = \sqrt{\frac{\Sigma(X - M)^2}{N}}$	• Data on interval and ratio scales • Most appropriate for normally distributed data
Variance	$s^2 = \frac{\Sigma(X - M)^2}{N}$	• 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.

of peers. This peer group, called a *norm group,* might be either the other participants in a research study or, instead, a nationwide group of individuals who have responded to the same assessment instrument.

You 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, we know that she's quite outgoing—more so than 95% of the people who have taken the test. But you should recall that percentile ranks have a definite limitation: They are ordinal data rather than interval data, and thus we can't 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 negative numbers—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:

$$\text{New standard score} = (z \times s_{new}) + M_{new}$$

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.[6] 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 doesn't 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 haven't 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!

[6]A 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 6.1 in Chapter 6), between emotional state and physical health, and 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, and 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, it shouldn't surprise you to learn that there's a positive correlation between personal conscientiousness and general achievement level: People who are internally driven to get things done tend to be higher achievers both at school and in the workplace (von Stumm, Hell, & Chamorro-Premuzic, 2011). In contrast, a negative number indicates an inverse relationship, or **negative correlation**: As one variable increases, the other variable *de*creases. For example, there's 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., Laursen, Bukowski, Aunola, & Nurmi, 2007).
- ***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 within a few days of each other: People tend to get very similar scores on both tests, especially if the 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 can't 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., Shavinina & Ferrari, 2004), 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

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., males vs. females, Republicans vs. Democrats).
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	$r_{12 \cdot 3}$	The relationship between two variables exists, in part, because of their relationships with a third variable, and the researcher wants to statistically "factor out" (i.e., eliminate) the effects of this third variable (e.g., what is the relationship between motivation and student achievement when IQ is held constant?).
Multiple correlation	$R_{12 \cdot 3}$	A single variable is correlated with two or more variables; here the researcher wants to compute the single variable's *combined* relationship with the others. This statistic is used in multiple regression.
		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., $N < 10$).

statistics as well. As is true for measures of both central tendency and variability, 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're apt to see this statistic used frequently in research reports, the reports rarely explain what it means. By *variance,* we are specifically referring to a particular 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, and vice versa. This knowledge will allow us to statistically control for—and thus *reduce*—some of the variability in our data through such statistical procedures as partial correlation and analysis of covariance (described in Table 11.4 and later in Table 11.5, respectively).

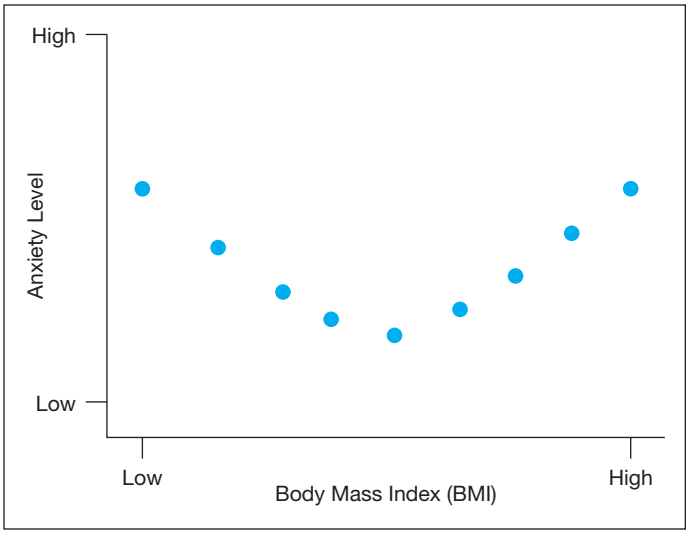


FIGURE 11.11
U-Shaped Relationship Between Body Mass Index (BMI) and Anxiety

Source: Based on Scott et al. (2008).

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 research study (Scott, McGee, Wells, & Oakley 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).

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 of Assessment Strategies Affect Correlation Coefficients

The extent to which you find a statistical correlation between two characteristics depends, in part, on how effectively you are assessing those characteristics. Even if there really *is* a correlation between two variables, you won't necessarily find one if the assessment techniques you're using have poor validity and reliability. For instance, you're less likely to find a correlation between age and reading level if the reading achievement test you use is neither a valid (accurate) nor reliable (consistent) measure of reading skills.

Over the years, we authors have had many students find disappointingly low correlation coefficients between two variables that they expected to be highly correlated. By looking at the correlation coefficient alone, you can't determine the reason for a low correlation any more than you 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 assess both characteristics with a reasonable degree of validity and reliability.* We refer you back to the section "Validity and Reliability in Assessment" in Chapter 4, where you can find strategies for determining and enhancing both of these essential qualities of sound assessment practices.

A Reminder About Correlations

Whenever you do 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 positive correlation between self-esteem and annual salary level, you can't necessarily conclude that people's self-esteem *influences* their salary level, nor can you assume the reverse—that earning a high salary directly *enhances* people's self-esteem. The words *influence* and *enhance* both imply a cause-and-effect relationship—something that a correlation between two variables doesn't necessarily reflect. Only experimental studies, such as those described in Chapter 7, 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 encountering 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 won't be able to answer this question.

INFERENTIAL STATISTICS

By definition, inferential statistics enable 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 don't have the space to venture too far into these areas; statistics textbooks can give you more detailed information. However, the general concepts and principles we discuss in the following sections can help you appreciate just how useful inferential statistics can be.

Estimating Population Parameters

Especially when we conduct quantitative 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 (e.g., the mean, or μ), variability (e.g., the standard deviation, or σ), and proportion (P). These values in the population compare with the M or $\overline{X}$, the s, and the p of the sample (again see Table 11.1).

A simple example can illustrate this idea of estimation. Imagine that a woman named Jane 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's too large, the arm will stick and refuse to budge. Several customers have complained that some of the pins are faulty. Jane 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, Jane 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 here is to determine population parameters on the basis of sample statistics. From the sample, Jane can estimate the mean and variability of the pin diameters and the proportion of acceptable pins within the pin universe, as represented by the symbols μ, σ, and P, respectively.

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 its selection has been *biased* in some way—the statistics we calculate may be poor reflections of the population from which the sample has been drawn.

An Example: Estimating a Population Mean

Now imagine that we want to estimate the average (mean) height of 10-year-old boys in the state of Iowa. 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.

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 data on the heights of 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 one has been randomly selected from the same population—will almost certainly yield slightly different estimates for that 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 variable that has been measured—the variable for which we've calculated all the means—for the overall population.[7]

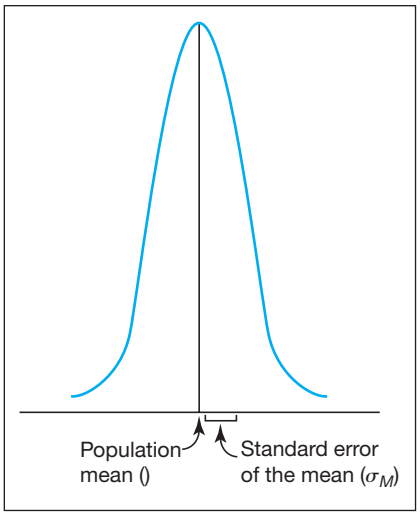
This situation is depicted in Figure 11.12.

The parameter just mentioned in Number 3 of the preceding list—the standard deviation for the distribution of sample means—is the **standard error of the mean**. This index tells us how much the particular mean we obtain 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 for any variable being measured is related to the standard deviation *for the variable itself* in the following way:

$$\sigma_M = \frac{\sigma}{\sqrt{n}}$$

[7]Taken together, these assumptions form the core of what statisticians refer to as the *central limit theorem.*

FIGURE 11.12
Distribution of Sample Means



Here we are faced at once with a problem. The formula just presented involves using the population standard deviation (σ), but the purpose of using a sample is to *avoid* having to assess 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 (σ_M) is directly related to the standard deviation of the characteristic being measured (σ or *s*): More variability in the population leads to a larger standard error of the mean—that is, it leads to greater variability in the sample means 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 Jane has selected a random 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're 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, Jane 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 Jane has done is to determine that the standard error of the mean is 0.002 (see the previously presented formula for estimated σ_M). Jane knows that sample means fall in a normal distribution (look again at Figure 11.12). She also knows that normal distributions have predictable proportions within each section of the curve (look again at Figure 11.6). In particular, Jane 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. For more information and guidance, we urge you to consult one or more statistics textbooks, such as those listed in the "For Further Reading" section at the end of the chapter.

Testing Hypotheses

It is time to return to a distinction previously made in Chapter 2: the distinction between research hypotheses and null hypotheses. Our discussions of hypotheses thus far have focused almost exclusively on research hypotheses—those tentative explanations that researchers put forth either for phenomena they propose to investigate or for data they have just collected and analyzed. In contrast, a **null hypothesis** (often symbolized as H_0) postulates that any observed pattern in a data set is due to chance alone—to the luck of the draw, so to speak. For instance, if we were to compare the average performances of two groups of people—and more specifically, if we were to compare the two group *means* for performance—our null hypothesis would be that both groups are parts of the same population and that any difference between the group means is 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 entirely to chance. If, for example, we find that this big of a difference between two group means would occur by chance *only one time in a thousand,* we could reasonably conclude that the group-mean difference is *not* due to chance—that, instead, something in the situation we are studying (perhaps an experimental treatment administered to only one of the groups) is systematically leading to a difference in the means. This process of comparing the results we've actually obtained with the results we would expect from chance alone is called *testing the null hypothesis.*

At what point does a researcher decide that a result has *not* occurred by chance alone? One commonly used 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 a certain 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 ($p < .01$)." 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 $(p = .003)$." 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 roundabout 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 is often our *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: The two groups are *different.* This is, to be sure, a backdoor approach to finding evidence for a research hypothesis, yet it's the approach most often taken in statistical analyses. The reasons are too complex to be dealt with in a text such as this one. Suffice it to say that it's mathematically much easier to test a hypothesis that an equivalence exists than to test a hypothesis that a difference exists.

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 isn't 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 thus 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 sign" that certain dynamics are operating within the data and merit attention.

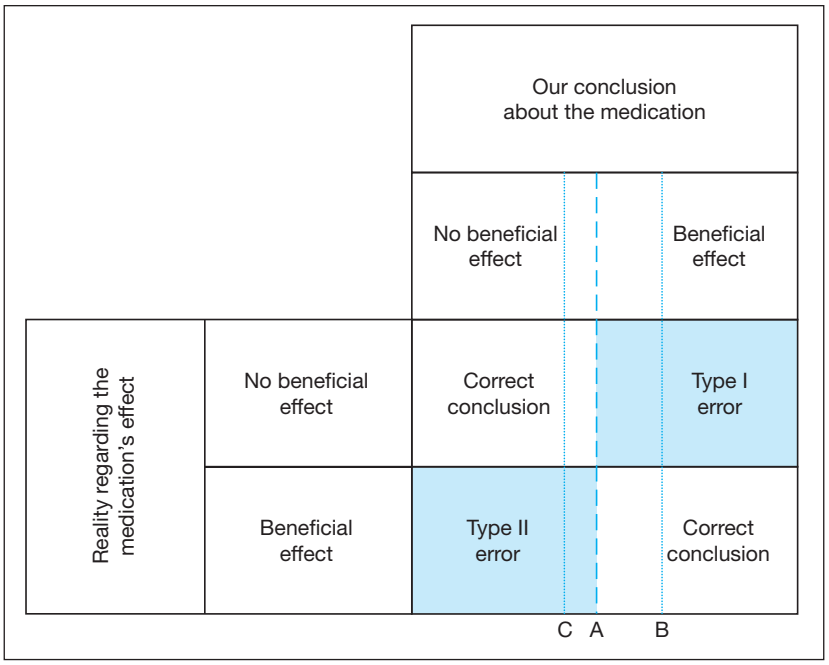
Making Errors in Hypothesis Testing

It's possible, of course, that we make a mistake when we decide that a particular result isn't the result of chance alone. In fact, *any* result might 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 wasn't 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're testing the relative effects of a new medication versus the effects of a placebo in reducing low-density lipoprotein cholesterol (the "bad" kind of cholesterol) in people's blood. 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 discover 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's 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's 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.

FIGURE 11.13 The Trade-off Between Type I and Type II Errors



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 *doesn't* reduce cholesterol.
3. We mistakenly conclude that it's 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 won't 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.

For a novice researcher, it can be extremely frustrating to get insignificant results—those that, from a statistical standpoint, could have been due to chance alone. Following are four 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.[8]

[8]Formulas exist for computing the power of statistical procedures for varying sample sizes. For example, see Lipsey (1990) or Murphy, Myors, and Wolach (2009).

- ***Maximize the validity and reliability of your assessment instruments.*** Assessment instruments in a research study rarely have perfect (100%) validity and reliability, but some instruments are more valid and reliable than others. Research projects that use instruments or other assessment strategies with high validity and reliability are more likely to yield statistically significant results. (Again we refer you to the section "Validity and Reliability in Assessment" in Chapter 4.)
- ***If logically defensible and logistically practical, obtain repeated measures of your dependent variable.*** Other things being equal, repeated measures of a certain behavior or characteristic for every participant enhance the power of statistical analyses. For example, experience-sampling methods (ESM; see Chapter 6) and experimental designs with one or more within-subjects independent variables (see Chapter 7) lend themselves well to statistically powerful analyses. Repeated measures enhance statistical power by (a) improving the *reliability* of overall assessment scores for each participant and/or (b) statistically controlling for between-subjects differences as possible sources of variability in assessment scores.
- ***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 a null hypothesis. When characteristics of the data meet the assumptions for parametric statistics, then, we urge you to use these statistics.

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 one time in every 20. And now let's say that as we analyze our data, we perform 15 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. In this case of testing 15 hypotheses at a .05 significance level, the probability of making at least one Type I error is .5367—in other words, better than a 50-50 chance![9]

Desperate for findings that might warrant a journal article or other noteworthy research report, some researchers sift through their data in search of a few statistically significant results amid many more *non*significant results. Such a practice, which is known by such terms as *data dredging* and *p-hacking,* is widely disparaged, in large part because the chance of making a Type I error—and thus of reporting an essentially bogus finding—is so high (e.g., see DeCoster, Sparks, Sparks, Sparks, & Sparks, 2015; Ledgerwood, 2016). Once again, we authors repeat that all-important rule: *Absolute honesty and integrity are assumed in every statement a scholar makes.*

Examples of Statistical Techniques for Testing Hypotheses

Table 11.5 lists many commonly used parametric and nonparametric statistical techniques for testing hypotheses. We authors 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 sophisticated analysis (e.g., multiple regression or structural equation modeling), parametric statistical procedures—with certain underlying assumptions about the nature of your data—are your only viable option.

[9] 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.

TABLE 11.5 Examples of Inferential Statistical Procedures and Their Purposes

Statistical Procedure	*Purpose*
	Parametric Statistics
Student's *t*-test	To determine whether a statistically significant difference exists between two means. A *t*-test takes slightly different forms depending on whether the two means come from separate, independent groups (an *independent-samples* t-test) or, instead, from a single group or two interrelated groups (a *dependent-samples* t-*test*). *Note:* In 1908, William Sealy Gosset published an explanation of this procedure under the pseudonym "Student"—hence the name "Student's *t*-test" (Student, 1908).
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 *t*-tests, ANOVAs take slightly different forms for separate, independent groups and for a single group; in the latter case, a *repeated-measures* ANOVA is called for. If an ANOVA yields a significant result (i.e., a significant value for *F*), the researcher should follow up by comparing various pairs of means using a *post hoc comparison of means.*
Multiple analysis of variance (MANOVA)	To simultaneously examine differences among means for two or more dependent variables rather than for the single dependent variable that characterizes *t*-tests and ANOVAs. A MANOVA takes into account any correlations that exist among the various dependent variables and, as a result, can enhance the statistical power of the analysis.
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 *covariate*). This technique can be statistically more powerful than ANOVA (i.e., it decreases the probability of a Type II error).
t-test for a correlation coefficient	To determine whether a Pearson product moment correlation coefficient (*r*) is larger than would be expected from chance alone.
Regression	To examine how accurately one or more variables enable predictions regarding the values of another (dependent) variable. A *simple linear regression* generates an equation in which a single independent variable yields predictions for the dependent variable. A *multiple linear regression* yields an equation in which two or more independent variables are used to predict the dependent variable. The researcher must keep in mind, however, that an independent variable's accuracy in predicting a correlated dependent variable does *not* necessarily indicate a cause-and-effect relationship.
Factor analysis	To examine the correlations among a number of variables and identify clusters of highly interrelated variables that reflect general underlying categories, or *factors,* 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 (*paths*) among the variables. SEM encompasses such techniques as *path analysis* and *confirmatory analysis* and is typically used to test a previously hypothesized model of how variables are causally interrelated. SEM enables a researcher to identify a *mediator* 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 *moderator* 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 *correlational* in nature; thus, any conclusions about cause-and-effect relationships are speculative at best.
	Nonparametric Statistics
Mann–Whitney *U*	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 *t*-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 *t*-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. nonsmokers and presence vs. absence of heart disease) are significantly correlated. This is one nonparametric alternative to a *t*-test for Pearson's *r.*
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 *t*-test for Pearson's *r.*

We urge you to consult one or more statistics textbooks 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 conduct defensible 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 the results together into a 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 have yielded which conclusions, which studies have contradicted other studies, and what methodological differences might have accounted for inconsistencies in findings. 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 doesn't choose arbitrarily among studies that have been reported about the research problem. Instead, the researcher uses some systematic and far-reaching approach (e.g., searching in several prespecified professional journals, using certain keywords in a search of online databases) to identify studies that have addressed the topic of interest.
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. The researcher may further restrict the chosen studies to those that involve particular populations, settings, assessment instruments, or other factors that might impact a study's outcome.
3. ***Converts each study's results to a common statistical index.*** Previous researchers haven't necessarily used the same 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 (ES)** is calculated for each study; that is, the researcher determines how much of a difference each study's intervention makes (in terms of standard deviation units) relative to a control group or other comparison group. The effect sizes of all the studies are then used to compute an average effect size for that intervention.[10]

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 make important contributions 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.

[10]Increasingly, researchers are including effect sizes in their reports of single research studies. Procedures for calculating effect sizes differ somewhat depending on the circumstances (e.g., see Cumming, 2014; Fidler & Cumming, 2013).

MyLab Education Self-Check 11.3

MyLab Education Application Exercise 11.2: Finding and Estimating Central Tendency

MyLab Education Application Exercise 11.3: Reducing Type I and II Errors

USING STATISTICAL SOFTWARE PACKAGES

We have already mentioned that general-purpose spreadsheet programs (e.g., Microsoft Excel) 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 many statistical software packages now widely available for use on personal computers. Some of these (e.g., SPSS, SAS, Stata, SYSTAT, Minitab, Statistica) are commercially available; many universities purchase them for use by all students and faculty members. Other packages (e.g., AdaMSoft, R) are *freeware* that virtually anyone can download from the Internet at no cost (to find R on the Internet, search for "R Project").

Statistical software packages typically have several advantages:

- ***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.
- ***User-friendliness.*** As statistical software programs become increasingly comprehensive, 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. However, selections of the proper statistics and interpretations of the results are still left to the researcher.
- ***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.
- ***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.

For researchers 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 EasyCalculation (easycalculation.com) and QuickCalcs (graphpad.com/quickcalcs). An Internet search for "online statistics calculator" can identify additional helpful websites.

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 can't make logical, theoretical, or pragmatic sense of what your analyses have revealed, then all your efforts have been for naught. Ultimately *you* must be in control of your analyses; you must know what calculations are being performed and why. Furthermore, you must have an intimate knowledge of the data you have collected; otherwise, you won't be in any position to detect what is almost certainly an absurd, nonsensical result—say, a mean of 106 for a set of scores that range between 20 and 45.

INTERPRETING THE DATA

Subjecting data to elegant statistical analyses 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.

Certainly statistical analyses are invaluable tools that enable us to find patterns in our data and thus help us detect possible dynamics working within those data. But we must never stop with the numerical values that statistical analyses yield. We must also *interpret* those values and give them meaning. The latter process includes the former, but the two should never be confused.

At the beginning of the chapter, we presented a hypothetical data set for 11 schoolchildren and discovered that the five girls in the sample had higher reading achievement test scores than the six 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 the girls' scores all 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 unique to the nautilus, or are there other similar relationships in nature?

Knowledge springs from questions like these. But we must be careful not to make snap judgments about the data we've collected. It's 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 aren't 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. On the following day, the same test given to another 11 children might yield different results.

In general, interpreting the results of quantitative analyses means several things:

1. ***Relating the findings to the overall research problem or question, as well as to specific subproblems 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 the phenomenon under investigation.
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 at hand. 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. 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 might 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 undesirable 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 human minds—including those of the researcher, practitioners in the field of medicine, and individuals with unhealthy cholesterol levels—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.

PRACTICAL APPLICATION Analyzing and Interpreting Data in a Quantitative Study

You can gain a clearer understanding of various 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

Choosing Statistical Procedures

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 statistic? _____

_____ Calculate variability? If so, with which statistic? _____

_____ Calculate correlation? If so, with which statistic? _____

_____ 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 statistically analyzed your data, what information will you then have?

_____

_____

_____ 6. From a research standpoint, what will your interpretation of the data entail? How will the statistical analyses help you address some or all of your research problem?

_____

_____

_____ 7. Will some of your data be inappropriate for statistical analysis? If so, how might you analyze and interpret them?

_____

_____

MyLab Education Self-Check 11.4

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 (K. J. Mitchell, 1998). Mitchell 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 excerpts refer 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 Mediational) 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.

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.)

Comments

The author capitalizes the names of the three theories. Sometimes, instead, 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 provide conceptual frameworks *that reflect how some variables may influence other variables, perhaps directly or perhaps indirectly through additional,* mediator *variables.*

The term 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 of the three theoretical frameworks on which the author bases her study.*

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.*

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.

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.

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.

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.

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 (related to) alcohol use, drug use, and eating. In turn, these cognitive factors interact to predict health-risk behavior.

(The author continues with additional conclusions, 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 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. . . .

The author's use of the term cross-sectional *is somewhat different from our use of it in Chapter 6. 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.*

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 the accuracy and dependability (i.e., validity) of the results, rather than to reliability as we have previously defined the term.*

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.

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.

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 for 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. . . .

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.

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.

By perception differences, *the author is presumably referring to how participants may have idiosyncratically 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.*

(The researcher concludes the discussion by talking about potential implications of her findings for clinical practice and social policy.)

Note: 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.

APPLICATION EXERCISES

Similar exercises are available on Pearson MyLab Education.

1. Henry is interested in how children learn to talk about their emotions. He knows that parents use a lot of different emotion words when they talk to their children about daily activities. He also knows that many parents read to their children regularly, and he wonders if there are certain characteristics of children's picture books that facilitate learning about emotions. He thus decides to conduct a study in which he first makes a list of words that describe emotions that a storybook character might be feeling, then gathers a large sample of books and marks each use of a word on the list. Review his data and answer the questions that follow.

afraid	59
angrier	2
angry	56
fearful	2
frightened	32
glad	39
happier	3
happiest	3
happy	109
joyful	1
joyous	2
mad	30
mournful	1
sad	58
sadder	1
scared	28

terrified	9
unhappy	10
Total number of different word types used	18
Total number of words (tokens)	445
Type/token ratio	0.04

i. Henry has a list of the words he found in the storybooks (18 different words) and the number of times each word was found (known as the "count"). The words are organized in alphabetical order. Reorganize these data to make them easier to think about and interpret, and put them in a two-dimensional table. What are the benefits and the drawbacks of the organization you chose?

ii. Would it make sense for Henry to graph these data? Explain why or why not.

2. IQ is a standard score, and IQ scores are normally distributed, with $\mu = 100$ and $\sigma = 15$.

   i. What does it mean to say that IQ scores are normally distributed? What are μ and σ?

   ii. What percentage of the population has an IQ over 130?

   iii. Imagine that you are a psychometrician, and your job is to administer and score IQ tests. The 15 individuals comprising your sample come in for testing on a Monday. Their scores are as follows: 123, 98, 103, 94, 136, 118, 126, 153, 114, 99, 121, 125, 139, 107, 126. Find the sample mean, median, and mode.

   iv. Compare your sample statistics to the population following parameters provided in the table. Use the concept of *error* to explain why you cannot accurately estimate the population parameters using only your sample.

	population	sample
mean	$\mu = 100$	$M = 118.8$
standard deviation	$\sigma = 15$	$s = 16.67$

3. Review the definitions of Type I and Type II errors, and answer the following questions.

   i. Jason works for a company that develops and tests new drugs that the company hopes will treat or cure medical problems. The company is working on a new drug that company leaders believe will be effective against leukemia: people's lives will be saved, and the company will make money. Which type of error, Type I or Type II, should the company researchers most vigilantly try to avoid when testing the effectiveness of the drug?

   ii. At the same time, the company researchers should also try to avoid the other type of error. Explain why.

   iii. This scenario clearly shows the importance of balancing the possibility of Type I errors against the possibility of Type II errors. In this scenario, what are the two best strategies for achieving this balance? Explain why.

FOR FURTHER READING

Agresti, A., & Franklin, C. A. (2013). *Statistics: The art and science of learning from data* (3rd ed.). Hoboken NJ: Pearson.

Azen, R., & Walker, C. M. (2011). *Categorical data analysis for the behavioral and social sciences.* New York, NY: Routledge.

Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2013). *Analyzing social networks.* London, England: Sage.

Card, N. A. (2012). *Applied meta-analysis for social science research.* New York, NY: Guilford Press.

Coladarci, T., & Cobb, C. D. (2014). *Fundamentals of statistical reasoning in education* (4th ed.). New York, NY: Wiley.

Coolidge, F. L. (2013). *Statistics: A gentle introduction* (3rd ed.). Los Angeles, CA: Sage.

Cooper, H. (2016). *Research synthesis and meta-analysis: A step-by-step approach* (5th ed.). Los Angeles, CA: Sage.

Cumming, G. (2011). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis.* New York, NY: Routledge.

Cumming, G., & Finch, S. (2005). Inference by eye: Confidence intervals and how to read pictures of data. *American Psychologist, 60,* 170–180.

Fennessey, J. (1968). The general linear model: A new perspective on some familiar topics. *American Journal of Sociology, 74,* 1–27.

Fidler, F., & Cumming, G. (2013). Effect size estimation and confidence intervals. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology: Vol. 2. Research methods in psychology* (2nd ed., pp. 142–163). Hoboken, NJ: Wiley.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Los Angeles, CA: Sage.

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2011). *Applied longitudinal analysis* (2nd ed.). Hoboken, NJ: Wiley.

Fowler, F. J., Jr. (2014). *Survey research methods* (5th ed.). Los Angeles, CA: Sage. [See Chapter 10.]

Gast, L., & Ledford, J. R. (Eds). (2014). *Single case research methodology: Applications in special education and behavioral sciences* (2nd ed.). New York, NY: Taylor & Francis.

Gonzalez, R. (2009). *Data analysis for experimental design.* New York, NY: Guilford Press.

Gravetter, F. J., & Wallnau, L. B. (2014). *Essentials of statistics for the behavioral sciences* (8th ed.). Belmont, CA: Wadsworth/Cengage.

Grissom, R. J., & Kim, J. J. (2012). *Effect sizes for research: Univariate and multivariate applications* (2nd ed.). New York, NY: Routledge.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis.* New York, NY: Guilford Press.

Heiman, G. W. (2014). *Basic statistics for the behavioral sciences* (7th ed.). Boston, MA: Cengage.

Jose, P. E. (2013). *Doing statistical mediation and moderation.* New York, NY: Guilford Press.

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York, NY: Guilford Press.

Knapp, H. (2016). *Introductory statistics using SPSS* (2nd ed.). Los Angeles, CA: Sage.

Kranzler, J. H. (2011). *Statistics for the terrified* (5th ed.). Hoboken, NJ: Pearson.

Kraska-Miller, M. (2013). *Nonparametric statistics for social and behavioral sciences.* New York, NY: Routledge.

Lind, D. A., Marchal, W. G., & Wathen, S. A. (2013). *Basic statistics for business and economics* (8th ed.). New York, NY: McGraw-Hill.

McGrath, R. E. (2015). *Creating and verifying data sets with Excel.* Los Angeles, CA: Sage.

Montgomery, D. C. (2013). *Design and analysis of experiments* (8th ed.). New York, NY: Wiley.

Phillemer, D. B. (1994). One- versus two-tailed hypothesis tests in contemporary educational research. *Educational Researcher, 20*(9), 13–17.

Prell, C. (2012). *Social network analysis: History, theory, and methodology.* London, England: Sage.

Rosner, B. (2016). *Fundamentals of biostatistics* (8th ed.). Boston, MA: Cengage.

Rowntree, D. (2004). *Statistics without tears: A primer for non-mathematicians*. Hoboken, NJ: Pearson.

Salkind, N. J. (2016). *Excel statistics: A quick guide* (3rd ed.). Los Angeles, CA: Sage.

Salkind, N. J. (2016). *Statistics for people who (think they) hate statistics* (6th ed.). Los Angeles, CA: Sage.

Schmidt, F. L, & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Los Angeles, CA: Sage.

Shadish, W. R. (2014). Statistical analyses of single-case designs: The shape of things to come. *Current Directions in Psychological Science, 23,* 139–146.

Silver, N. (2012). *The signal and the noise: Why so many predictions fail—But some don't.* New York, NY: Penguin.

Sweet, S. A., & Grace-Martin, K. (2012). *Data analysis with SPSS: A first course in applied statistics* (4th ed.). Hoboken, NJ: Pearson.

Tay, L., Parrigon, S., Huang, Q., & LeBreton, J. M. (2016). Graphical descriptives: A way to improve data transparency and methodological rigor in psychology. *Perspectives on Psychological Science, 11,* 692–701.

Terrell, S. R. (2012). *Statistics translated: A step-by-step guide to analyzing and interpreting data.* New York, NY: Guilford Press.

Vogt, W. P., & Johnson, R. B. (2016). *The Sage dictionary of statistics and methodology: A nontechnical guide for the social sciences* (5th ed.). Los Angeles, CA: Sage.

Wagner, W. E. (2016). *Using IBM® SPSS® for research methods and social science statistics* (6th ed.). Los Angeles, CA: Sage.

Wheelan, C. (2013). *Naked statistics: Stripping the dread from the data.* New York, NY: Norton.

This page intentionally left blank

Chapter 12 Analyzing Qualitative and Mixed-Methods Data

In qualitative research, we closely examine the data to find the meanings that lie within them. In most qualitative methodologies, data interpretation begins almost immediately, and initial interpretations are apt to drive subsequent data collection. As we engage in such an iterative data-collection-and-interpretation process, we must remember that although the human mind is capable of amazing insights, it can be easily influenced by expectations and biases that prejudice our understandings.

Learning Outcomes

12.1 Describe general strategies you might effectively apply when organizing and analyzing qualitative data.

12.2 Explain how qualitative data analysis is typically an iterative, spiral-like process.

12.3 Explain what the term *researcher-as-instrument* means in qualitative research, and discuss its implications for data analysis.

12.4 Identify appropriate and efficient approaches to analyzing data in a mixed-methods study.

12.5 Explain how you might conduct a systematic review of qualitative and mixed-methods studies regarding a particular topic.

In the quantitative designs described in Chapter 6 and Chapter 7, data collection, data analysis, and data interpretation are, in large part, three separate steps in a research project. In most qualitative designs, however, they are closely intertwined. Not only do data analysis and data interpretation often go hand in hand, but they also begin early in data collection and may subsequently drive further data collection. For instance, data collected early in a qualitative study might suggest certain patterns or dynamics in the phenomenon being investigated, leading the researcher to seek additional information in an attempt to confirm, clarify, or disconfirm the hypothesized patterns or dynamics. Thus, in qualitative studies, analysis is an iterative process in which researchers move back and forth among data collection and data analysis/interpretation, with initial analyses and interpretations driving later data collection.

Strategies for analyzing qualitative data are less prescriptive than those for analyzing quantitative data. Also, they tend to rely heavily on *inductive* reasoning processes: The researcher observes a few specific situations or events and, from them, imposes specific meanings on them—often by *coding* them in some way—and then draws conclusions about a more general state of affairs. Such flexibility and open-endedness in data analysis strategies certainly have benefits—they can yield insights that might not come to light any other way—but they make it extremely difficult for a researcher to analyze data with total objectivity. Any qualitative researcher must continually acknowledge, both to self and to others, that personal attitudes and opinions are inevitably creeping into and biasing observations and interpretations.

No matter how you proceed, your data analysis in a qualitative study is apt to be a complex, time-consuming process. You must wade through a great deal of information, some of which will be useful and some of which will not. Furthermore, the data you obtain are apt to be multifaceted and may simultaneously reflect several distinct layers of meaning. This is not to say, however, that the process of analyzing qualitative data is dull and tedious. On the contrary, it can be stimulating, challenging, illuminating, and quite enjoyable—a very personally rewarding enterprise.

QUALITATIVE ANALYSIS STRATEGIES

Experienced qualitative researchers have offered many strategies for effectively analyzing and interpreting qualitative data, and the particular strategies employed must depend, in large part, on the nature of the research problem and the types of data for which the problem calls. In the upcoming pages we offer a few general considerations and suggestions—enough, we hope, to guide you as you plan a qualitative research project. For more explicit strategies, we urge you to consult one or more of the sources listed in the "For Further Reading" section at the end of the chapter.

General Strategies for Organizing and Analyzing Qualitative Data

Following are general strategies that might apply to virtually any qualitative study. They are meant—only very loosely—to be executed in the order we list them here. But please remember that many qualitative research studies involve an iterative process in which a researcher goes back and forth among data collection, analysis, and interpretation. You, too, should go back and forth among the strategies as needed.

1. ***Convert the data into one or more forms that will be easy to organize and analyze.*** For example, transcribe audiotaped interviews. Put handwritten field notes into word-processing documents or a spreadsheet. Electronically scan photocopied documents.

2. ***Organize the data in a preliminary, superficial way that will enable you to locate them easily as you proceed.*** For example, in a case study, this initial organizational scheme might involve putting notes about various incidents and events in chronological order. In an ethnographic study, it might involve sorting electronic documents into desktop e-folders with such labels as *Field Observations, Interviews,* and *Artifacts.* In a content analysis, initial organization might involve separating the entities to be analyzed according to a variable central to the research problem—perhaps (in a study of historical trends) the chronological time periods in which different items were created or (in the field of child development) the ages of the children who created various writing samples, drawings, or other artifacts.

You might think of Steps 1 and 2 as "getting your ducks in a row"—a commonly used idiom that essentially means gathering and arranging all needed materials before beginning an activity. With all the materials close at hand and in readily usable forms, you can more efficiently forge ahead and tackle the many mental challenges that your data analysis will ultimately involve.

3. ***Identify preliminary categories that are likely to be helpful in coding the data.*** One of the biggest challenges in qualitative data analysis is to determine how best to organize it *meaningfully*—something the preceding ducks-in-a-row strategy doesn't fully accomplish. Oftentimes the first step in the meaning-making process is to identify a list of potentially helpful ways of categorizing and coding the data. For example, a researcher studying the course of a political campaign might think in terms of "campaign strategies," "fund-raising activities," "news media accounts," "setbacks," and the like.

In some cases a researcher begins with a *start list*—a predetermined list of categories or themes derived from the research problem and its subproblems or, instead, from a particular theoretical or conceptual framework. In other cases the researcher peruses the collected data in search of general themes that seem to "pop out" as being important considerations in the phenomenon under investigation. For example, a beginning list of codes might include codes related to some or all of the following:

- Specific topics
- Characteristics and attributes
- Actions
- Processes
- Emotions
- Beliefs
- Values
- Evaluations

Typically the researcher also identifies subcategories—*subcodes*—for some or all of the codes. For example, under a general code *Emotions,* a researcher might include such subcodes as *Joy, Anxiety,* and *Depression.* And sometimes researchers pull certain codes or subcodes directly from participants' own language and terminology; such codes are called *in vivo* (i.e., "living") codes.

Lists of codes and subcodes are rarely meant to be mutually exclusive. In many studies, data might be simultaneously coded in two or more ways. Whatever approach you use to analyze your data—and we strongly recommend using one of the computer software programs we mention later in the chapter—your approach should readily accommodate multiple and possibly overlapping cycles of coding your data.

It's important to note, however, that not all qualitative researchers specify preliminary categories in advance. Some researchers solely employ inductive processes in their data analyses: They refrain from identifying any codes at the beginning of the coding process, opting instead to engage in a strategy called **open coding**. That is, they initially pass through the data in an open-minded search for meanings in an effort to allow codes to emerge from and be strongly grounded in the data. Such researchers believe that any identification of codes and labels before the formal coding process might lead to overly prescriptive and deductive analyses of data and thus might prevent them from identifying unexpected beliefs and perspectives. In some cases, prespecified codes could cause them to look only for evidence that supports their preconceived assumptions and working hypotheses. Open coding will be discussed in more detail later in the chapter, as it is a predominant strategy used in grounded theory studies.

4. ***Divide the data into meaningful units that will be individually coded.*** In some qualitative studies the meaningful units to be coded are clear-cut; for example, this might be true when the objects of study are paintings, television commercials, or specific movements in classical music. But especially in verbal materials, such as interview data or lengthy written works, the data need to be systematically broken into small segments—perhaps individual phrases or sentences, or paragraphs—that will be coded separately.

5. ***Apply the initial coding scheme to a subset of the data.*** In a sense, this strategy involves pilot testing the list of codes to determine whether it will adequately capture the multidimensional meanings that the data hold. Such a pilot test is likely to reveal certain weaknesses in the list. For example, initial codes might:

- Be too vague to enable consistent, reliable categorization of the data
- Reflect overlapping ideas (and thus might be combined into a single code)

- Be too item-specific to be of use in making generalizations about the data
- Omit certain themes that appear to be important in the overall data set

This might also be a good time to identify parts of the data set that are, for all intents and purposes, irrelevant to the research problem and can be omitted from future analyses.

This process of refining and strengthening codes is sometimes referred to as *coding clean-up* and often involves deleting, merging, and renaming codes in order to reduce the number of codes and organize them. Initial coding cycles can produce a large number of codes, and such clean-up helps researchers develop and organize a manageable list of codes to apply to data.

6. ***Construct a final list of codes and subcodes, and define each code and subcode as specifically and concretely as possible.*** The goal here is to create a list that ensures reasonable consistency in coding the data. The list should be definitive enough that two different researchers would be highly likely to code any single item in the same way. One common strategy for enhancing coding consistency is to include specific examples (e.g., quotes from participant interviews) to illustrate each code and subcode. As an illustration, let's return once again to Christy Leung's dissertation study, a project briefly mentioned in Chapter 2 and subsequently revisited in Chapter 8. Leung's research questions centered around Chinese mothers, their experiences in immigrating to the United States, and their parenting practices. Her start list and final list of codes included many categories and subcategories related to immigration and parenting. Figure 12.1 presents short excerpts from the two lists—in particular, codes related to the interviewee's negative and positive immigration experiences. Notice that the start list includes only general ideas to be coded, whereas the final list includes a definition and illustrative quote for each code. Notice, too, that the final list omits some codes in the start list (e.g., "Financial hardship"), substantially revises other codes, and includes new topics (e.g., "Instrumental support") that presumably emerged during Leung's early inspection of the data. Also, the category "Positive Immigration Experiences" in the final list includes a subcategory, "Perceived Support at the Initial Stage of Migration and/or Throughout Their Transition."

As you can see in Figure 12.1, Leung used numbers and number–letter combinations to label the various categories in her start list. For example, the number "I-3(1)" near the top of the leftmost column essentially means that "racial discrimination" is the first subcategory under the more general category "negative immigration experiences"; the corresponding code "I-NIE-RAC" in the rightmost column is short for "Category 1, negative immigration experiences, subcategory racial discrimination." Later, in her final list, Leung eliminates the numbers but continues to use letters to designate general categories and their subcategories. Such abbreviations presumably enabled the coders to assign codes more quickly than would be the case otherwise. However, an obvious downside to using numbers and a category–subcategory hierarchy at the very beginning of the coding process is that it imposes a predetermined "order" on the data that may not accurately characterize those data. Some researchers (e.g., see Saldaña, 2013) recommend against using such numbering systems and abbreviations, viewing them as unnecessary and actually encumbering the decoding/encoding process.

7. ***Consider using two or more raters to code the data independently.*** When only a single researcher codes the data, it's all too easy for the researcher's expectations and biases to influence the codes assigned to each piece of data. To minimize such an influence, the researcher might, instead, have two or more people independently code the data. Here we are talking about **interrater reliability**: the extent to which two or more individuals provide identical judgments. When two individuals consistently rate the great majority of the data in similar ways, a researcher not only documents the reliability of the coding scheme but can also enhance the credibility of the results obtained.

FIGURE 12.1 Coding the Interviews of Chinese Immigrant Women: Excerpts from a Start List and a Final List of Codes and Subcodes

EXCERPTS FROM START LIST:

Number	Descriptive Label	Code
I-3	Negative Immigration Experiences	I-NIE
I-3(1)	Racial discrimination	I-NIE-RAC
I-3(2)	Language barriers	I-NIE-LAN
I-3(3)	Financial hardship	I-NIE-FIN
I-3(4)	Job limitations	I-NIE-EMP
I-3(5)	Lack of social support/networks	I-NIE-SSN
I-3(6)	Cultural shock	I-NIE-CUL
I-4	Positive Immigration Experiences	I-PIE
I-4(1)	Strong sense of family	I-PIE-FAM
I-4(2)	Perception of equal opportunities	I-PIE-EOP
I-4(3)	More educational opportunities or better quality education	I-PIE-EDU
I-4(4)	More employment opportunities or job advancement	I-PIE-EMP
I-4(5)	Higher standard of living or better living environment	I-PIE-LIV
I-4(6)	Cultural experiences	I-PIE-CUL

EXCERPTS FROM FINAL LIST:

Code	Full Name	Definitions	Direct Quote
		Negative Immigration Experiences	
NIE-DEL	Difficulties with Everyday Life	Mothers encountered difficulties with everyday lives, such as a lack of instrumental support from family or a lower standard of living in the U.S.	Our parents [who stayed in our home country] could not help us. We had to rely on ourselves for everything, [but] we could only do things within our own abilities. Of course, that was difficult.
NIE-LAN	Language or Cultural Barriers	Mothers had communication problems or job limitations due to language barriers or unfamiliarity with the cultural and social norms in the U.S.	[It] is mainly the language barrier. Because I cannot express myself very well, so it is very difficult to be recognized/valued at work, which is due to [my] limitations in language.
NIE-SSN	Small Networks or Disconnectedness	Mothers had a small social network in the U.S., or expressed a sense of disconnectedness from the mainstream society	The social network here [in the U.S.] is quite limited, whereas having a lot of friends, classmates and colleagues around in China makes us feel more like at home. Here, we also socialize with some Chinese and American friends, but the feeling is different. Sometimes, I feel lost and irritable.
NIE-RAC	Limitations or Discriminations	Mothers encountered various challenges such as job limitations, complicated immigration processes, or racial discrimination in the U.S.	[I] gave up [my job] after [I] came. In fact, [I] completely abandoned my previous work [and therefore,] the level of [my] work is declining.
		Positive Immigration Experiences	
PIE-GPE	General Positive Experiences	Mothers had positive social encounters, reunited with their spouses, or gained valuable life experiences after migrating to the U.S.	I learned many things after coming to the U.S.; [such experience] has widened my vision and changed my perspectives on many things. [I] realized that my thoughts were very narrow before. [I] just feel that many of my thoughts have changed now.
PIE-EDU	Education/Career Opportunities	Mothers had access to quality education, more employment opportunities, or better career development in the U.S.	In the U.S., the educational system gives you freedom to choose the subjects you like, whereas in Hong Kong, your major is actually based on the [academic] grades you received.

(continues)

FIGURE 12.1 Coding the Interviews of Chinese Immigrant Women: Excerpts from a Start List and a Final List of Codes and Subcodes *(continued)*

PIE-EQU	Justice, Freedom, Equality	Mothers believed that social order and justice are better maintained and freedom and individual rights are guaranteed in the U.S.	The legal system is relatively sound; there is law in every aspect. Everyone obeys the legal system.
PIE-LIV	Better Living Environment	Mothers had a better living environment or a less stressful lifestyle after their migration to the U.S.	The U.S. is quite good overall. Life here is relatively simple [and I] like the environment. It's very convenient to live here.
Perceived Support at the Initial Stage of Migration and/or Throughout Their Transition			
HEP-INS	Instrumental Support	Tangible help with applications, employment, housing, financial difficulties, child care, transportation, or everyday needs.	[We received] a lot of help from our families. My parents helped us take care of our children and supported [us] financially.
HEP-INF	Informational Support	Information about universities or professional schools, or advice on everyday life and available resources in the U.S.	When I first came, she [a friend here in the U.S.] gave me a lot of advices on how to adapt to the life here, where to find a place to live, and on many things over everyday living.
HEP-ESS	Emotional or Spiritual Support	Emotional/psychological support, encouragement, and social contact, or spiritual support	When I had difficulties, I had love and support [from friends from church]. Then I had the courage to overcome any difficulties. My friends and classmates also gave me psychological support. They encouraged and comforted me.
HEP-CUL	Culture or Language	Assistance in learning the American culture or improving their English language skills	I went to the ESOL class offered in the county community college. They gave me a lot of help. They help new immigrants to learn English.

Source: From *The Immigration Experiences, Acculturation, and Parenting of Chinese Immigrant Mothers* (pp. 299, 303–305). by C. Y. Y. Leung, 2012, unpublished doctoral dissertation, University of Maryland, Baltimore County. Reprinted with permission.

When using two or more raters, qualitative researchers use strategies such as these to enhance interrater reliability of the final codes assigned:

- Include concrete definitions or significant characteristics of each code, along with specific examples from the data set (as illustrated in Figure 12.1).
- Provide training and practice in applying the codes to samples of data; continue until raters are consistently rating each data point.
- If necessary, revise the final list of codes or definitions to address ambiguities.
- After raters have independently coded the data, have them discuss and reach agreement about any inconsistently rated pieces of data.

We should mention, however, that some qualitative researchers think that achieving interrater reliability is an impossible goal, because different researchers will inevitably employ their unique, somewhat idiosyncratic lenses and interpretative frameworks in their efforts to make sense of the data.

8. ***Identify noteworthy patterns and relationships among the codes.*** For example, you might discern common themes that underlie many participants' experiences. If you look once again at Figure 12.1, you might notice that the codes "Racial discrimination" and "Job limitations" in Leung's start list became parts of a single larger category "Limitations or Discriminations" in her final list. Likewise, "More educational opportunities . . . " and "More employment opportunities . . ." became "Education/Career Opportunities."

Other patterns are apt to emerge as well. For instance, you might identify key sources of disagreement or conflict among participants, or you might detect inherent social

hierarchies among members of a certain social group or community. You might also discover that certain events almost always precede other events or that certain thoughts and feelings almost always precede certain actions. However, a qualitative researcher must always keep in mind that despite such *this-happens-and-then-that-happens* regularities, the researcher cannot conclude that the first event in the sequence definitely *causes* the second event. As we authors have stated before, *correlation does not necessarily indicate causation.* Although qualitative research can be especially useful in *describing* complex phenomena, only well-controlled experimental research can truly pin down causal, *this-variable-influences-that-variable* relationships.

Some patterns and relationships within the data might be readily apparent early in an analysis—and thus are reflected in a final list of codes—whereas others might come to light only after all the data have been coded. We must emphasize once again that in qualitative research, data analysis is often a back-and-forth, iterative process.

We previously introduced the term *open coding* as a strategy in which a researcher first goes through the data with an open mind, looking for various possible meanings and categories. Subsequent, more integrative analyses of the data and initial codings could involve *second-cycle coding* approaches such as *pattern coding, domain analysis, axial coding,* and *selective coding.* You will learn more about some of these strategies in upcoming discussions of grounded theory and ethnographic analyses.

9. ***Be alert for exceptions, tensions, and contradictions within the data set.*** A good qualitative researcher actively—and conscientiously—resists temptations to find commonalities that don't truly exist, make sweeping generalizations that aren't accurate, or jump to quick but unwarranted conclusions. Qualitative data are inevitably messy data that can't easily be wrapped up in one tight little package. Remember that the goal of qualitative analysis is not necessarily to achieve a single truth, or one answer, but rather to uncover multiple meanings and experiences.

Such is the *nature of nature:* Diversity reigns supreme. And the more complex the phenomenon under investigation—and human beings are incredibly complex beings—the more diversity one is likely to see.

10. ***Interpret the data in light of your research problem and questions.*** You must never forget the reason for conducting your study in the first place: to address a particular problem or question. Perhaps the data unequivocally provide one or more answers. Perhaps, instead, they only hint at certain answers and suggest that one or more follow-up investigations are in order. Only rarely do they offer nothing of value—an unfortunate outcome that is most common when a novice researcher naively embarks on a qualitative project with little or no advance planning and preparation.

The nature of your final interpretations—the *meanings* you ascribe to your data—can take a variety of forms. Following are common meaning-making strategies in qualitative research (Miles, Huberman, & Saldaña, 2014):

- Quantifying frequencies or probabilities of certain noteworthy characteristics or events
- Making comparisons or drawing contrasts within the data
- Connecting findings to one or more existing theories
- Developing a new, coherent theory to account for the findings
- Using metaphors to capture key phenomena or dynamics (e.g., describing the life of retirees as involving an "empty nest")
- Speculating about *possible* cause-and-effect relationships and *possible* mediating or moderating variables influencing those relationships (recall the discussion of mediating and moderating variables in Chapter 2)
- Creating graphics (e.g., concept maps, tables, flowcharts) that summarize general patterns in the data

You should also look for convergence (*triangulation*) within your data: Many separate pieces of information should all point to the same conclusions.

A strategy introduced in Chapter 8—writing *memos,* and more specifically, *analytical* memos—can be particularly useful in your analysis of qualitative data. As you interpret your data and begin to develop assertions from them, such ongoing notes and writings can help you identify general themes and patterns and then connect them to certain theoretical perspectives.

MyLab Education Self-Check 12.1

MyLab Education Application Exercise 12.1: Mapping the Initial Steps in Qualitative Data Analysis

Creswell's Data Analysis Spiral

We must repeat once again that qualitative data analysis is an iterative process, and thus a good qualitative researcher is apt to go back and forth a bit among the strategies just presented. Nevertheless, the researcher must gradually move forward, spending more time on the later strategies and leaving the earlier ones behind. Creswell (2013) has described a *data analysis spiral* that, in our view, offers a helpful perspective on how qualitative data analysis can reasonably proceed. Using Creswell's 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 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.

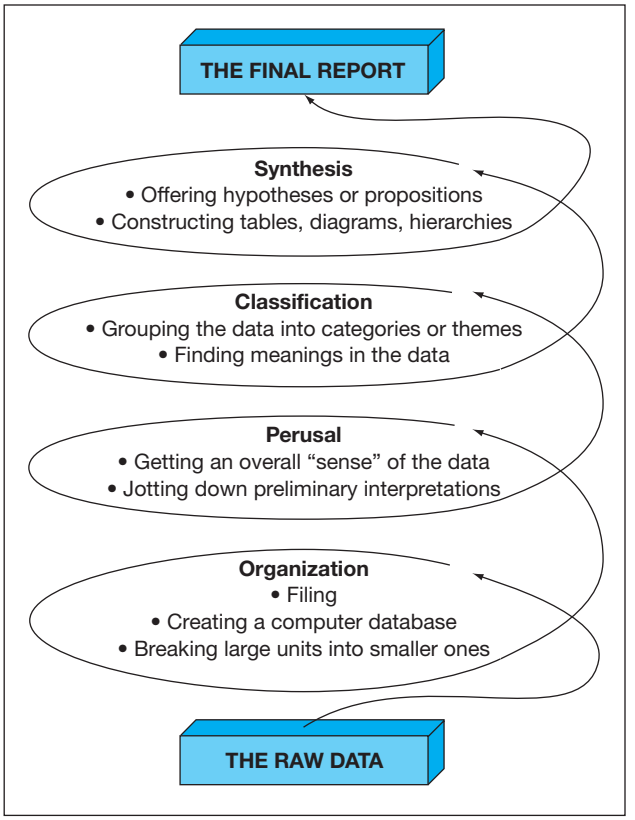
We depict this spiral graphically in Figure 12.2.

An Example: Data Analysis in a Grounded Theory Study

In many types of research, and especially in experimental quantitative designs, a major goal is to test hypotheses related to one or more existing theories. But as you should recall from Chapter 8, the chief goal of a grounded theory study is to develop a *new* theory, one that arises from (i.e., is grounded in) the data themselves. In no other design, then, are data analysis strategies more central to the entire research effort.

Experts disagree about the best approach to analyzing data in a grounded theory study (e.g., see Charmaz, 2014; Corbin & Strauss, 2008; Glaser, 1992). One widely used approach is that proposed by Corbin and Strauss (2008; Strauss & Corbin, 1990), who have suggested the following steps:

FIGURE 12.2 The Data Analysis Spiral (Based on Creswell, 2013)



1. ***Open coding.*** The data are divided into segments and then scrutinized for commonalities that reflect general categories or themes. After meaningful categories are identified, the data are further examined for *properties*—specific attributes or subcategories—that characterize each category. Properties are typically reflected in adjectives, adverbs, and adjectival and adverbial phrases in verbal material. For example, properties might be related to desirability (e.g., whether a certain object or event is to be sought or avoided), intensity (e.g., how angry someone feels), or duration (e.g., how long an event lasts). During this phase of analysis, data continue to be collected until it is clear that no new information is coming forward, no new categories are coming to light. At this point the categories are considered to be conceptually *saturated.*
2. ***Axial coding.*** During the process of open coding, one or a few categories might emerge as being central to the phenomenon under investigation. In axial coding, one of these categories is chosen as a *core category,* and other categories are identified as possibly reflecting:

   - The context in which the core category is embedded
   - Conditions that give rise to the core category
   - Strategies people use to manage or carry out the core category
   - Conditions that influence how the strategies are carried out
   - Consequences of those strategies

   In other words, the core category serves as an *axis* around which certain other categories appear to revolve in some way. Often the intercategory relationships are depicted as a diagram that encapsulates possible dynamics among various categories. If appropriate and helpful, the axial coding process is repeated using other codes as core categories.

3. ***Selective coding and theory development.*** A single category is chosen as *the* core concept in the phenomenon, and a theory is developed based on this concept and its interrelationships with other categories. This theory—which might take the form of a verbal statement, visual model, or set of hypotheses—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. The result tends to be a general *story line* that describes "what happens" in the phenomenon being studied.

We have described these steps only in the most general terms. Corbin and Strauss's (2008) *Basics of Qualitative Research* 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 experienced researchers' diverse approaches to data analysis techniques (e.g., see Charmaz, 2014; Corbin & Strauss, 2008; Garson, 2013; Martin & Gynnild, 2011).

An Example: Data Analysis in a Content Analysis Study

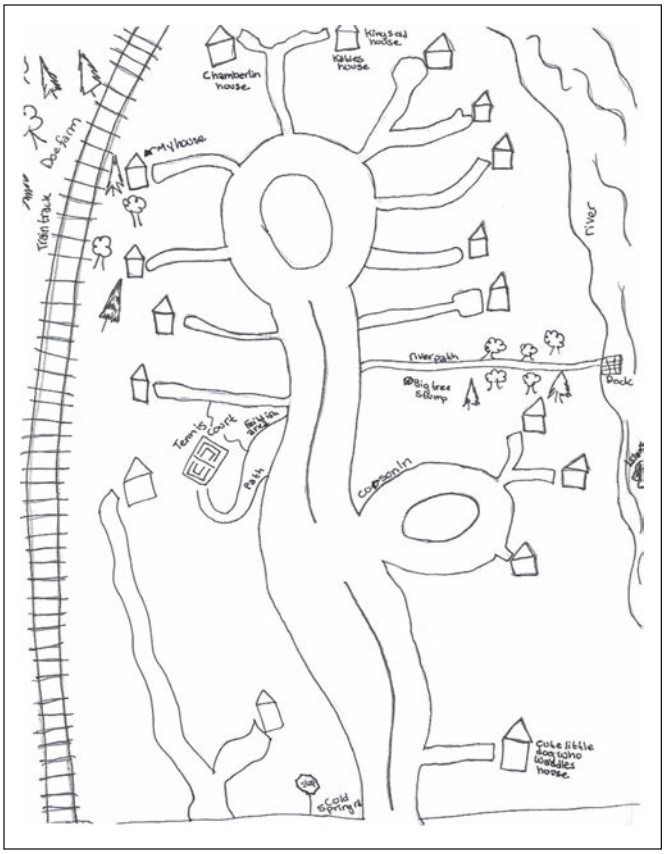
In one sense, virtually any qualitative study involves analyzing the contents of collected data. But in a content analysis study, determining the particular contents of verbal or nonverbal materials—typically human-created materials intended to communicate thoughts, ideas, and perceptions—is the *whole point* of conducting the study. For instance, a content analysis study might be aimed at identifying evolving techniques in ancient Greek sculptures, ideological biases in cable news channels, or popular topics on adolescents' Facebook pages.

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 related to the research question. The researcher then uses the tabulations and statistical analyses to interpret the data as they reflect on the problem under investigation.

Sometimes the research problem is specific enough that the researcher can identify, in advance, certain characteristics that will be searched for and counted. For example, when archeologists want to determine which pre-Mayan cultural group lived at a newly discovered Peruvian archaeological site, they might look for specific characteristics in artifacts and architectural structures known to be used by various cultural groups in pre-Mayan times.

In other situations, however, the research problem is less precise, such that initial data analysis involves an open-minded perusal of the data for ideas about significant characteristics to consider. Such was the case for one of us authors, who with several graduate students once asked the question, "What general developmental trends characterize children's conceptions of geographical space, as reflected in the maps they create of their local environments?" (Forbes, Ormrod, Bernardi, Taylor, & Jackson, 1999). To answer the question, the team had children at various grade levels create maps of their local neighborhoods from

FIGURE 12.3 ■ What Qualitative Differences Are Evident in These Two Neighborhood Maps Drawn by a 6-Year-Old (left) and 13-Year-Old (right)?



memory. The two maps shown in Figure 12.3 can give you an idea of the kinds of maps the children created. The first (showing the route from home to school) was drawn by a 6-year-old boy; the second (showing a bird's-eye view of several streets and surrounding geographical features) was drawn by a 13-year-old girl.[1]

At the time the study was conducted, no systematic research had been conducted on the nature of children's maps; thus, the research team had to start from scratch in its determination of helpful coding categories for the maps. To get a sense of what this process might involve, look closely at the maps in Figure 12.3. What characteristics pop out at you as being of possible interest? Amount of detail? Area of geographical space depicted? Perspective (e.g., side view vs. bird's-eye view of features)? Consistent scale? (For instance, notice the disproportionately large stop sign, tree, and picket fence in the 6-year-old's map.) Labels? Signage? Actions? (For instance, in the first map, a child says "[Th]ank you" to a school bus driver, and the second map mentions a dog that waddles.) Ultimately the research team had to make decisions about which of these and other characteristics were most important to focus on—which ones were most significant from a developmental perspective—as they analyzed their data.

An Example: Data Analysis in an Ethnographic Study

A primary goal of coding in any qualitative study is to identify patterns and thematic threads within the entire data set. In ethnographies, researchers have a particular interest in the identification and description of cultural patterns. As they situate themselves in, participate in, and observe various settings and contexts, they collect a good deal of field

[1]At the time the study was conducted, neither the researchers nor the children's parents/legal guardians anticipated that any of the data might appear in a book such as this one. The two maps shown in Figure 12.3 have been created more recently, with the parents giving appropriate written permission to publish them in this book. We use them here only for illustrative purposes.

data (field notes, interview transcripts, and artifacts) in order to gain insights not only into people's daily practices and behaviors but also into their **cultural knowledge**—the concepts, beliefs, attitudes, values, problem-solving skills, and other mental phenomena they have internalized through their daily interactions with other group members.

After ethnographers have collected a good amount of descriptive data but while they are still in the field, they often conduct a **domain analysis**, in which they identify certain cultural patterns that can guide further data collection and analyses (Spradley, 1980). In conducting such an analysis, they identify certain parts or elements of cultural meanings and organize them into larger units, or **cultural domains**, each of which has three elements:

1. A *cover term* is a general label for the domain (e.g., "teacher roles," "problem-solving skills").
2. *Included terms* are labels for smaller categories within that domain (e.g., "mentor," "trial and error").
3. *Semantic relationships* link the included terms to the cover term (e.g., "is one kind of," "can be used to").

Typically researchers will identify cover terms and included terms for various cultural domains during initial close readings of their data. In general, they look for *kinds* of things, people, and/or places within their data. Whenever possible, they should use *folk terms* as labels—that is, they should use participants' own language and terminology—although in some cases they might need to create their own *analytic terms* to describe domains (e.g., "religious beliefs," "moral codes").

As an illustration of what a domain analysis might involve, imagine that a researcher wants to conduct an ethnography of a community garden project. The researcher might initially notice gardeners talking about and using various kinds of planters; thus, *planters* would be an obvious cover term for this particular domain. As the researcher reviews field notes and interviews, included terms such as *clay pots, buckets,* and *plastic bins* might emerge. As analysis progresses further, these categories might be divided further, and new categories and semantic relationships could emerge. The initial domain analysis might lead the researcher to conduct focused observations of setting and activities to learn more about planters and planting and further refine categories, to determine the functions of planters or to identify various kinds of planter materials. In addition to "a kind of," semantic relationships could include "a way to," "a part of," "a reason for," "a stage in," or "a use for" (Spradley, 1980).

The ethnographic researcher pursuing this investigation of a community gardening project would identify the various elements of other domains as well—perhaps including domains related to "planting schedule," "ongoing nurturance of the plants," "division of labor," and "apportionment of crops." After completing such a domain analysis, the researcher can focus further data collection to better reveal the various domains and subsequently to create a complex taxonomy of the community's cultural knowledge.

PRACTICAL APPLICATION Using Computer Databases to Facilitate Data Organization and Analysis

We authors are assuming that you will use computer software to record interviews and keep track of other qualitative data. By storing your data on a computer, you can easily retrieve any piece of information using one or more relevant keywords, and you can sort your data quickly and in multiple ways. As a precaution against some unforeseen catastrophe (e.g., a fire or flood), you should regularly back up your data either on an external storage device (e.g., flash drive or external hard drive) or in an electronic dropbox or other Internet-based

storage mechanism. Regardless of the specific backup strategy you use, you should keep at least two copies of your data in different physical or electronic locations.

Most qualitative researchers also use computer software to organize and interpret a large body of 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, EthnoNotes, HyperQual, HyperRESEARCH, Kwalitan, MAXQDA, Nvivo, QDA Miner, Qualrus). Such programs provide a ready means of storing, segmenting, and organizing lengthy field notes, interview transcripts, and visual materials, and they're specifically designed to help you find patterns in your data. Some of them also enable you to compare how different raters have coded the same material, thus providing a means of determining interrater reliability. Typically you can import (i.e., transfer) data from word-processing files or electronic spreadsheets into the programs; some programs let you include graphic materials, photographs, audiotapes, and videotapes as well. Any one of these programs may take some time 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.

MyLab Education Self-Check 12.2

MyLab Education Application Exercise 12.2: Coding Qualitative Data

ACKNOWLEDGING THE ROLE OF RESEARCHER-AS-INSTRUMENT IN QUALITATIVE RESEARCH

Any research study requires the researcher to make important, informed decisions about how to proceed with data analysis. In quantitative designs, those decisions primarily involve selection of statistical procedures most appropriate for the research question and best suited to characteristics of the data. In quantitative data analysis, then, most important decisions occur up front, before data are collected. In contrast, qualitative research requires the researcher to make significant decisions and judgments throughout the data analysis process, not only about what strategies to use in general but also about which data are most likely to be noteworthy and how to evaluate and code specific pieces of data.

Because qualitative data analysis involves so much decision making—so many *judgments*—it can be especially vulnerable to a researcher's predispositions, expectations, biases, and values, reflecting the notion of *researcher-as-instrument* previously mentioned in Chapter 8.

True objectivity probably isn't possible in qualitative research (if, in fact, it's ever possible in *any* research project). Accordingly, good qualitative researchers take certain precautions to enhance the trustworthiness and credibility of their findings. Following are three widely recommended strategies.

1. ***Strive for balance, fairness, and completeness in data analysis and interpretation.*** As an alternative to maintaining complete objectivity—which, as just pointed out, is probably impossible, Wolcott (1994) has proposed that qualitative researchers should instead strive for *rigorous subjectivity*—in other words, they should take steps to ensure that their analyses and interpretations will ultimately be credible and defensible in the eyes of colleagues and other well-informed individuals. Such steps include the following:

- Triangulate multiple data sources.
- Intentionally look for exceptions and contradictions both within the sample selected and within the data collected.
- Remember that participants won't necessarily give you the Ultimate Truth; rather, they may tell you what they (a) *believe* to be true, (b) *wish* were true, or (c) think you want to hear.

- Continue to collect data until you have completely *saturated* the categories—that is, until you are no longer gaining new insights about the phenomenon of interest.
- Have two or more individuals independently code the data.
- Seek feedback from both participants and professional colleagues about your findings and interpretations.

All of these strategies should look familiar because, in one context or another, we have mentioned them all in earlier chapters. We bring them up again here because, taken as a whole, they can greatly enhance the credibility and defensibility of a researcher's findings and conclusions.

2. ***Carefully document your analysis procedures.*** In a quantitative study, a researcher may need to describe only the specific statistical analyses performed and perhaps also the specific software programs used to perform them. In contrast, a qualitative researcher must document and defend every step along the way—for example, how an initial list of codes was identified, how and in what ways the list was modified for final coding, how raters were trained to code the data reliably, and how raters resolved discrepancies in their judgments about particular pieces of data. Ideally, data analysis procedures should be described in enough detail that another researcher could replicate them and obtain similar results.

3. ***In your final report, be upfront about your personal biases.*** Credible qualitative researchers don't claim that they have approached a project with complete objectivity. (If they did, the rest of us would never believe them!) Instead, they carefully look inward, reflecting on and then describing possible beliefs, expectations, and cultural values that might have predisposed them to interpret their data in particular ways. With such knowledge, readers of the final report can better evaluate the credibility of a researcher's findings and interpretations. They are also more likely to perceive the researcher to be a person of honesty and integrity.

PRACTICAL APPLICATION Planning Data Analysis for a Qualitative Study

As you have probably realized by now, 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, might 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.

The following checklist can help you pin down specific strategies you might use as you consider how best to analyze qualitative data.

CHECKLIST

Pinning Down the Data Analysis in a Qualitative Study

WHAT WILL BE THE NATURE OF YOUR DATA?

_____ 1. What form(s) will your data take?

- _____ Verbal materials
- _____ Concrete objects/artifacts
- _____ Video recordings

_____ Audio recordings

_____ Other (explain): _____

_____ 2. Will some or all of your data need to be broken down into smaller segments for coding? If so, what criteria might you use to identify appropriate segments?

_____

_____

How Will You Code Your Data?

_____ 3. Do your research problem and/or its subproblems suggest particular categories you might use in coding? If so, what are they?

_____

_____ 4. Do existing theories or previous research studies related to your problem offer possible categories? If so, what are they?

_____

_____ 5. What steps will you take to create a final list of codes and subcodes?

_____

_____

_____

_____ 6. Will you have two or more people code the data? If so, what steps will you take to ensure reasonable interrater reliability?

_____ Concretely define each category and subcategory

_____ Provide examples of each category and subcategory

_____ Give explicit training and practice in applying the codes

_____ Establish a procedure for resolving inconsistent ratings (explain): _____

_____

_____ Other (explain): _____

_____

How Will You Analyze and Interpret the Coded Data?

_____ 7. What software program(s) might be helpful in analyzing the data?

_____

_____ 8. What strategies will you use to identify general trends within the data?

_____

_____

_____

_____ 9. What kinds of exceptions and inconsistencies might you find within the data?

_____

_____

_____ 10. What personal or theoretical biases do you have relative to the outcomes of your study? How can you minimize their influence on your analyses and interpretations?

_____

_____

_____ 11. Will it be helpful to analyze the coded data quantitatively as well as qualitatively? If so, what computations might you perform?

- _____ Frequency counts
- _____ Percentages
- _____ Measures of central tendency (modes, medians, means)
- _____ Measures of variability (ranges, interquartile ranges, standard deviations)
- _____ Correlation coefficients
- _____ Other (explain): _____

_____ 12. How will you document your data analysis procedures?

_____

_____

_____ 13. What strategies will you use to ensure that you and others can have confidence in your findings and interpretations?

- _____ Triangulating multiple data sources
- _____ Using two or more coders/raters
- _____ Saturating categories
- _____ Using direct quotes from interview data to support your conclusions
- _____ Getting feedback from participants
- _____ Getting feedback from experts and professional colleagues (if yes, list these individuals): _____
- _____ Other (explain): _____

To the extent that you would find it helpful to conduct certain *quantitative* analyses of qualitative data (as indicated in your responses to Item 11 in the preceding checklist), some might argue that you are really conducting a mixed-methods study rather than a strictly qualitative one. We authors don't entirely agree with this perspective; in our minds, computing a few simple statistics for qualitative data—perhaps means or percentages—doesn't magically transform a qualitative design into a mixed-methods one. Rather, a mixed-methods study is one in which the researcher plans *in advance* for the collection of (a) some data that will either be quantitative in nature or else will eventually be quantified during data analysis and (b) some data that will remain qualitative throughout the data analysis process.

MyLab Education Self-Check 12.3

MyLab Education Application Exercise 12.3: Minimizing the Impact of the Researcher-as-Instrument

ANALYZING AND INTERPRETING MIXED-METHODS DATA

Strategies for analyzing and interpreting data in a mixed-methods study vary widely, depending on the general design of the study. One decision you must make early in the game—ideally, before you collect your data—regards whether you will (a) give 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 ignore one set of data simply because you don't like what it tells 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.

In general, any data analysis in a mixed-methods study will probably include at least three of the following four steps (Onwuegbuzie & Teddlie, 2003):[2]

1. ***Condensing the data.*** As noted in Chapter 1, the human mind can think about only so much—and not *very* much—at any single point in time. Trying to consider every single data point separately as you look at your data . . . well, you just can't do it. Instead, you need to boil it down to more compact entities your mind can reasonably handle at once. Condensing quantitative data might involve calculating basic descriptive statistics, such as means or medians, standard deviations or ranges, or correlation coefficients. And as you've learned earlier in this chapter, condensing qualitative data often involves inspecting the overall data set to get preliminary ideas about general themes or potentially useful codes.
2. ***Depicting general patterns in the data in ways that enable quick visual inspection.*** For example, quantitative data might be presented in line graphs, bar graphs, or statistical tables. Qualitative data might be depicted as diagrams, matrices, or hierarchical taxonomies.
3. ***Transforming qualitative data into quantitative entities (an optional step).*** Occasionally mixed-methods researchers transform quantitative data into a qualitative form—say, a general narrative—but more typically the transformation is in the other direction. For example, after coding discrete elements within a data set—perhaps participants' responses to interviews or perhaps works of literature undergoing content analysis—many qualitative researchers count instances of each code. Alternatively, a researcher might use rating scales to assign rough indicators of magnitude or intensity to discrete pieces of qualitative data (see Miles et al., 2014).
4. ***Systematically comparing and integrating results from the two types of data.*** As you learned in Chapter 9, some mixed-methods researchers use data from one or more early phases to guide the planning and implementation of one or more subsequent phases; this is especially likely to be true for exploratory designs and multiphase iterative designs. Whether or not such preliminary data integration has occurred in an early phase, it *must* occur after all data have been collected and analyzed. Again, there are no prescriptions for integrating quantitative and qualitative findings; the only generalization we can make is: *It all depends.* Such a generalization certainly isn't helpful for novice researchers. But we can offer a few possible strategies you might consider (Bazeley, 2010; Creswell, 2014, 2015):

   - Create a two-column table—one column for quantitative findings and another for qualitative findings—that can reveal consistencies or inconsistencies in findings related to each subproblem or research question within the overall research problem.
   - Identify concrete examples within the qualitative data that *give meaning and depth* to the numbers.
   - After conducting both a quantitative paper-and-pencil survey and open-ended, qualitative interviews, count the number of times each theme appears in the interviews and then compare the frequency counts with the quantitative analysis of survey responses.
   - Divide participants into two or more groups based on certain demographic variables (e.g., age group, gender, political party membership), and statistically compare the frequencies of various codes or themes within the qualitative data for each group.

[2]Onwuegbuzie and Teddlie (2003) described seven discrete steps. We have consolidated their Steps 4, 5, 6, and 7 into a single Step 4.

We illustrate the last of these strategies with a two-phase sequential study conducted by Harkness and seven colleagues (2011). These researchers' overall research question might be summarized like this: "What similarities and differences might there be among various Western cultural groups in children's normal daily activities and in the meanings that parents ascribe to those activities?" Participants were parents from six different countries—Italy, the Netherlands, Poland, Spain, Sweden, and the United States—that were presumed to represent somewhat different cultural groups. As part of a much larger international research project, the parents had already provided information about various demographic variables—for instance, about their children's ages and genders and about the parents' ages, educational backgrounds, and occupations.

Harkness and colleagues (2011) used two strategies to collect data that might help them answer their research question:

1. In *parent diaries,* parents reported on what their children specifically did each day over a particular 1-week period. These diaries consisted of researcher-created forms on which parents recorded every activity in which a child engaged, along with the time and place at which the activity occurred and the other individuals who were present during the activity.
2. In follow-up *semistructured parent interviews* conducted at a time convenient for both mother and father, parents were asked about the importance and meanings they attached to each activity previously described in their parent diaries.

The researchers began their data analyses by using approximately 80 previously identified categories to code the qualitative data collected from the diaries and interviews, with the option of adding new categories as needed. From these initial codes, they subsequently derived more general, overarching codes—eight for the children's activities recorded in the diaries (e.g., "meals, including snacks," "family activities," "play," and "watching TV") and nine for the general themes in the meanings that parents attached to the activities (e.g., "cognitive development," "rules and values," "family relationships," and "health, motor skills, and exercise"; Harkness et al., 2011, p. 803).

After completing the qualitative coding processes, the research team moved on to various quantitative analyses. For instance, they computed the average daily amounts of time that each child spent in the eight categories of activities and then used such statistical procedures as analysis of variance (ANOVA) and multiple analysis of variance (MANOVA) to identify cultural-group differences in the kinds of activities in which children of various age levels engage. (You can find brief descriptions of these two statistical procedures in Table 11.5 in Chapter 11.) Also, they compared the relative frequencies with which parents in the six cultural groups ascribed various meanings to the children's activities, as reflected in the nine general themes that had emerged from the qualitative analysis of parents' interview responses—for instance, the extent to which participants from the six different countries described certain activities as serving the purpose of promoting cognitive development or, instead, physical well-being.

Of the five mixed-methods designs previously presented in Chapter 9, the study we've just described can probably best be categorized as a two-phase *explanatory sequential design,* but with a notable exception. In its simplest form, an explanatory sequential design involves (a) collection of quantitative data followed by (b) collection of qualitative data (e.g., participants' explanations of or reflections about their quantitative responses) that can help a researcher make sense of the numbers. In this study, however, both phases involved the collection of qualitative data that were subsequently transformed into quantitative data. Using the symbols previously introduced in Chapter 9, then, the study can probably best be described as having a QUAL+quan → QUAL+quan design.

In integrating their quantitative and qualitative findings in their final research article, Harkness and her colleagues (2011) began by reporting their quantitative findings and then gave depth and meaning to those data with illustrations from the qualitative data. For

example, although all six cultural groups talked about the importance of healthy family relationships, they differed somewhat in the strategies they used to foster such relationships. In this statement, one Polish mother expressed her strong commitment to spending a good deal of time together as a family:

> On Sundays we operate on the understanding that it is solely time for our family. We as parents do not work, we are not preparing to work, we don't go shopping. We spend the entire day together. We organize family outings to the cinema or to visit the grandparents. We invite cousins, we play chess or Scrabble, or soccer. [In addition] we eat supper together every day and we celebrate all family events. We pray together. We try to have quality time with the children on a daily basis, just being together or talking with them. (p. 809)

In contrast, some parents in the United States seemed to struggle a bit in their efforts to balance their work and home lives, as an American mother revealed:

> We need to get into, especially with the school year coming and Meredith being . . . I have one friend that their family night is like Monday night. So it's like we need to get into a family night where we can just . . . you know . . . it's just family time. Like he [father] said, usually by the time he gets home from work it's dinner time and it's almost time for the kids to go to sleep. So with family time . . . we need to set around and read a book or go out together or whatever. At least by September we need to schedule a day. (p. 810)

And more generally, Harkness and colleagues found that even though family activities were *more* frequent in the U.S. group than in any of the five European groups (a quantitative finding), nonetheless, the American parents seem to be the most frustrated by and concerned about the relative lack of time they could spend with their children each day (a qualitative finding). The researchers could only speculate about possible reasons for the parents' frustration and concern. Perhaps it was the result of longer work days for the American parents than for the European parents. Perhaps, instead (or in addition), it was due to the fact that the American families didn't seem to eat together as often as the European families did. Other factors may have been sources of frustration and concern as well. As the research article's authors pointed out, such hypotheses would have to be tested through further data collection and analysis.

Remember, like any other researcher, a mixed-methods researcher can't just stop after reporting the results of data analyses. The researcher must also interpret those results—in this case, not only the qualitative results and the quantitative results but also the *combination* of those two sets of findings.

PRACTICAL APPLICATION Using Computer Software Programs to Facilitate Mixed-Methods Data Analysis

We authors hope that at this point in the book we have convinced you of the importance of using one or more software packages to assist you in data analysis. Some software programs are especially suited for mixed-methods data analysis; examples include Dedoose, MAXQDA, NVivo, and QDA Miner. Such software programs typically enable a researcher to:

- Code a variety of qualitative data (interview transcripts, field notes, artifacts, visual data and images, social media documents, videos, and/or audio files)
- Sift through, search, and sort data that might be relevant to particular subproblems or research questions
- Convert qualitative data into simple quantitative data
- Create matrixes, graphs, and/or visual representation that summarize certain aspects of a data set (e.g.,*word clouds* that use varying font sizes to depict the relative frequency of noteworthy words in verbal materials)
- Integrate and compare results from quantitative data and coded qualitative data
- Compare data obtained from various demographic groups

As mixed-methods research studies have become increasingly popular in recent years—and also increasingly complex—software developers have continued to revise and expand the capabilities of programs suitable for mixed-methods data analysis. Probably the best resource for learning more about the *current* features of these programs is the Internet. For example, a simple search for "mixed-methods data analysis software" on Google, Bing, or another search engine should lead you to many helpful websites.

MyLab Education Self-Check 12.4

SYSTEMATIC REVIEWS OF QUALITATIVE AND MIXED-METHODS STUDIES

When a researcher wants to consolidate the results of many quantitative studies related to a particular research problem, the researcher might conduct a *meta-analysis*—a statistical analysis of all of the studies' individual statistical results—in order to discern general trends in the findings. But when a researcher wants to synthesize the results of many previous qualitative or mixed-methods studies, a statistical meta-analysis obviously isn't possible, at least not for most of the reported data.

A viable alternative for qualitative and mixed-methods studies is a **systematic review**, in which *research reports,* rather than individual people or other individual entities, are the objects of study (Harden & Thomas, 2010; Petticrew & Roberts, 2006). The review is *systematic* in the sense that a researcher identifies and implements an explicit, rigorous method for selecting and analyzing the reports. For example, the researcher is likely to:

- Conduct an extensive search for studies related to the research problem—for instance, by using certain keywords in appropriate online databases and then including all relevant peer-reviewed qualitative and/or mixed-methods research reports (e.g., articles, dissertations, conference presentations) in the sample.
- Evaluate the quality of each report (e.g., were rigorous methods used to collect and analyze data? are obvious researcher biases affecting conclusions?) and then possibly either exclude or give less credence to certain reports.
- Code the contents of the reports for key ideas, concepts, themes, and/or theories.
- Perform one or more meta-analyses of any statistical findings reported in the studies.

Aside from taking steps such as these, there is no single "best" way to conduct a systematic review. As is true for most qualitative and mixed-methods research, the specific strategies used depend on the nature of the research problem and the particular methodologies used in the individual studies included in the sample.

A study described by Harden and Thomas (2010) illustrates what a systematic review might involve. In response to a request from the United Kingdom's Department of Health, these two researchers and several colleagues wanted to review current research regarding effective strategies for getting children to eat more healthfully. The research team identified their overall research question and three more specific questions as follows:

> What are the barriers to and facilitators of healthy eating?
>
> - Which interventions are effective in promoting healthy eating habits?
> - What are children's perspectives on, and experiences of, healthy eating?
> - What are the implications of Questions 1 and 2 for intervention development? (Harden & Thomas, 2010, p. 760)

In an extensive search of relevant research with children ages 4 to 10, the researchers located 33 studies that included both a treatment group and a control group to evaluate the effectiveness of an intervention designed to improve eating habits. They also found

eight studies that examined children's perspectives about healthful eating. These 41 studies included a variety of research methodologies (e.g., experiments, surveys, interviews) and, in combination, yielded a large body of quantitative and qualitative data.

The "data" that the team initially analyzed were the verbatim contents of sections in the reports that were labeled either "results" or "findings." Analyses of these data involved three major steps:

1. Separately coding each line of the text for one or more possible meanings, sometimes using words that the study authors themselves had used
2. Grouping the Step 1 codes into more general descriptive themes and subthemes related to what children's understandings of healthy eating involved and what factors affected their food choices
3. Identifying six overall "analytical themes" (i.e., general conclusions) within the data that each addressed at least one of the team's research questions and either directly or indirectly had implications for intervention strategies (e.g., children like to make their own choices about food, children don't seriously consider the long-term consequences of poor eating habits)

The six analytical themes, in turn, became the focuses of subsequent qualitative and quantitative analyses. For example, the team examined various interventions to determine which ones best incorporated principles consistent with the analytical themes identified in Step 3. It then calculated *effect sizes*[3] for the interventions and asked the question, "Did interventions consistent with one or more of the identified analytical themes have a greater impact than other interventions?" The answer was *yes,* leading the research team to make some tentative recommendations for future policies and practices.

PRACTICAL APPLICATION Observing How Experienced Researchers Have Conducted Qualitative and Mixed-Methods Research

In virtually any complex human endeavor—for instance, in carpentry, figure skating, musical composition, and computer programming—novices gain new skills and increasing expertise through careful observation of what experienced individuals do. This principle certainly applies to research methodology: Beginning researchers gain considerable proficiency by observing what expert researchers do, in some cases by watching them in the field or in a lab and in other cases by reading their research reports. Especially in qualitative and mixed-methods research—where possible strategies for data collection and analysis are virtually endless—discovering how expert researchers have tackled research problems similar to your own can give you many useful ideas for planning a methodology and data analysis strategies. Thus, before you embark on a qualitative or mixed-methods study, we urge you to read many, many qualitative or mixed-methods research reports in search of ideas about how you might best carry out your own study.

As you read these reports, however, you must keep in mind that not all published research studies are good ones; some are apt to have methodological weaknesses that adversely affect the credibility of the results obtained and conclusions drawn. When you find one or more weaknesses in a particular study, think creatively about how the researcher might have strengthened the study—and, indirectly, also how you might strengthen your own study.

[3]As noted in Chapter 11, an *effect size* is a statistical indicator of how much of an impact an experimental intervention has in comparison with outcomes for a control group or other comparison group.

Drawing from our discussions in this chapter, Chapter 8, and Chapter 9, as well as from guidelines offered by Good (1993) and Miles and colleagues (2014), we offer the following checklist to help you evaluate other researchers' qualitative studies and final research reports.

CHECKLIST

Evaluating a Qualitative or Mixed-Methods Study

	METHODOLOGY	YES	NO
_____	1. Is the context/setting of the study adequately described?	_____	_____
_____	2. Are techniques for data collection appropriate for the research problem/purpose? Are these techniques 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 or materials presented? Is the sample described in sufficient detail?	_____	_____
_____	6. Are potential ethical concerns appropriately addressed?	_____	_____
_____	7. Are the roles of the researcher and participants made clear?	_____	_____
_____	8. Might the researcher's presence at the research site influence what participants do or say?	_____	_____
_____	9. Does the researcher identify any assumptions, beliefs, values, or biases that might influence data collection or analysis?	_____	_____
_____	10. Does the researcher actively look for evidence that might disconfirm any tentative hypotheses and explanations?	_____	_____
	FINDINGS AND INTERPRETATIONS		
_____	11. Are the data analysis techniques appropriate for the research question, methodology, and theoretical framework?	_____	_____
_____	12. Are data analysis techniques explicitly described?	_____	_____
_____	13. Do data analysis techniques allow for revision and reinterpretation as new data come to light?	_____	_____
_____	14. Are various data sources triangulated?	_____	_____
_____	15. If used, are tables, figures, and other graphics easy to read and interpret? Do they enhance the reader's ability to understand the study?	_____	_____
_____	16. Are sufficient data reported to support the conclusions drawn?	_____	_____
_____	17. Are any irrelevant and unnecessary data reported? If so, what should be deleted? _____ _____	_____	_____
_____	18. Are discrepant data discussed and reconciled?	_____	_____
_____	19. Have the setting and observations been sufficiently described to present a convincing case?	_____	_____
_____	20. Are participants' "voices" used to support the assertions and present multiple perspectives?	_____	_____
_____	21. Is the report detailed enough that the findings can be compared to other studies in other contexts?	_____	_____

_____	22. Is the discussion congruent with the research question and rationale for the study?	_____	_____
_____	23. Are implications for theory and/or practice discussed?	_____	_____
_____	24. Have other scholars in the field reviewed the proposal or report? If so, do they agree that the approach, methodology, and conclusions are appropriate?	_____	_____
_____	25. Have participants in the project read the report? Do they agree with its findings?	_____	_____

MyLab Education Self-Check 12.5

A SAMPLE DISSERTATION

In Chapter 2, Chapter 8, and this chapter's Figure 12.1, we have previously looked at brief excerpts from Christy Leung's doctoral dissertation regarding Chinese mothers' reasons for immigrating to the United States, experiences prior to and after immigrating, and adjustment and parenting strategies. We now close the chapter by examining some of her qualitative data analysis strategies, which effectively illustrate several of the chapter's recommendations.

Leung conducted her study within the context of a large mixed-methods research project for which her doctoral advisor, Dr. Charissa Cheah, was the principal investigator. She collected and analyzed both quantitative and qualitative data related to several intertwined research questions. Many of the excerpts you will see in the following pages focus on Leung's efforts to learn about the women's "positive and negative experiences as well as their perceived support received at the initial stage of migration and/or throughout their time in the U.S." (Leung, 2012, p. 116).

DISSERTATION ANALYSIS 9

3.5.2 Preparation of qualitative data. Interviews were conducted with 50 mothers (27 boys and 23 girls) who migrated to the U.S. at the age of 13 years or above. . . . The 50 audio-taped interviews were transcribed and checked by multiple bilingual graduate and undergraduate students for accuracy. Prior to the content analyses, the transcripts of the interviews conducted in Chinese were translated into English and then reviewed by multiple bilingual graduate and undergraduate students. Students discussed any discrepancies in translations with each other until they agreed with the final version of the translation to ensure translation accuracy.

In order to provide a full description of the context of the qualitative content analysis, background information about the two coders and the auditor, the consensus process, and issues with potential biases are discussed below.

Comments

Leung used numbers in her headings to indicate her overall organizational scheme. Section 3.5.2 is a subsection of Section 3.5 ("Data Preparation") in Chapter 3 ("Methodology").

Many of the interviews were conducted in the participants' native Chinese. Notice the use of multiple raters to translate interview responses into English, thus ensuring that the data to be analyzed were accurate representations of what participants had said.

Coders and auditor. A team of two coders, which comprised a graduate assistant (the author of this dissertation project) and an undergraduate assistant, primarily developed the coding scheme and then coded the interview data. An auditor reviewed the coding process and analyses and provided feedback to the coders (Hill et al., 2005). . . . Regarding the first coder, the graduate assistant is a first-generation Chinese American who immigrated to the U.S. during her late adolescence. She is fluent in both spoken and written Chinese. She supervised all of the translation of research materials and the transcription and translation of interview transcripts. Prior to the coding of interview data for the present study, she was trained by the principal investigator of the larger project to independently code semi-structured interview data regarding: (a) the expressions of love and care as well as the long-term socialization goals of Chinese immigrant mothers, (b) the socialization of Aboriginal-Canadian and European-Canadian mothers, and (c) the social cognitive reasoning of Chinese-Malaysian adolescents.

The use of two coders and an auditor reflects the author's concern about interrater reliability. The descriptions of the coders in this paragraph and the following one should reasonably persuade readers that the two individuals had sufficient background to knowledgeably code the data. Note that the first coder was the author herself.

Regarding the second coder, the undergraduate assistant is a second-generation Chinese American who grew up in an urban neighborhood with a large Chinese immigrant population in New York. She is fluent in spoken Chinese. She reviewed the English or English-translated version of all interview transcripts prior to the coding. She has knowledge of the literature of immigration, acculturation, and adaptation of Chinese immigrants. Under the mentorship of the principal investigator of the larger project and the first coder, she completed an independent research project on Chinese mothers' experiences in coming to the U.S. and expectations for their children's adaptation and success in the U.S. . . . Given their background and experiences, the two coders are familiar with Chinese immigrant communities. They can understand the narratives and experiences of Chinese immigrants as cultural insiders.

(In a subsequent paragraph, the author also describes the background and qualifications of Dr. Cheah, who served as auditor for the project.)

Consensus process. Using multiple coders is a preferable practice in qualitative research (Hill et al., 1997, 2005). The multiple perspectives and insights generated from the coders' discussions on coding disagreements help minimize their subjective bias and improve the quality of data analysis (Barbour, 2001). On the other hand, the process of developing a coding scheme and coding of interview transcripts involves consistent discussion and lengthy consensus between coders. Moreover, the nature of the consensus process requires that coders need to respect each other, value others' perspectives, have confidence to share their opinions, feel comfortable to negotiate disagreements, and welcome feedback from others (Brodsky et al., 2004; Hill et al., 1997, 2005).

Here the author accurately points out that using two coders can minimize the influence of any single coder's biases. It is possible, of course, that two coders would be influenced by the same assumptions, expectations, and/or values—a potential source of bias that any qualitative researcher should keep in mind.

Thus, both of the coders came to a mutual understanding that equal voice, willingness to share different opinions, and openness to negotiate intellectual conflicts were essential during the coding process. Despite the effort to develop a supportive relationship and facilitate open discussions among the coders, both of the coders recognized the existing power differentials between themselves due to the graduate/undergraduate student dynamics, mentor/mentee academic hierarchy, and differences

Notice the author's concern that the two coders have equal say during the coding process. Regular involvement of the auditor helped to minimize the possibility that one coder's opinions would dominate over those of the other coder.

in research knowledge and experiences. In order to avoid the potential dominant voice of one coder over the other coder, they consulted with the auditor throughout the consensus and coding process. . . .

Potential biases. Addressing potential biases at the initial stage of data analysis is also critical in qualitative research (Hill et al., 1997, 2005). Both of the coders believed that Chinese cultural values continue to play a significant and influential role in Chinese family socialization. Specifically, the two coders articulated their expectations that the Confucian ethic of filial piety would be reflected in the socialization goals and parenting practices of Chinese immigrants. To prevent biases from interfering with coding and data analysis, the two coders and the auditor acknowledged these expectations and came to a mutual understanding that the coders put aside their expectations throughout the process.

Here the author acknowledges that the coders might have had similar biases in how they coded the data. Making a conscious effort to put aside expectations can help reduce their biasing effect but would probably never eliminate it. (Look once again at Figure 1.3, "Common Pitfalls in Human Reasoning," in Chapter 1.)

In addition, they reviewed and came to a consensus on some general coding guidelines, such as coding by themes, coding based on evidence in the transcripts, and multiple coding (Hill et al., 1997, 2005). In order to understand the context of the responses, coders (did) not utilize the line-by-line coding strategy because fragmenting sentences within a response might fail to capture the context of the response. Instead, coders coded responses by themes or ideas. Thus, one sentence might be coded more than once across different segments of a response. However, coders (had to) identify evidence in the transcripts to support their coding decisions. They might make short notes to indicate their interpretations of specific responses and to justify their coding decisions. Moreover, multiple codes (could) be assigned to one response as long as the response contained multiple themes.

Notice how the researcher/author (who was also the first coder) involved the second coder in the creation as well as the application of the coding scheme.

Here the author provides a rationale for the particular approach she used to code the data.

Notice that the coders kept track of their reasons for applying various codes to different parts of the data. Such notes-to-self might be useful in situations where the coders disagreed about how to code certain parts of an interview.

(At this point the author begins Section 3.6, "Planned Analyses," and first addresses analyses of her quantitative data. We pick up the document again when she returns to her analyses of the qualitative data.)

3.6.2 Qualitative content analysis. *(In the first paragraph of this section, the author reminds readers of the research questions that the qualitative data were collected to address.)*

Based on Hill et al.'s (2005) consensual qualitative research method (CQR), a qualitative content analysis (Fereday & Muir-Cochrane, 2006) was conducted to systemically classify and describe the themes regarding the mothers' immigration experiences and parenting revealed from the audio-taped interview. Utilizing a mixed approach that integrates the deductive thematic analysis (Crabtree & Miller, 1999) with the inductive coding process (Boyatzis, 1998), some codes were derived theoretically to address the research questions at the initial phase, whereas other codes were generated to capture the new themes that emerged in the interviews (Fereday & Muir-Cochrane, 2006; Miles & Huberman, 1994). During the content analysis, a coding scheme was developed to create categories of themes, and a cross-analysis was conducted to identify similar and different themes across mothers with the four acculturation strategies.

Notice how some initial codes had a theoretical basis, whereas others emerged during initial inspections of the interview data. The author cites several well-regarded qualitative researchers, who have presumably recommended her methodological choices.

The cross-analysis involved comparing responses of four distinct groups of mothers whose quantitative data had revealed differing acculturation patterns (i.e., differing ways of adapting or not adapting to U.S. culture).

Open coding and development of a coding scheme. A start list of codes was created based on the theoretical concepts reviewed in the literature and the research questions stated in the present study (Miles & Huberman, 1994; see Appendix L). To develop a coherent and structured start list, each research question served as a higher-order code which was accompanied by a subset of corresponding codes that captured the themes in response to the specific research question. The start list was reviewed and discussed by the coders and then checked by the auditor to reach an agreement regarding the classification and definition of codes.

As you should recall, excerpts from both her start list and her final list appear earlier in the chapter (see Figure 12.1). Notice how the researcher uses her research questions as a basis for organizing the start list.

Based on the start list and the coding guidelines, open coding was conducted where the two coders independently analyzed and coded two interviews, chosen randomly by using the random numbers table. In coding the two interviews, each coder also identified new, distinct ideas and created conceptual codes based on similarities in the meaning of distinct ideas from the open-ended responses. Next, the team compared the codes, discussed any discrepancies, and consulted with the auditor to reach a consensus in order to generate new codes. They also discussed and revised the labeling and definition of the codes to better capture the meaning of each code. The coders repeated this process with an additional 22 interviews. They modified and refined the coding scheme and finalized the coding and classification guidelines in order to establish a common understanding of the coding scheme. The process of discussion, consultation, and consensus process involved coding two interviews at a time. They discussed any discrepancies in their coding and consulted with the auditor to resolve disagreement, to seek advice in interpretations of responses, or to clarify applications of the established codes. . . .

Notice the iterative process described here: The coders moved back and forth between revising the list of codes and applying it to additional data. This was undoubtedly a painstaking endeavor, but it helped to ensure that the final list of codes would capture the multiple nuances and meanings of the interview data.

Axial coding and cross-analysis. The remaining 26 interviews were coded based on the established coding scheme. At this point in the data analysis, the coding application and revision were conducted by the first coder, and then reviewed by the auditor. Based on a series of conjoint review and discussion sessions, they modified and refined the coding until any disagreements in coding were resolved by consensus (Hill et al., 2005). Given that the coding scheme was relatively comprehensive, no new codes were identified during the process of coding the additional 26 interviews. Next, axial coding was conducted where some of the codes representing similar or overlapping themes were combined and the definitions of those codes were revised to reduce the number of codes. Finally, a cross-analysis was conducted to compare themes across the mothers in this sample (see Appendix M for the final list of codes).

Here the author presents convincing evidence ("no new codes") that the list of codes developed from the first 24 interviews captured all noteworthy meanings within the interview data.

Based on the guidelines provided in Hill et al.'s (2005) CQR update, the percentages of the codes were categorized by using labels in order to determine how often the codes applied to the mothers in this sample. . . In the present study, a code category was identified as *general* when it applied to *almost all or a majority of the cases* (80%–100%), *typical* when it applied to *most of the cases* (50%–77%), *sometimes* when it applied to *some of the cases* (20%–45%), or *rare* when it applied to a *small number of the cases* (1%–19%). . . .

Calculating percentages and defining what is meant by such terms as "typical" and "sometimes" are two simple yet effective ways through which the author enhances her ability to analyze and represent her data in an unbiased manner.

(After two additional sentences in the preceding paragraph, the author begins Chapter 4, which she straightforwardly calls "Results." We pick up the dissertation again when she turns to qualitative results related to the mothers' negative and positive immigration experiences. We present only small portions of the section in order to illustrate how a researcher might effectively communicate qualitative findings.)

4.2.2 Qualitative results The second research aim in the qualitative portion was to explore Chinese immigrant mothers' (a) negative and positive immigration experiences, and (b) perceived support at the initial stage of migration and/or throughout their transition to the U.S. . . .

Once again, the author reminds the reader of her second research question relative to the qualitative data.

Negative immigration experiences. Regarding the negative experiences that the mothers in this sample had with their migration or since being in the U.S., four themes emerged from the analyses . . . : (a) difficulties with everyday life, (b) language or cultural barriers, (c) small social networks or a sense of disconnectedness, and (d) limitations or discrimination (see Table 6). Different patterns of responses across the four acculturation groups were identified. . . .

The author uses many subheadings within the section to organize her findings, and in the first paragraph of this subsection, she provides an advance organizer for the upcoming discussion (recall the discussion of advance organizers in Chapter 1).

Many of these mothers, such as Ms. Gao, a marginalized mother who migrated at 30 years of age, discussed the problems they encountered communicating with individuals outside their own ethnic group due to language barriers: "When you are talking (with others) outside (the home), people sometimes do not understand what we mean by what we say." Thus, these mothers indicated that they often cannot express themselves well in English and experience a feeling of not being understood by the others. Moreover, several mothers mentioned that they had difficulties finding a job or suffered from a decline in their occupational status with a lower earning potential because of their poor English skills.

The term marginalized refers to one of the four acculturation groups the author had identified; the other three groups were labeled "assimilated," integrated," and "separated."

"Ms. Gao" and other names in the report are pseudonyms, thus maintaining confidentiality and protecting the participants' right to privacy.

In addition, mothers in all four acculturation groups had to face various challenges at work because they were not familiar with the cultural and social norms in the U.S. Specifically, some of these mothers had difficulties building professional relationships because they were not competent in socializing with their American colleagues. Other mothers believed that they missed professional opportunities because they had been socialized to be modest and reserved in their home country, whereas their American colleagues had been taught to be confident and assertive in the U.S. Mothers such as Ms. Hui, an assimilated mother who migrated at 25 years of age, also expressed that Chinese women might not present themselves with confidence effectively at work due to such cultural differences: "I can speak out only when I know something very well. However, I find out that it is not the case with others sometimes. Some people can speak out within one minute (even if) they only know a little bit."

Notice how the author effectively uses quotes from the interview data to illustrate her generalizations. Bracketed words in quoted responses reflect changes that the author or a translator had made in order to clarify what the women were talking about. In reporting interview data, changing words within brackets is both appropriate and helpful, provided that it does not change the meanings of participants' statements.

Note: From *The Immigration Experiences, Acculturation, and Parenting of Chinese Immigrant Mothers* (pp. 109–113, 116–119, 146–149), by C. Y. Y. Leung, 2012, unpublished doctoral dissertation, University of Maryland, Baltimore County. Reprinted with permission.

APPLICATION EXERCISES

Similar exercises are available on Pearson MyLab Education.

1. Shannon is a college student who is conducting a qualitative research study of mentoring among high school choir students. Shannon received permission to sit in the back of the choir room, and she took extensive field notes on social interactions between choir members during practices. She also videotaped interviews with each choir member, as well as choir performances.

   i. Describe the first steps Shannon must take before she can begin the analysis of her data.
   ii. Describe useful strategies that would help Shannon to organize her data in a preliminary way that will enable her to locate them more easily as she proceeds.
   iii. Explain how Shannon might identify preliminary categories that are likely to be helpful in coding the data.

2. In a qualitative study, the interpretation of data is influenced by the researcher's biases and values to some extent, reflecting the notion of researcher-as-instrument. Qualitative research requires the researcher to make significant decisions and judgments throughout the data analysis process, not only about what strategies to use in general but also about which data are most likely to be noteworthy and how to evaluate and code specific pieces of data. Consider the following questions about strategies the researcher can use to minimize bias in the final analysis of data.

   Imagine that a political science researcher decided to conduct a qualitative case study of the political campaign of a candidate for statewide office whom he believed to be running an unusually honest and ethical campaign. In the 6 months between the primary and the general election, the researcher observed the weekly community organizing meetings of the campaign staff, collected all of the press releases distributed to members of the media, videotaped all of the public statements made by campaign spokespeople, and conducted structured interviews with each member of the paid campaign staff, as well as with the candidate.

   i. As an alternative to maintaining complete objectivity—a goal that is probably impossible—describe at least three strategies this researcher can use to strive for balance and fairness in his data analysis.
   ii. Describe how the participants in this qualitative study might themselves introduce bias in response to the researcher.
   iii. Explain how and why this researcher should deal with his potential biases when the time comes to write up his final research report.

FOR FURTHER READING

Altheide, D. L., & Schneider, C. J. (2013). *Qualitative media analysis* (2nd ed.). Los Angeles, CA: Sage.

Boland, A., Cherry, G., & Dickson, R. (2014). *Doing a systematic review: A student's guide.* London, England: Sage.

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). Los Angeles, CA: Sage.

Corbin, J., & Strauss, A. C. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Los Angeles, CA: Sage.

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Los Angeles, CA: Sage.

Denzin, N. K., & Lincoln, Y. S. (2013). *Collecting and interpreting qualitative materials* (4th ed.). Los Angeles, CA: Sage.

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods, 5*(1), 1–11.

Glaser, B. (1992). *Basics of grounded theory analysis.* Mill Valley, CA: Sociology Press.

Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews* (2nd ed.). London, England: Sage.

Grbich, C. (2007). *Qualitative data analysis.* Los Angeles, CA: Sage.

Guest, G., MacQueen, K. M., & Namey, E. E. (2012). *Applied thematic analysis.* Los Angeles, CA: Sage.

Harding, J. (2013). *Qualitative data analysis from start to finish.* London, England: Sage.

Hill, C. E., Knox, S., Thompson, B. J., Williams, E. N., Hess, S. A., & Ladany, N. (2005). Consensual qualitative research: An update. *Journal of Counseling Psychology, 52,* 196–205.

Hill, C. E., Thompson, B. J., & Williams, E. N. (1997). A guide to conducting consensual qualitative research. *The Counseling Psychologist, 25,* 517–572.

Kuckartz, U. (2014). *Qualitative text analysis: A guide to methods, practice, and using software.* London, England: Sage.

Martin, V. B., & Gynnild, A. (Eds.). (2011). *Grounded theory: The philosophy, method, and work of Barney Glaser.* Boca Raton, FL: BrownWalker Press.

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.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Los Angeles: Sage.

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide.* Oxford, England: Blackwell.

Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). London, England: Sage.

Spradley, J. P. (1980). *Participant observation*. Belmont, CA: Wadsworth.

Tashakkori, A., & Teddlie, C. (Eds.). (2010). *Mixed methods in social & behavioral research.* Los Angeles, 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, NY: Guilford Press.

Wolcott, H. F. (1994). *Transforming qualitative data: Description, analysis, and interpretation.* Los Angeles, CA: Sage.

Yin, R. K. (2011). *Qualitative research from start to finish.* New York, NY: Guilford Press.

This page intentionally left blank

Chapter 13 Planning and Preparing a Final Research Report

Ultimately, what you put on paper and how you put it there reveals your knowledge about your topic, the quality of your thinking, and your standards of excellence with greater clarity than anything else you do.

Learning Outcomes

13.1 Describe several essential components of a good research report, including discussions of (a) the research problem or question, (b) study design and data-collection methods, (c) the obtained data and their analyses, (d) interpretations of the data, (e) possible weaknesses of the study, and (f) relevance of the study to a broader context.

13.2 Explain how you can best maintain your academic integrity as you write a final research report.

13.3 Describe the possible contents of front matter and end matter in a research report.

13.4 Describe potentially effective organizational schemes for quantitative and qualitative research reports.

13.5 Identify effective strategies for (a) writing a good report within a reasonable time frame and (b) critiquing your own or someone else's final or almost-final report.

13.6 Discuss various means by which you might succinctly describe your research study and findings to a broad audience.

Bringing a research effort to its rightful conclusion involves writing a report that is faithful to the data and 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 address the research problem or question. Like the research proposal, it makes no pretense at being a fine work of literature. It must, however, be well organized and easy to understand so that readers can quickly grasp what the researcher has done and found. It must also be flawless in its sentence and paragraph structures, punctuation, and spelling. The research document one writes is a clear reflection of one' 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.

Also, university guidelines aside, different academic disciplines tend to adhere to different style manuals in research reports. Differences among styles are most noticeable in the

formats used for citations and reference lists. For example, in this book we use APA style—the style recommended by the American Psychological Association. Alternatively, scholars in certain other disciplines tend to use Chicago style, CSE (Council of Science Editors) style, or MLA (Modern Language Association) style. You can find information about all four of these styles in Table 3.3 in Chapter 3.

Surfing the Internet for Writing Assistance

Many Internet websites offer assistance on the nitty-gritty details of different styles. The rightmost column of Table 3.3 in Chapter 3 provides 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 Bing 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 websites offer more general suggestions for writers. One excellent resource is the Online Writing Lab (OWL) at Purdue University (owl.english.purdue.edu). If you go to OWL's website and click on "Site Map," you can find links to discussions of a wide range of topics related to writing—for instance, how to argue for a particular point of view, how to enhance sentence clarity, when to use various pronouns and verb tenses, and in what circumstances to use hyphens. 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 Global database, an online resource to which many university libraries subscribe. Also, look once again at some of the research articles and reports you previously read for your literature review; these, too, might give you ideas about how to structure and write your own report.

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 reasonably conclude that the report's author is *not* a writer whose style you want to emulate!

ESSENTIAL ELEMENTS OF A RESEARCH REPORT

Any research report should achieve six main objectives. More specifically, it should:

1. Give readers a clear understanding of the research problem or question and why it merited an in-depth investigation.
2. Describe exactly what methods were used in an effort to address the research problem or question.
3. Present the collected data with enough precision and thoroughness that they can substantiate the interpretations and conclusions that will follow.
4. Interpret the data for readers and demonstrate either how the data resolve the research problem or why they *don't* completely resolve the problem.
5. Alert readers to possible weaknesses of the study—for instance, what its design flaws may have been and what assumptions and biases might have affected results and interpretations.
6. Conclude by summarizing the findings and connecting them to contexts beyond the study itself—for example, relating findings to current theoretical perspectives or drawing implications for future policies or practices.

In the following sections, we discuss each of these matters. Then, in a subsequent section, we remind you once again about the importance of being completely truthful and forthcoming about what you have accomplished.

Explanation of the Research Problem or Question

The beginning pages of any research report serve an essential purpose: to create a meeting of minds between the writer and readers of the report. Typically the first section of a report provides a statement of the research problem or question and any other information that readers will need to understand it. After a few introductory comments that provide background and a rationale for your study—perhaps in a few sentences or paragraphs—your report should set forth clearly and unmistakably the overall problem or question you have addressed. Readers should be able to comprehend *from the report alone* what your problem was and why it was an important one to investigate.

Often, an appropriate subheading can draw readers' attention to the research problem. If you have divided your problem into subproblems, these should be presented following the statement of the problem and announced with proper subheadings. And, of course, any *a priori* hypotheses should be clearly stated early in the report. By presenting the problem, its subproblems, and any hypotheses, you give readers a clear and complete understanding of the principal thrust of the research effort.

Also, you should define any terms that may have multiple meanings or in some other way might be ambiguous (see the section "Defining Terms" in Chapter 2). For a meeting of minds, it is imperative that you and your readers share a common understanding of key concepts around which the research effort has revolved.

The extent to which related literature is presented in an introductory 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 Study Design and Data-Collection Methods

The general *design* of the study should be clear early in the report. 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 a survey, a longitudinal study, a single-group time-series study, a 2-by-2 factorial design, an ethnography, a grounded theory study, or some combination of approaches.

Almost without exception, a research report should include a specific section labeled "Methods," "Methodology," or something similar. The research setting, sample, assessment strategies, and procedures should be described with as much precision as possible. Ideally, readers should know—from this description alone—exactly what was done, to the point where readers could replicate the study and, in most cases, get similar results.

Presentation of the Data and Their Analyses

After readers fully understand what the problem was and the manner in which it was investigated, the next question is, "What is the evidence?" You have collected a large body of data. You have also analyzed the data in one or more ways, perhaps by performing appropriate statistical computations or perhaps by systematically coding interview responses. In most quantitative research reports, the obtained results are described in a single section appropriately called "Results" or "Findings," with interpretations to follow in a separate section. In some qualitative reports, however, the data and the researchers' interpretations of them are woven together in one or more topic-specific sections following the "Methods" section.

If you have performed statistical analyses of quantitative data, you should include your rationale for employing the particular procedures you have used. It's important for readers 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.

Many qualitative data aren't appropriate for statistical analysis, of course. Accordingly, you should present your qualitative data in such a way that they speak for themselves. As previously mentioned in both Chapter 4 and Chapter 8, qualitative researchers often engage in *thick description,* presenting their data in such detail that readers can see for themselves what kinds of observations have been made.

So that readers don't get lost in the data presentation, it's often helpful to begin the discussion of the data with an *advance organizer* in which you lay out the overall organization of your data presentation. We refer you back to the Guidelines section "Writing to Communicate" in Chapter 1 for a description of advance organizers.

Usually data are presented as they relate to (a) the overall research problem or question, (b) its subproblems, and (c) if applicable, any *a priori* hypotheses. One logical approach is to devote a separate section to each subproblem and its pertinent quantitative and/or qualitative data.

Yet you need not limit your discussion of results only to the problem, subproblems, and any hypotheses you have posited. In fact, 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 suggestion to "Check for consistency" in Chapter 6). If you have used two or more raters to code your data in some way, you should describe both the interrater reliability and the steps taken to maximize it (see the section "Enhancing the Reliability of an Assessment Strategy" in Chapter 4). 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 hadn't initially considered.

Your data should be presented thoroughly and, of course, accurately. In many cases, some data can be synthesized 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; for instance, it might be a graph, chart, map, flowchart, photograph, or drawing. Sometimes a picture really is worth a thousand words. However, we caution you to include graphics only when (a) the data are sufficiently complex to warrant one or more visual depictions and (b) you can represent the data without distorting their nature in such a way that you mislead readers. For an excellent discussion of what *not* to do when creating graphs and other visual representations of research data, we refer you to Wainer's (1984) classic article "How to Display Data Badly" in the journal *The American Statistician.*

Many computer software programs can help you create tables and figures; for example, see Microsoft Word (for tables), Microsoft Excel (for graphs), Inspiration (for flowcharts), and Maptitude 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 university's graduate school office for its own requirements about such things.[1]

[1]When 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.

Descriptions of data in quantitative studies are often 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.

Regardless of how you present your data, it's imperative that you present them in enough detail that they will support 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.

Interpretations of the Data

All too frequently, novice 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 underlying meaning of the data, no resolution of a research problem or question is possible.

In interpreting the data, however, you must be careful not to go *too far beyond* the data. Beginning researchers often lose sight of what they've actually found; they're so enthusiastic 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 schoolchildren 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 have previously made in several earlier chapters bears repeating one more time: *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 children 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 something beyond the shadow of a doubt. For example, even the most powerful inferential statistics are based only 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, possibly 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 you must never let your enthusiasm interfere with your open-mindedness—and as much impartiality as you can muster—as you interpret and draw conclusions from your data. The answer to your research problem or 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 your analyses, tables, graphs, illustrative examples, and other exhibits solidly support what those underlined words declare. That is good research.

What if the data *don't* support any expectations you might have had or any predictions you might have made? Does this mean that your theoretical or conceptual framework wasn't appropriate or that any *a priori* hypotheses you formed were wrong? Not necessarily. Look once again at your methodology and data analyses to see if you can identify one or more weaknesses in what you have done. Perhaps one of your assessment strategies had lower validity or reliability than you had anticipated and therefore didn't yield accurate and dependable measures of a critical factor in your study. Maybe you gave participants misleading instructions or asked them misleading questions. Perhaps your statistical analyses lacked power—maybe your sample was too small or your assessment strategies too unreliable—and so you made a Type II error.[2]

At the same time, maybe your expectations and predictions *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 for the results you obtained, perhaps in the form of alternative problems or questions that future research efforts might investigate.

In the final analysis, *your data must speak for themselves.* You are only their mouthpiece. You may not like what the data say. They may not confirm your fondest hopes or support your preconceived notions, but a researcher must be 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.

Identification of Possible Weaknesses of the Study

You should report any weaknesses in your study that may have influenced its outcome. For example, throughout your report, you must be up front about any assumptions that may have influenced your methods, analyses, and interpretations. You must also alert readers to personal, social, political, or philosophical biases that may have limited your ability to study the research problem with complete objectivity. But acknowledging your biases involves first *consciously thinking about* what your biases might be—a process that qualitative researchers call *reflexivity* (see Chapter 8)—and then speculating about how these may have affected what methods you used, what data you collected, and how you interpreted your results.

Delimitations and limitations of the study should be clearly set forth as well. All who read the research report should know precisely how far the research effort extended and what its boundaries were. Into what relevant areas did the research effort *not* inquire? What aspects of the problem were *not* studied? What methodological flaws emerged during the project? Delimitations—what a researcher never intended to do—are typically stated near the beginning of the report. Limitations—weaknesses that may cast doubts on results and interpretations—can be mentioned wherever they are most relevant, but they definitely *should* be mentioned.

For guidance on pinning down such things, we refer you back to the sections "Stating Assumptions" and "Identifying Delimitations and Limitations" in Chapter 2, as well as to the concept of *researcher-as-instrument* in the Practical Application feature "Ensuring That Qualitative Data Collection Yields Credible, Trustworthy Data" in Chapter 8.

[2]A *Type II error* occurs when the probability is too high that an obtained result was simply due to chance (see the section "Making Errors in Hypothesis Testing" in Chapter 11).

Summary and Connections to a Broader Context

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" or simply "Discussion."

In this final section, you should clearly summarize your findings and interpretations relative to your overall research problem or question and to any subproblems and hypotheses. 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. Readers should be able to see the entire 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.

In addition, you need to address the question, *So what?* In what way does the study contribute to our collective knowledge about some aspect of our physical or biological world or of human experience? The connections you make here might take one or more of several forms. You might compare your findings with other, previously reported research findings and point out similarities and dissimilarities. You might argue that your results either support or disconfirm an existing theoretical or conceptual framework. You might draw implications for instruction, social services, medicine, animal welfare, or business. You might offer suggestions for how future research could further advance the frontiers of knowledge about your topic. And if you have developed new methodological strategies or assessment tools to study your problem, you might justifiably argue that they are, in and of themselves, valuable contributions to the research literature in your field.

MyLab Education Self-Check 13.1

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 a few basic facts 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 excluded from your research sample and explaining why you excluded them
- Describing the limitations of your assessment strategies—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 predictions
- Explicitly identifying any potential confounding variables that may cast doubt on conclusions about cause-and-effect relationships

We must emphasize a point we have 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 the facts or mislead readers in any other way—no matter how well-intentioned those actions might be—is to potentially lead a community of scholars astray in their quest for knowledge.

MyLab Education Self-Check 13.2

FRONT MATTER AND END MATTER

The topics described in the preceding section typically appear in the main body of a research report. But 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 now 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 usually comes first; this includes the author and, typically, a university affiliation and date. In a thesis or dissertation, a page for signatures of the faculty advisor and research committee also appears at the beginning. Next are an abstract, possibly a dedication page (this is optional), acknowledgments, a table of contents, lists of any tables and figures, and, if desired, a preface.

In some instances, a copyright page is also included in the front matter. **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 don't register it with the United States Copyright Office. Nevertheless, registering it often provides reassurance and peace of mind. As this book goes to press in 2017, the fee for filing a research report with the copyright office is either $35 or $85, depending on whether the application is filed electronically or on paper. You can get more information about U.S. copyright laws and procedures at the Copyright Office's website (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 has addressed a topic of concern to them and thus merits their further attention. The abstract you write is likely to be included in one or more online databases (e.g., ProQuest Dissertations & Theses Global) available at many research libraries around the world. Thus, it's critical 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 or question 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 provide enough information for readers to 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 other 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 about your research topic.

A brief **acknowledgments** section graciously recognizes the assistance of those people through whose kindness the research effort has been possible. These individuals may include those who introduced you to data sources that helped you in your project or those who guided your study and gave counsel or support—perhaps an academic dissertation committee, peers who have helped you with data collection or analysis, a meticulous proofreader, and key family members. 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. Any lengthy research report should include a **table of contents**. The table of contents is a bird's-eye view of what the document contains, how it is organized, and on which page each section and subsection begins. Often following the main table of contents are two more specific ones, one for tables and another for figures that appear throughout the report.

Footnotes and Endnotes

Generally, **footnotes** (appearing at the bottoms of relevant pages) and **endnotes** (appearing at the end of the text) 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. For example, researchers who adhere to Chicago style use footnotes to acknowledge publications, conference presentations, and other resources that have provided a basis for their statements.

Second, footnotes and endnotes 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 copyright holder (typically the publisher or author). Immediately after including the material, you can indicate its original source in a footnote or endnote, followed by the words "Reprinted by permission" or other wording stipulated by the copyright holder.

A third important function of footnotes and endnotes 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 footnotes or endnotes becoming overly long and involved, sharpen your ideas and integrate them into the body of your 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's important that reference information be both complete—it should include references for *all* of your citations—and accurate. The reference list isn't 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 perused 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 (typically in accordance with a particular style manual), and you should follow the format that your institution or your discipline requires. Furthermore, you should apply that format consistently throughout your reference list. Most of the bibliographic software programs described in Chapter 3 (e.g., EndNote, RefWorks, Zotero) can quickly format your references in whatever style you need.

One widely used style in reference lists is that of the American Psychological Association (APA)—for instance, it's the style we use in this book—and so we describe it in some detail here for illustrative purposes. Following are key elements of each entry in an APA-formatted reference list.

Author In an APA-style reference list, the author's name appears with the surname first, followed by the author's first initial and any middle initials. When multiple authors are involved, the names are separated by commas. Commas are always used between the names (even between only two names), and an ampersand (&) is used before the last name in the list. Note, however, that, especially in the biological and medical sciences, some research articles have a large number of coauthors. In APA style, the first six authors are listed, followed by a 3-dot ellipsis (. . .) and then the final author's name. For example, see the entry for Abraham and colleagues (1941) in this book's reference list.

Date of Publication After the names of one or more authors is the year of publication in parentheses, followed by a period. Magazines, newsletters, newspapers, and presentations also include the month and, if necessary to pin down the particular issue, the day.

Occasionally a reference list includes two or more sources by the same author(s) in the same year. In such instances, lowercase letters *a, b,* and (if needed) subsequent alphabet letters follow the year inside the parentheses. For example, if you wanted to cite two sources written by sole author Deanna Kuhn in 2001—in this case, a journal article and a chapter in an edited book—you would list them in alphabetical order by title like this:

Kuhn, D. (2001a). How do people know? *Psychological Science, 12,* 1–8.

Kuhn, D. (2001b). Why development does (and does not) occur: Evidence from the domain of inductive reasoning. In J. L. McClelland & R. S. Siegler (Eds.), *Mechanisms of cognitive development: Behavioral and neural perspectives* (pp. 221–249). Mahwah, NJ: Erlbaum.

Any citations in the text would then be either "Kuhn, 2001a" or "Kuhn, 2001b."

Title of the Work In APA style, the title of the article, book, or other source follows the publication year. If you are referencing a journal article using APA style, the title of the article is *not* italicized, but the title of the journal in which it appears *is* (e.g., see the first of the two Kuhn citations just presented). The title of a book is always italicized. So, too, is the title of a conference presentation or doctoral dissertation (e.g., see the entry for Laura Lara's [2009] dissertation in this book's reference list).

Be sure to pay attention to the rules for capitalization in whatever style manual you are using. Can you determine what the APA's rules are from the entries in this book's reference list? As you should notice, the first word in a book, article, or presentation title is the only one capitalized unless (a) the word is a proper noun or proper adjective or (b) it follows a colon. In contrast, 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 such information is formatted in this book's reference list. All redundancy is eliminated; there are no extra words such as *volume, issue,* and *pages.* By their specific locations 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 names of publishers are usually short and succinct; with a few idiosyncratic exceptions, they exclude 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 information includes either (a) the document's **Digital Object Identifier (DOI)** or (b) the Internet address at which the document was found—that is, the **Uniform Resource Locator (URL)**. We have previously described both DOIs and URLs in Chapter 3.

Whenever you have a choice, we urge you to include a document's DOI rather than the URL at which you found it. As you should recall from Chapter 3, a document's DOI should remain constant for the foreseeable future, whereas its actual location on the Internet might change over time. 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 print. 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. https://doi.org/10.1177/0956797611400615

Readers who see this entry in a reference list can track down the document by typing the entire DOI number—beginning with "doi.org/"—in an Internet browser.[3] Alternatively, readers can go to the International DOI Foundation's website (doi.org) and then enter the article's DOI—in this instance, 10.1177/0956797611400615—in the "Submit" box.

DOIs are a fairly recent development for Internet-based documents; hence, you are likely to see them only for documents posted since the year 2000, and only for *some* of those documents. Thus, including the URL at which you originally found a document is an important backup strategy. For example, in APA style, a 2002 article by Amrein and Berliner in the online journal *Education Policy Analysis Archives* 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), 1–71. Retrieved from http://epaa.asu.edu/ojs/article/view/297/423

And a 2008 report written by Keeter and colleagues for the Pew Research Center would be referenced using this format:

> 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 from Pew Research Center for the People and the Press website: http://people-press.org/files/legacy-pdf/391.pdf

As this book goes to press in 2017, these two documents can still be found at their original URLs.

You should note that DOIs and URLs can be quite lengthy—often so long that they spill over onto a second line on the page. If you need to split one up, you should do so before

[3]We should point out that "doi.org/10.1177/0956797611400615" is actually a URL—but in this case it's a relatively permanent one. We make this point only in a footnote here so as not to unnecessarily confuse our readers.

a period (.), a forward slash (/), or a hyphen (-) within the sequence of letters, numbers, and punctuation marks.

We have presented only a few examples of how entries in a reference list might be formatted, and in all of our examples we have used APA style. A complete reference list is apt to include many types of resources—books, articles, government reports, conference presentations, posters, videos, website pages, blog postings, and so on—and 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 reference list, you can include any appendices that might help readers more completely understand your research study but aren't essential to readers' general comprehension of your study. A rule of thumb is that *material appearing in an appendix enables readers to go further in understanding the method and/or results if so desired.* For instance, an appendix might include informed consent letters, questionnaires and other assessment instruments, response sheets, field notes, preliminary and final coding categories, statistical computations, or extensive data tables.

In reporting research, nothing is hidden. All of the data are laid out. The researcher's integrity is thereby preserved, and the results and conclusions of the study can be readily verified.

MyLab Education Self-Check 13.3

ORGANIZING A RESEARCH REPORT

Research reports for most quantitative studies are similar in their organizational format. After any necessary preliminary pages (e.g., title page, acknowledgments, table of contents), they typically have five major sections: an introduction (which includes the research problem, an explanation of its importance, assumptions, definitions of terms, etc.), a review of the related literature, a description of the methodology, a presentation of findings and specific interpretations (e.g., whether the data do or do not support any *a priori* hypotheses), and general conclusions (including implications and suggestions for future research). Reports of qualitative and mixed-methods studies are less consistent; their organizational schemes are apt to depend somewhat on the nature and design of the studies themselves. Virtually all research reports include a reference list. Appendices are less common, although they are usually important components of dissertations and theses.

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 primarily quantitative, quasi-experimental study with an embedded qualitative component; 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 within each chapter.

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

MyLab Education Self-Check 13.4

MyLab Education Application Exercise 13.1: Outlining a Research Report

WRITING—AND FINISHING!—A REPORT

A final research report is exactly that—a *report.* You need to describe what you have done over the course of your research effort. In the process, you need to acquaint your readers with your problem or question, the means you used to gather data that might address that problem or question, the ways in which you analyzed and interpreted your data, and the ultimate conclusions you reached.

Remember that your report is *you.* Whether or not you intend it to do so, a report can say a great deal about you to your mentors and professional colleagues—not only about your abilities as a researcher and scholar but also about your diligence in completing a project in a timely manner. In the following two practical applications, we offer strategies for writing a clear, coherent report and for completing your report within a reasonable timeframe.

PRACTICAL APPLICATION Writing Your Final Report

As you begin your final report, we urge you to revisit five of the Guidelines sections included in earlier chapters: "Writing to Communicate" and "Using the Tools in Word-Processing Software" (both in Chapter 1), "Writing a Clear and Cohesive Literature Review" (in Chapter 3), and "Writing the First Draft" and "Revising Your Proposal" (both in Chapter 5). Most of the guidelines in these sections are as applicable to a final research report as they are to a research proposal or any other piece of scholarly writing. At the risk of repeating a few suggestions presented in those earlier Guidelines sections, we now offer some general recommendations to keep in mind as you begin to write.

GUIDELINES Writing a Clear, Coherent Report

Writing a research report isn't just a matter of mindlessly putting on paper the things you have done and discovered in your research project. Rather, writing a good report involves consciously and conscientiously striving to *communicate* what you have done. The following guidelines can greatly enhance the effectiveness with which you tell readers about your research project.

1. ***Choose an appropriate style for your intended audience.*** Most research reports use a somewhat formal and impersonal style. There are exceptions to this rule, however; for example, ethnographic researchers sometimes describe their findings in a personal, storytelling manner. Perusing research reports in your own academic discipline can give you a good sense of the writing style that is most prevalent—and thus most acceptable—in your field.

As briefly noted in Chapter 5, various disciplines and their preferred style manuals have different standards regarding active versus passive voice. For example, the APA (2010) *Publication Manual* 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).

Also, although writing in the third person was the preferred style throughout most of the 20th century (e.g., "The researcher analyzed the data"), increasingly researchers are using a first-person style 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 didn't take part in your own study! Whichever style you choose, you should be consistent in using it 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.

2. ***Create and follow a logical overall structure.*** To facilitate readers' "journey" through your research report, you should create a logical, predictable structure for the report. You should also provide regular guideposts in the forms of headings, transitional words and paragraphs, and other means of helping readers follow your train of thought.

After you have written 10 or 20 pages of your report, go back and read your headings and subheadings and the paragraphs following each one. 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 and text show a logical progression from one idea to the next? They should.

3. ***Be clear and precise in your wording.*** A research report must be crisp with clarity and 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 report should present what you have thought, done, and discovered in a straightforward manner. Show how your data resolve your subproblems and how the subproblems help to resolve your main problem or question. Lead your readers through your own thought processes step by step. Such tactics provide evidence that you have approached your entire research endeavor in a thoughtful, systematic manner.

4. ***Use appropriate verb tenses.*** In writing a final report, some novice researchers simply cut and paste large portions of their research proposals into their final reports, the unfortunate result sometimes being that the report describes what "will" be done when it already *has* been done. Attentiveness to such details as verb tenses is yet another sign of a good, conscientious researcher and writer. Past tense is most appropriate for describing things that have already happened. The APA (2010) *Publication Manual* suggests that simple past tense be used for things that occurred at a specific time (e.g., "Participants were randomly divided . . .") and present-perfect tense be used for actions and events that either happened over a lengthy period or are continuing into the present (e.g., "Theorists have suggested"). When drawing conclusions and suggesting implications, however, present tense is appropriate.

The grammar checker in your word-processing software may enable you to check for inconsistencies in your use of active versus passive voice or past versus present tense. Exactly where you find such options in your program depends on whether you are using a PC or Macintosh computer and which version of a software program you are using. Our best advice is to search the Internet for instructions relative to your specific platform and software.

5. ***Strive for as much objectivity and impartiality as possible in what you say and how you say it.*** Although it's inevitable that certain personal, theoretical, or philosophical biases will influence what you say, you should actively strive to avoid making judgments as much as you can. Even small changes in wording can make a difference. For example, when describing your findings in an ethnographic study, 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.

6. ***Regularly summarize what you have just said.*** After your lengthy and intensive involvement in your study, you're apt to have a keen understanding of your master plan and of the relationships and contributions of various components of the study to your overall research endeavor. Your entire research project—the problem, the data and their organization, the relationships and interrelationships—are likely to be crystal clear in your mind. Your readers, however, won't be so fortunate. As they proceed through your report, they may need to stop occasionally to consider and reconsider how a certain piece fits into the total investigation. Thus, especially in a lengthy research report such as a dissertation, you can help readers immensely by providing a brief summary at the close of each extended discussion.

A report that merely rambles on and on, going blindly forward from one topic to another, can lead to psychological numbness, bewilderment, and confusion. Frequent summaries help to minimize such reader disorientation.

7. ***Submit a neat, clean final copy that strictly adheres to required formats.*** 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 document's 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 perfectly aligned columns.

Check and double-check for spelling and grammatical errors. Keep an eye out for possible misuse of homonyms, such as using *their* when you mean *there* or using *affect* when you mean *effect* (see Chapter 5). Lack of attention to such details—some of which may seem trivial and picayune to you—can leave lasting impressions on others.

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 isn't something you can whip up in a week or two. 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 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 Pinning Down and Adhering to a Workable Schedule

The suggestions we offer here have emerged from our own experiences in writing research reports and other lengthy documents, including our own doctoral dissertations—and also this textbook!

1. ***Identify small, easily accomplishable goals within the overall project.*** A lengthy research report 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 few days' time at most.

2. ***Write easier sections first.*** No rule says you have to write the sections of a research report in order, starting at the beginning and working your way to the end. Most researchers find it easier to write some parts than others. For instance, writing a Methods section can be fairly simple and straightforward. What were the general characteristics of your sample? How did you recruit human participants? What assessment strategies did you use? What procedures did you follow? And after you have conducted appropriate statistical analyses of quantitative data, creating one or more tables to summarize them should be a relatively easy task. Once you have completed easier parts of the report, you may begin to see a glimmer of light at the end of your report-writing tunnel and be reinvigorated to tackle more challenging sections.

3. ***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? (One of us authors has learned from experience that she can usually write only

about 8 to 10 double-spaced manuscript pages a day. After that, she's essentially brain-dead and should wait until the following morning before she continues.)

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'll never stick to your schedule; you'll be doomed to failure even before you start.

4. ***Reward yourself each time you reach one of your goals.*** Give yourself a treat of some sort after you successfully finish each significant piece of your report. Watch a movie, read a magazine or short mystery novel, clean the house, play a few games on your computer, surf the Internet—whatever you need to do to get refreshed and ready to tackle the next task on your schedule.

5. ***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.

6. ***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 doctoral dissertation. An editorial review board carefully scrutinizes any manuscript submitted to a scholarly journal. A formal 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: You're mentally and emotionally very close to your research project and, at the end, are equally close to the report you have written about the project. So close, sometimes, that omissions, errors, and logical inconsistencies that may be blatantly obvious to others aren't at all obvious to you. Furthermore, other people might have useful ideas about better ways to organize a discussion, suggestions about additional analyses that may shed further light on the data, or new sources of literature that are potentially 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 acceptance for publication 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 Final Research Report

We have previously given you a checklist for "Evaluating a Research Article" in Chapter 3. But by now you should be in a position to answer many more questions—and more *specific* questions—about the quality of a research report than you would have been able to answer

early in the book. The following checklist can both help you evaluate other researchers' final reports and serve as a guide as you assess your own writing.[4] Admittedly it's a very long list, but we remind you that a good researcher must consider a wide range of issues related to any research endeavor.

CHECKLIST

Criteria for Critiquing a Research Report

	YES	NO
STEP 1. THE MAIN RESEARCH PROBLEM OR QUESTION		
Is the main problem or question clearly and concisely stated, perhaps in an explicit purpose statement?	_____	_____
Is the problem or question adequately narrowed down into a researchable issue?	_____	_____
Is the problem or question important enough to warrant a formal research investigation?	_____	_____
STEP 2. SUBPROBLEMS AND ANY RELATED HYPOTHESES		
Has the main problem or question been explicitly broken down into subproblems and, if so, does each subproblem warrant data collection and interpretation in its own right?	_____	_____
Are any *a priori* hypotheses presented and, if so, are they clearly related either to the main problem/question or to one or more of its subproblems?	_____	_____
If there are *a priori* hypotheses, can they be either supported or discredited by the collection of new data?	_____	_____
STEP 3. THEORETICAL OR CONCEPTUAL FRAMEWORK		
Is the theoretical or conceptual framework clearly applicable to the overall problem or question (as opposed to being a "stretch" in which the framework is only marginally relevant)?	_____	_____
Are concepts central to the framework adequately defined, and are relationships among these concepts clearly identified?	_____	_____
If applicable, do the *a priori* hypotheses flow logically from the theoretical or conceptual framework?	_____	_____
STEP 4. LITERATURE REVIEW		
Is the literature review logically organized?	_____	_____
Are classic, widely cited theoretical perspectives and empirical findings relevant to the topic included in the literature review?	_____	_____
Does the review include critiques of relevant studies and syntheses of similar findings?	_____	_____
Are gaps in knowledge about the research problem identified?	_____	_____
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?	_____	_____

[4]A similar checklist that appeared in earlier editions of this book was loosely based on a list of questions proposed by Beck (1990). However, Beck's questions were slanted heavily in the direction of quantitative research; hence, we have substantially revised the list in this 12th edition to be equally applicable to qualitative and mixed-methods research.

	YES	NO
STEP 5. RESEARCH DESIGN		
Is the research design clearly described?	_____	_____
Is the design appropriate for the research problem or question?	_____	_____
Does the research design adequately address issues related to the credibility and, if applicable, generalizability of the study?	_____	_____
If causal relationships are hypothesized, does the design credibly enable conclusions about these relationships?	_____	_____
STEP 6. SAMPLING		
Is an appropriate method used for selecting the sample?	_____	_____
Are the criteria for selecting the sample clearly identified?	_____	_____
Is the sample size adequate?	_____	_____
Is any sampling bias in the method acknowledged?	_____	_____
STEP 7. DATA-COLLECTION METHODS		
Are data-collection methods appropriate for the research problem or question?	_____	_____
If applicable, are assessment instruments adequately described?	_____	_____
Do assessment strategies appear to have reasonable validity and reliability for their purposes in the study?	_____	_____
STEP 8. DATA ANALYSES		
Is the discussion of results clearly and logically organized?	_____	_____
Are analyses appropriate for each set of quantitative and qualitative data? For example, is each analysis of quantitative data appropriate for the type of measurement scale (nominal, ordinal, interval, ratio) that the data reflect?	_____	_____
Is each analysis relevant to answering the research problem/question or to one of its subproblems?	_____	_____
Are tables and figures clear and understandable?	_____	_____
STEP 9. INTERPRETATION AND DISCUSSION OF THE FINDINGS		
Does the investigator clearly distinguish between actual results, on the one hand, and interpretations of those results, on the other?	_____	_____
Are interpretations based on the data obtained?	_____	_____
Are the findings discussed in relation to previous research and to a theoretical/conceptual framework?	_____	_____
Are all generalizations warranted and defended?	_____	_____
Are limitations of the results and interpretations identified?	_____	_____
Are implications of the results discussed?	_____	_____
Are recommendations for future research identified?	_____	_____
Are all conclusions justified and defensible?	_____	_____

MyLab Education Self-Check 13.5

MyLab Education Application Exercise 13.2: Critiquing a Report

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, we strongly urge you to seek a wider audience.

One easy way to gain a broader audience for a thesis or dissertation is to publish it in the online database ProQuest Dissertations & Theses Global (proquest.com). In fact, some universities *require* doctoral students to submit their final dissertations to ProQuest. Also, you might post details about your study in topic-specific online discussion forums, bulletin boards, and other electronic discussion groups. But probably the two most effective ways to get the word out about your study—and to enhance its credibility—are conference presentations and journal articles, in large part because these two means of communication are apt to require *peer review and approval* of your work.

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 focused on particular interest areas (e.g., family violence, Piaget's theory of child development). Organizers of many of these conferences eagerly seek presentations (a.k.a. *papers*) from new researchers as well as from more experienced ones. Some conferences also include poster sessions, which (as the word *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 and answer questions about a study's design, methods, and findings.

If you would like to present a paper or poster at a professional conference, you will probably need to submit a 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 of these 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 that you're in good company won't make your public-speaking jitters go away, 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 nerves 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 3 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 discovered. Instead, you should 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 or question and, if applicable, your *a priori* hypotheses
- The 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 or question and to any hypotheses
- Your interpretations of and conclusions drawn from your data

Many posters also include a one-page abstract immediately after the title and author(s), 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 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 will probably be available at the front of the presentation room, and click on a mouse or wireless clicker to advance to successive slides.[5] 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.

We urge you not to clutter up your presentation or poster with too many visual effects, such as distracting and irrelevant images and animations. However, simple images and animations—for instance, having individual bullet points "march" across the screen as you introduce them (which PowerPoint lets you do)—are quite appropriate and can catch and hold people's attention.

[5]Most IBM and IBM-compatible personal computers (PCs) laptops have standard outlets for hooking up to an LCD projector. In our experience, however, Macintosh laptops require special adaptors; check with your local Apple retailer to be sure you have the appropriate adaptor for your own laptop.

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 limit. It's helpful, too, to review your speaking notes within an hour or two of your presentation so that they're fresh in your mind. 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 its "Print" feature, PowerPoint lets you specify how many slides you print on a single page; 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 the manuscript. Alternatively, you might direct interested audience members to a website at which they can download the paper or poster. And in some instances—this is most often the case when someone hasn't sufficiently 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'll 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, affiliation, and mailing and e-mail addresses.

7. ***After the conference, post your paper in the sponsoring organization's repository of conference papers.*** For a variety of reasons, some members of a professional organization either won't be able to attend the conference at which you're presenting or won't be able to attend your particular session. In reviewing the conference program, some of these individuals might be quite intrigued by your topic and frustrated that they can't see or hear you in person. To accommodate such individuals, many associations now either encourage or require presenters to submit electronic copies of their papers to an online repository of conference papers. We urge you to take this simple follow-up step, as it can definitely expand your audience and possibly enable other scholars to incorporate your research into their own thinking about your topic and to build on what you have done and found.

Journal Articles

An even more effective way to disseminate your findings and interpretations—and certainly a more visible and permanent one—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 that it's the right place for your article. Determine whether the journal includes research reports, including reports about your general topic. 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 determine whether you should be sole author or share authorship with one or more other individuals. For example, when presenting or publishing reports based on master's theses or 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 coauthors 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 analyses, typing, or minor editing—but who have not contributed *intellectually* to the work—usually don't 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 mentioned in the 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 project or your writing and thus, we hope, 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. In some cases, however, conference program chairs and journal editors simply don't have the space for every good research project that comes their way.

Rejection letters are always disappointing, but we urge you not to let them discourage you. 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 to let your frustration and disappointment dissipate a bit—and then try to look objectively and dispassionately at what the reviewers have had to say. Occasionally a reviewer will be unforgivably 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.

You can take heart in the fact that even very experienced researchers occasionally get rejection letters. *Good* researchers constructively use negative reviews, perusing them closely

for important ways in which they might strengthen their reports and make them the very best that they can be.

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!

MyLab Education Self-Check 13.6

MyLab Education Application Exercise 13.3: Dissemination Strategies

A CLOSING THOUGHT

As you prepare to write your final research report, you might do well to read one or more books on effective writing. One classic source is Strunk and White's *The Elements of Style* (e.g., 2009); you can find many others in the "For Further Reading" sections in this chapter and in Chapter 1.

Whatever you do, remember to keep your thoughts and statements clear, precise, and concise. Look closely at your choices of nouns, verbs, and adjectives. Avoid exaggerated and unwarranted claims. Stick to the facts. Report them accurately, but in so doing, enliven your prose with variety in sentence structure and sentence length.

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 of 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.

APPLICATION EXERCISES

Similar exercises are available on Pearson MyLab Education.

1. Alan conducted a study of 124 low-achieving second-grade learners' preferences for instructional methods in mathematics and is preparing the report. In one group, math concepts were initially introduced using manipulatives. In the other group, the same concepts were initially introduced through an instructor-led presentation. Both before and after instruction, Alan assessed understanding of the concepts with a set of math problems. His research hypothesis was that those who were initially taught through the instructor-led presentation would perform better than those initially taught through the use of manipulatives. He found no differences between the groups.

   i. After any front matter such as a title page or abstract, how should Alan begin his research report? What are the most important elements of this first section?

   ii. What information should Alan include in the description of the methods of his study?

   iii. Alan found no differences between the groups. He still should provide an interpretation of the data in his research report. What might he include in this section of his report, and what conclusion should he *avoid* drawing?

2. This activity is based on a research report by DuBois, Erickson, and Jacobs (2007), which is available at no cost through the Education Resources Information Center. To download the article, visit http://eric.ed.gov/, enter the authors' names or the title of the article in the search box, and set the keyword appropriately to title or author. The resulting record will contain a direct link to the report, which you should read carefully. The full reference is as follows:
   DuBois, K., Erickson, K., & Jacobs, M. (2007). *Improving spelling of high-frequency words for transfer in written work*. Masters' thesis, Saint Xavier University & Pearson Achievement Solutions. Retrieved from http://www.eric.ed.gov/

Your task is to critique key elements of the report. Note that you may have to search a bit for the information because the thesis format used in this report does not match the checklist given in the text. Nevertheless, you can ask the same key questions about the content.

i. What type of study is reported? What is the basic design, and is it appropriate for the authors' purpose?
ii. State the problem addressed in this research project. Is it one that had to be addressed via systematic research? Is it clearly stated?
iii. What methods did the researchers use to evaluate student spelling ability? Did they use valid and reliable measurement tools? Explain your reasoning.
iv. The researchers conclude that their intervention was successful. What evidence do they present in support of this conclusion? Is it enough evidence to persuade you that they've solved their problem?

3. A graduate student has just successfully defended her dissertation, a qualitative research study examining the role of popular culture and media in children's literacy development in the early grades. She is interested in sharing her research with a broader audience and would like to publish her results.

i. What venues should she consider for publication of her work? What should she do to prepare her work for publication?
ii. In addition to publication, what other venues might this student consider for disseminating research? What steps should she take in order to disseminate her research?
iii. The student has been accepted to present her work at a national conference. How should she prepare?

FOR FURTHER READING

American Psychological Association. (2010). *Concise rules of APA style* (6th ed.). Washington, DC: Author.

American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.

American Psychological Association. (2012). *APA style guide to electronic references* (6th ed.). Washington, DC: Author.

Barzun, J. (2001). *Simple and direct: A rhetoric for writers* (4th ed.). New York, NY: HarperCollins.

Boice, R. (1993). Writing blocks and tacit knowledge. *Journal of Higher Education, 64*(1), 19–54.

Booth, W. C., Colomb, G. G., Williams, J. M., Bizup, J., & Fitzgerald, W. T. (2016). *The craft of research* (4th ed.). Chicago, IL: University of Chicago Press.

Chicago manual of style (17th ed.). (2017). Chicago, IL: 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. (2014). *Scientific style and format: The CSE manual for authors, editors, and publishers* (8th ed.). Reston, VA: Author.

Dahlberg, B., Wittink, M. N., & Gallo, J. J. (2010). Funding and publishing integrated studies: Writing effective mixed methods manuscripts and grant proposals. In A. Tashakkori & C. Teddlie (Eds.), *Mixed methods in social & behavioral research* (2nd ed., pp. 775–802). Los Angeles, CA: Sage.

Day, R. A., & Sakaduski, N. (2011). *Scientific English: A guide for scientists and other professionals* (3rd ed.). Santa Barbara, CA: Greenwood.

Evergreen, S. D. H. (2018). *Presenting data effectively: Communicating your findings for maximum impact* (2nd ed.). Thousand Oaks, CA: Sage.

Gastel, B., & Day, R. A. (2016). *How to write and publish a scientific paper* (8th ed.). Santa Barbara, CA: Greenwood.

Greene, A. E. (2013). *Writing science in plain English.* Chicago, IL: University of Chicago Press.

Hacker, D., & Sommers, N. (2015). *A writer's reference* (8th ed.). New York, NY: Bedford/St. Martin's.

Hacker, D., & Sommers, N. (2016). *Rules for writers* (8th ed.). New York, NY: Bedford/St. Martin's.

Lunenburg, F. C., & Irby, B. J. (2008). *Writing a successful thesis or dissertation: Tips and strategies for students in the social and behavioral sciences.* Los Angeles, CA: Corwin.

Matthews, J. R., & Matthews, R. W. (2014). *Successful scientific writing: A step-by-step guide for the biological and medical sciences* (4th ed.). Cambridge, England: 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.

Meloy, J. M. (2001). *Writing the qualitative dissertation: Understanding by doing* (2nd ed.). Hillsdale, NJ: Erlbaum.

Modern Language Association. (2016). *MLA handbook* (8th ed.). New York, NY: Author.

Munter, M., & Paradi, D. (2012). *Guide to PowerPoint 2010.* Hoboken, NJ: Pearson.

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.

Rocco, T. S., & Hatcher, T. (Eds.). (2011). *The handbook of scholarly writing and publishing.* San Francisco, CA: Jossey-Bass.

Seely, J. (2013). *Oxford guide to effective writing and speaking: How to communicate clearly* (3rd ed.). Oxford, England: Oxford University Press.

Turabian, K. L. (2013). *A manual for writers of research papers, theses, and dissertations: Chicago style for students & researchers* (8th ed.). Chicago, IL: University of Chicago Press.

U.S. Government Printing Office. (2016). *Style manual: An official guide to the form and style of federal government publishing.* Washington, DC: Author.

Wainer, H. (1984). How to display data badly. *American Statistician, 38,* 137–147.

Wolcott, H. F. (2009). *Writing up qualitative research* (3rd ed.). Los Angeles, CA: Sage.

Appendix A
Using a Spreadsheet: Microsoft Excel

On the surface, an electronic spreadsheet looks like nothing more than a table with many rows and columns. But the typical features of spreadsheet software enable you to do an almost limitless number of things with the data you put in the table. For example, you can recode your data, reorganize the data in various ways, and perform various 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 our discussion on the 2016 version of Excel for Macintosh.

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 either (a) the Internet address (Uniform Resource Locator, or URL) at which you retrieved the document or (b) for many documents posted since 2000, the document's Digital Object Identifier (DOI).

Let's organize such information with Excel by going to the "File" drop-down menu at the very top of the computer screen 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 might be too long for the cells, so let's do two things. First, let's go back to the "File" drop-down menu, click on "Page Setup," then on "Landscape," and finally on "OK." By doing this, we turn the page sideways and give ourselves more room across the page. Now let's move the cursor up to the row of alphabet letters labeling the columns. If we move the cursor to the line separating the A and B columns and jiggle it around just a bit, a cross-with-arrow-points icon should soon appear. 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 highlighting those cells and going to the "Format" drop-down menu, then clicking on "Cells," then "Font," then "Bold," and finally "OK." With such steps we've set up our list for keeping track of books.

Now let's 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"; then, in the first seven cells of Row 2, we can insert the headings "Authors," "Article Title," "Journal Title," "Date," "Vol(Iss)" (for "volume and issue"), "Pp." (for "page numbers"), and "DOI/URL" (for articles available only in online form). 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 to include columns labeled "Author/Organization," "Title," "DOI/URL," "Date Posted," and "Date Retrieved"; the two columns for dates may or may not be applicable, depending on how much the source is likely to change its content and/or Internet address over time. Sheet 3 should also have a column for any other pertinent information that might be important for identifying the resource.

Our workbook of three spreadsheets is now ready for us to enter information about our various library and Internet resources. We could certainly print out the sheets and add the necessary information in pen or pencil, but it would be far more efficient to take a laptop or tablet computer with us to the library and insert the information directly into the electronic file. Figure A.1 shows how the three spreadsheets might look for a few resources on the topic

BOOKS					
Authors/Editors	**Title**	**Date**	**Call Number**	**Publisher**	**Pub. Loc.**
Völlink, T., Dehue, F., &	Cyberbullying: From theory to intervent	2016	HV6773.15.C92 C	Routledge	London, England
Hitchcock, J. A.	Cyberbullying and the wild, wild web: W	2017	HV6773.15.C92 H	Rowman & L	Lanham, MD
Betts, L. R.	Cyberbullying: Approaches, consequenc	2016	HV6773.15.C92eb	Palgrave/Ma	London, England

ARTICLES						
Authors	**Article Title**	**Journal Title**	**Date**	**Vol/(Iss)**	**Pp.**	**DOI/URL**
Valkenburg, P. M., & P	Social consequences of th	Current Directions in Psychol	2009	18(1)	1–5	
Hinduja, S., & Patchin	Bullying, cyberbullying, a	Archives of Suicide Research	2010	14	206–221	
Smith, P. K., Mahdavi,	Cyberbullying: Its nature	Journal of Child Psychology a	2008	49(4)	376–385	onlinelibrary.w

WEBSITES					
Author/Organization	**Title**	**DOI/URL**	**Date Posted**	**Date Retrieved**	**Other Info.**
Megan Meier Foundation	Megan's story	meganmeierfoundation.org		May 15, 2017	
U.S. Department of Health	Cyberbullying	stopbullying.gov			
Cyberbullying Research Ce	What is cyber	cyberbullying.org			

FIGURE A.1 ■ Using Excel to Keep Track of Library and Internet Resources

of cyberbullying. Notice that many of the entries (e.g., the three 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 "DOI/URL" columns in the ARTICLES and WEBSITES sheets are in colored font rather than in black font. When you type an Internet address into a cell, Excel sometimes automatically makes it a *hyperlink*; clicking on the link will immediately take you to that address. If for some reason the address remains in black font, you can select the cell, go to the "Insert" drop-down menu, click on "Hyperlink," and then type in the address.

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, including (among other things) text, numbers, and dates. 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 four rating-scale items previously presented in Chapter 6:

		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 Chapter 6, researchers Middleton, Ormrod, and Abrams (2007) included these four *risk-taking* items in a questionnaire they created to assess a variety of traits in college education majors who were completing their teaching internship. The researchers also included four *perfectionism* items in the questionnaire, 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

FIGURE A.2 Hypothetical Data for 10 People Responding to Eight Rating-Scale Items Related to Risk Taking and Perfectionism

	A	B	C	D	E	F	G	H	I
1		RISK-TAKING				PERFECTIONISM			
2	Person	#11	#16	#39	#51	#19	#27	#38	#60
3	1	2	5	3	4	3	5	3	2
4	2	4	1	5	2	4	2	5	4
5	3	5	2	3	3	5	3	4	4
6	4	2	4	2	4	4	4	3	3
7	5	1	5	3	5	2	5	1	2
8	6	3	2	5	1	5	2	5	3
9	7	5	1	3	2	4	3	4	3
10	8	1	3	1	5	4	5	2	3
11	9	4	1	3	2	5	3	5	4
12	10	3	3	2	5	3	5	3	2

Let's create a spreadsheet to enter data for 10 hypothetical students' responses to these eight items. For a simple data set, the general convention is to assign each *row* in the spreadsheet to a particular participant and to assign each *column* to a particular variable assessed for each participant.[1] However, let's use the first two rows in our spreadsheet to insert headings for the columns. We actually begin with Row 2, in which we include the heading "Person" (to indicate the column in which each participant is identified by a number) and then eight headings for the eight questionnaire items; we increase the column widths slightly so that all nine headings can be seen in their entirety. We then move up to Row 1, where we type "RISK-TAKING" in cell B1 and "PERFECTIONISM" in cell F1 to remind us which characteristic each of the items was intended to assess; in this case, the two broader headings can also be seen in their entirety because cells to each one's immediate right are blank. Finally, we enter participants' numbers and item responses in Rows 3 through 12. The resulting spreadsheet is shown in Figure A.2.

Can we combine a person's responses to the four risk-taking items to create an overall risk-taking score and, similarly, combine responses to the four perfectionism items to create an overall perfectionism score? Not necessarily—it depends on whether the responses yield ordinal or interval data (see the section "Types of Measurement Scales" in Chapter 4 and the Practical Application "Using Checklists, Rating Scales, and Rubrics" in Chapter 6). But for the purposes of our example, let's say that the eight rating-scale items just presented *do* yield interval data. Even so, we can't just add up the numbers to get total scores. 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 *im*perfection. 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 while leaving 3s as they are. Thus, in their recoded forms, higher-number responses to these three items would indicate high rather than low risk taking or perfectionism. The following simple formula makes this conversion for us:

$$6 - \text{Original response} = \text{Recoded response}$$

[1]Some data sets may require more than one row for each participant and more than one column for each variable. For example, this might be the case if some categories are nested within other, broader categories. If you are working with such a data set, you will need more detailed instructions than those we offer here.

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

$$6 - 2 = 4$$

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:

$$= 6 - \text{B3}$$

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 *without doing anything else between the 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 cell immediately below the "#16" label; by doing this, we will now see "=C3" in the RtScore cell for Person #1.
4. Hold down the Shift key and hit the equals sign key, which will yield a plus sign (+); we will now see "=C3+" in the cell.

FIGURE A.3 Adding Three Columns for Reversing People's Responses to Certain Rating-Scale Items

	A	B	C	D	E	F	G	H	I	J	K	L
1		RISK-TAKING				PERFECTIONISM						
2	Person	#11	#16	#39	#51	#19	#27	#38	#60	Rev11	Rev39	Rev27
3	1	2	5	3	4	3	5	3	2	4	3	1
4	2	4	1	5	2	4	2	5	4	2	1	4
5	3	5	2	3	3	5	3	4	4	1	3	3
6	4	2	4	2	4	4	4	3	3	4	4	2
7	5	1	5	3	5	2	5	1	2	5	3	1
8	6	3	2	5	1	5	2	5	3	3	1	4
9	7	5	1	3	2	4	3	4	3	1	3	3
10	8	1	3	1	5	4	5	2	3	5	5	1
11	9	4	1	3	2	5	3	5	4	2	3	3
12	10	3	3	2	5	3	5	3	2	3	4	1

FIGURE A.4 ■ Adding Two Columns with Overall Scores for Risk Taking and Perfectionism

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		RISK-TAKING				PERFECTIONISM								
2	Person	#11	#16	#39	#51	#19	#27	#38	#60	Rev11	Rev39	Rev27	RtScore	PerfScore
3	1	2	5	3	4	3	5	3	2	4	3	1	16	9
4	2	4	1	5	2	4	2	5	4	2	1	4	6	17
5	3	5	2	3	3	5	3	4	4	1	3	3	9	16
6	4	2	4	2	4	4	4	3	3	4	4	2	16	12
7	5	1	5	3	5	2	5	1	2	5	3	1	18	6
8	6	3	2	5	1	5	2	5	3	3	1	4	7	17
9	7	5	1	3	2	4	3	4	3	1	3	3	7	14
10	8	1	3	1	5	4	5	2	3	5	5	1	18	10
11	9	4	1	3	2	5	3	5	4	2	3	3	8	17
12	10	3	3	2	5	3	5	3	2	3	4	1	15	9

5. Click on the first cell below the "#51" label.
6. Again type a plus sign (+).
7. Click on the first cell below the "Rev11" label.
8. Once again type a plus sign (+).
9. Click on the first cell below the "Rev39" label.

At this point, the entry for the first participant's RtScore should look like this:

= C3 + E3 + J3 + K3

We immediately hit the Enter or Return button and—*voila!*—the value 16 should appear. Now we copy the contents of this cell into the nine cells below it in the column, and we 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 cell entries 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 *also including our specific headings* for the columns in Row 2. We move the cursor to the "Person" heading, hold the mouse button down, drag the cursor to the right and down until the 14 headings in the second row and the 140 cells with data are all highlighted, and then release the mouse button so that those 154 cells remain highlighted. At this point we move the cursor to the "Data" drop-down 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 click on the empty space immediately below the heading "Column," at which point we can scroll down and click on "RtScore." And because we want to have the high risk takers appear at the top of the list, we click on "Largest to Smallest" below the heading "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 isn't required to complete our mission, but it helps us keep track of which statistics we're putting in which cells.

The procedure we follow next depends somewhat on the particular version of Excel we are using. In Excel 2016 for Macintosh, we now click on the cell representing the intersection of the "Mean" row and the "RtScore" column, then go to the "Insert" drop-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 "Formula Builder" 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"), click on "AVERAGE," and then click on the "Insert Function" button. At this point we need to tell Excel *which* numbers to use in calculating the average (mean); we can do so by highlighting the risk-taking scores ("RtScore") for all 10 participants and then clicking on the "Done" button in the "Formula Builder" box. The mean of the risk-taking scores (which is 12) 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 standard deviations for our two sets of scores, but this time we first click on the cell representing the intersection of the "SD" row and either the "RtScore" column or the "PerfScore." Once we get to the "Formula Builder" box, we select "STDEV" and then click on "Insert Function." As before, we highlight the scores for which we want standard deviations calculated, being sure that we *do not* include the *means* that are now in the spreadsheet. After clicking "Done" each time, we get standard deviations of 4.9889 and 4.0565 for the risk-taking and perfectionism scores, respectively. (If you do this yourself and instead get standard deviations of 4.7329 and 3.8484, you have mistakenly included the means in your calculations.)

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 insert the "CORREL" function in this cell. When we do so, boxes appear that ask for "Array1" and "Array2." After clicking inside the Array1 box, we highlight the 10 scores 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 into the Array2 box, click on the mouse, and highlight the 10 scores in the "PerfScore" column. What we will see in the two subboxes are the following:

M3:M12

N3:N12

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		RISK-TAKING				PERFECTIONISM								
2	Person	#11	#16	#39	#51	#19	#27	#38	#60	Rev11	Rev39	Rev27	RtScore	PerfScore
3	5	1	5	3	5	2	5	1	2	5	3	1	18	6
4	8	1	3	1	5	4	5	2	3	5	5	1	18	10
5	1	2	5	3	4	3	5	3	2	4	3	1	16	9
6	4	2	4	2	4	4	4	3	3	4	4	2	16	12
7	10	3	3	2	5	3	5	3	2	3	4	1	15	9
8	3	5	2	3	3	5	3	4	4	1	3	3	9	16
9	9	4	1	3	2	5	3	5	4	2	3	3	8	17
10	6	3	2	5	1	5	2	5	3	3	1	4	7	17
11	7	5	1	3	2	4	3	4	3	1	3	3	7	14
12	2	4	1	5	2	4	2	5	4	2	1	4	6	17
13	Mean												12	12.7
14	SD												4.99	4.06
15	Corr													-0.91

FIGURE A.5 ■ The Data Set as Reorganized, with Descriptive Statistics Calculated

We click on the "Done" button, and a correlation coefficient of -0.9114 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 isn't warranted from such a small sample size. We can limit the number of decimal places to 2 by going to the "Format" drop-down menu, then to "Cells," then to "Number," and then, under "Category," to "Number" again, where we will see a place to specify the number of decimal places we want. Our final calculations are shown in Figure A.5, along with the data as previously reorganized by the "Sort" function.

You can find many other simple statistical tests in Excel, including *t*-tests and chi-square (χ^2) tests. For fairly complex statistical 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.

CREATING A DATA SET

When you open SPSS, 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 (1996) conducted in preparation for her dissertation study (excerpts from her dissertation appear in Chapter 1 and Chapter 13). The data include the following information for 15 students in a college psychology class: (a) their gender, (b) their scores on three exams administered 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 proportions, such as .406 or .496) were theorized to facilitate better learning—and thus to be of better quality—than relatively nonintegrated 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:* In this column, we identify the label that will appear for each variable in the "Data View" table. Labels can include alphabet letters, numbers, and a few other meaningful symbols (e.g., "$").
- *Type:* In this column we identify the type of data each variable represents, perhaps a number (numerical data), a letter string, a dollar amount, a date, or something else altogether.
- *Decimals:* In this column we specify an upper limit on the number of digits that will appear to the right of a decimal point.
- *Label:* In this column we identify the labels that the variables will have when we create a table or graph—perhaps one to be included in a dissertation or research report.

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								

Data View | Variable View

	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											

Data View | Variable View

FIGURE B.2 ■ The "Variables" Table

- *Values:* In this column we can identify any labels we want to give 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." After we have finished doing so, we click on "Add" and then on "OK" to say that we have labeled all possible values of the Gender variable.
- *Measure:* In this column we specify whether each variable reflects a nominal scale or an ordinal scale; the category "ordinal scale" also encompasses interval and ratio scales. (You can find descriptions of the four kinds of scales in either Chapter 4 or Chapter 11.) 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."

FIGURE B.3 ■ The "Value Labels" Box for the "Variable View" Table

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 yield meaningless numbers). We move the cursor to the word "Analyze" at the top of the screen and click on the mouse. A drop-down menu appears, and we move the cursor down and click on "Descriptive Statistics," 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 (i.e., the sum of a student's three exam scores), quantity of notes, and 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 the "Analyze" drop-down menu at the top of the screen, choose "Correlate," and then move the cursor to the right to choose "Bivariate." 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 (see Figure B.5), which we can print out by going to the "File" drop-down menu and then to "Print." 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.

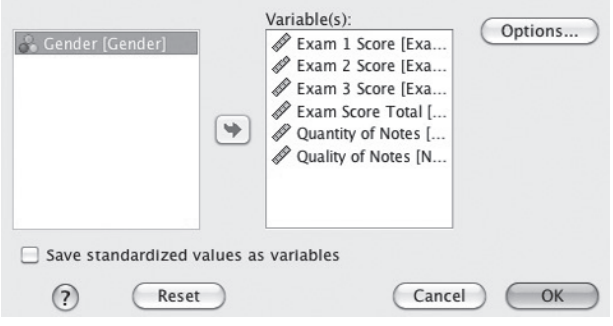


FIGURE B.4 ■ Identifying Variables for Which We Want Basic Descriptive Statistics to Be Calculated

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).

The middle number in any two-variable cell 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. (For a review of p values, see the section "Testing Hypotheses" in Chapter 11.)

The table in Figure B.5 marks with two asterisks (**) all rs 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 use a significance level (α) of .05 for all of our analyses. Any p value in the table that is *smaller* than .05 indicates that the variables probably *are* correlated in the population from which our sample has been drawn. For instance, 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 .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 the correlation necessarily indicates a causal relationship: Taking more notes doesn't necessarily cause a student to take better ones, nor does taking better ones cause a student to take more of them. Remember, correlational data alone *never* allow 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 choose "Compare Means" and then "Independent Samples T Test."[1] A box similar to

[1]As 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.

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

that shown in Figure B.4 appears, but this one has three smaller 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 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 (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 (ANOVA). 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 the "Analyze" drop-down menu, choose "Compare Means," and then choose "Paired-Samples T Tests." 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 (see Figure B.7). The first table gives us descriptive statistics; we've seen most of these before, but the column for the standard error of the mean is new (this statistic is described in Chapter 11). 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 alpha 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.

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)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	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

Glossary

***a priori* hypothesis** Hypothesis that is proposed before data collection begins and that to some degree guides research design and data analysis.

abduction In scientific inquiry, a theory-building process in which a researcher begins with the facts known about a phenomenon and then brainstorms ideas about a plausible best explanation for the phenomenon.

abstract In a research report, brief summary of the purpose, methods, and findings of a research project, along with general conclusions drawn from the project.

academic integrity Honesty and responsible behavior in scholarly activities; includes ensuring the privacy and well-being of research participants, accurately portraying findings and weaknesses of one's research, and appropriately crediting the works of others.

acknowledgments In a research report, a section that recognizes other individuals who have provided noteworthy assistance to the research endeavor.

action research Research focusing on an existing problem (usually in a helping profession), with the goal of improving current practices and desired outcomes.

advance organizer Introduction to a discussion that provides an overall organizational scheme for the discussion.

alpha (α) In statistical analysis, a predetermined probability level at which a researcher can reject a null hypothesis; also known as a *significance level.*

applied research Research that has immediate relevance to current procedures or policies and can inform decision making about practical problems.

assessment Identification and evaluation of a key characteristic or other variable in a research study; sometimes involves assigning numbers to (i.e., measuring) a variable.

assumption In a research study, a premise that is taken for granted without confirmatory evidence.

audit trail Written record of research activities and procedures as a means of documenting progress and enhancing the credibility of a research study; most often used in qualitative research.

autoethnography Type of ethnography in which one or more researchers conduct ethnographic research regarding their *own* perspectives, beliefs, experiences, and practices.

baseline data Measures of a dependent variable taken before any experimental treatment or intervention has been administered.

basic research Research intended to advance theoretical conceptualizations about a particular topic.

bias In a research study, any influential factor that can distort the data obtained or conclusions drawn.

case study Qualitative research design in which a single individual, program, or event is studied in depth for a defined period of time.

checklist List of characteristics that may or may not be evident in an individual performance or product; can be used to quantify the overall quality of the performance or product.

cohort-sequential study Descriptive quantitative study in which participants in (a) two or more age groups are followed over a lengthy period, (b) data are collected at two or more times and (c) developmental trends for the various age groups are compared.

collective coding Type of qualitative coding process in which members of a research team individually code pieces of data and then come together to generate and agree on a list of common codes that can be used to code other data; used primarily in action-oriented and participatory research models.

community-based participatory research (CBPR) Approach to research that employs a collaborative partnership in the study of community health issues, with the assumption that such a partnership will produce more useful findings and lead to improvements in community health outcomes.

community-based qualitative research (CBQR) Model of research that builds on the traditions of participatory action research (PAR), using qualitative methodologies and collaborating with community members and organizations to investigate and address local issues and problems.

community-based research (CBR) Group of research models in which research investigations are conducted in particular community settings and in partnership with certain community organizations/groups as a means of addressing salient issues and problems within the community.

conceptual framework Network of interconnected concepts that can be useful in helping to understand the underlying nature of a phenomenon under investigation.

confidence interval Statistically computed range within which a population parameter probably lies; also known as an *interval estimate.*

confirmability In qualitative research, a researcher's concerted effort to base conclusions on *actual data* as much as possible and to describe data-collection and data-analysis processes in enough detail that other researchers might reasonably replicate their study and draw similar conclusions.

confirmation bias Tendency to seek information that confirms rather than discredits one's current beliefs.

confounding variable Unexamined variable that is or might be correlated with both an independent variable and a dependent variable; must be controlled if conclusions about cause-and-effect relationships are desired.

construct validity Extent to which an assessment strategy yields credible results regarding a characteristic that cannot be directly observed but is assumed to exist based on patterns in people's behaviors or creations (such a characteristic is a *construct*).

constructivism Philosophical perspective based on the idea that any understandings of physical or psychological phenomena are inevitably only human creations and beliefs; that is, researchers *construct* understandings that can never reflect "true" reality (if such a reality exists).

content analysis Qualitative research design in which a body of material is systematically examined in order to detect general patterns, themes, or biases; the material typically involves some form of human communication.

content validity Extent to which an assessment instrument or procedure adequately reflects the full breadth of the characteristic being assessed.

continuous variable Variable that has an infinite number of possible values falling along a continuum.

control group Group of people in a research study who are given either no experimental treatment or a *placebo* treatment that is unlikely to have an effect on the dependent variable(s) of interest.

convenience sampling Approach to sampling that entails using readily available research sites and/or easily accessible individuals who can provide insights related to central research questions.

convergent design Mixed-methods design in which a researcher collects both quantitative and qualitative data in parallel, usually at the same time and with respect to the same general research problem; similar weight is given to the two types of data, with the hope that they will yield consistent or complementary findings.

copyright Legal right of the creator of an intellectual or artistic work to have exclusive use of that work unless the creator explicitly grants permission for someone else to use it.

correlation Extent to which two variables are associated, such that when one variable increases, the other either increases or decreases somewhat predictably.

correlational study Descriptive quantitative study that explores possible relationships among two or more variables.

correlation coefficient Statistic that indicates the strength and direction of an association between two variables.

course-based action research (CBAR) Type of research that integrates community-based research and participatory action research; brings together college students, faculty, and local community partners to undertake course-based collaborative research projects that yield benefits for the community partners.

credibility As a characteristic of a research study, the study's overall quality as judged by other scholars; includes the use of appropriate designs and methods, believable findings, and plausible interpretations.

criteria for admissibility of the data Criteria used to determine whether individual pieces of data warrant inclusion in a researcher's overall data set.

criterion validity Extent to which the results of an assessment strategy correlate with assessments of another, presumably related characteristic; may also be called *predictive validity* if the related characteristic is assessed at a subsequent time.

critical thinking Process of evaluating the accuracy, credibility, and worth of information and lines of reasoning.

cross-sectional study Descriptive quantitative study in which participants from two or more age groups are sampled and the groups' characteristics or behaviors compared.

cultural domain Unit of cultural knowledge regarding a particular topic that ethnographers use to organize and help make sense of a larger, more all-encompassing system of cultural knowledge.

cultural knowledge Concepts, beliefs, attitudes, values, problem-solving skills, and other mental phenomena that members of a particular culture or cultural group have internalized through their ongoing interactions with other members.

deductive logic Reasoning process in which a researcher begins with one or more premises ("givens") and then identifies conclusions that can undisputedly be drawn from the premise(s).

dependability In qualitative research, a concept that accounts for the ever-changing contexts within which research studies take place and thus requires researchers to provide in-depth descriptions of their data collection methods.

dependent variable Variable in a quantitative research study that is hypothesized to be influenced by an independent variable; hence, its value *depends* on the value of the independent variable.

descriptive research Research that enables researchers to draw conclusions about the current state of affairs regarding a situation or issue but not about cause-and-effect relationships.

descriptive statistics Statistics that summarize the nature of a particular set of numerical data (e.g., means, standard deviations, correlation coefficients).

design-based research (DBR) Model of research that aims to generate theories and knowledge about learning and teaching that are grounded in real-life contexts; typically involves conducting and refining experiments and interventions so that researchers can make evidence-based claims about the nature of learning and/or the effectiveness of certain strategies for enhancing it.

Digital Object Identifier (DOI) Unique, permanent number assigned to a document posted on the Internet; enables one to find the document even if its exact location (URL) changes over time.

discrete variable Variable that has a finite and usually small number of possible values.

discriminant sampling In a grounded theory study, a follow-up to theoretical sampling in which a researcher identifies one or more data sources that might help substantiate an emerging theory of the phenomenon under investigation.

DOI *See* Digital Object Identifier (DOI).

domain analysis Data analysis process that entails identifying and describing key cultural domains of knowledge, beliefs, and practices; used primarily in ethnographic research studies.

double-blind experiment Experiment in which neither participants nor research associates (e.g., those who administer interventions or analyze qualitative data) are aware of any participant's membership in a particular treatment or control group.

effect size (ES) Statistically determined estimate of the strength of a relationship between two variables (e.g., the degree to which an intervention has an effect on a dependent variable, or the strength of a correlation between two variables).

electronic discussion group Internet-based mailing list to which people can subscribe in order to get ongoing messages, newsletters, and other postings related to a particular topic.

electronic spreadsheet Software program that enables a researcher to enter and then manipulate data in a two-dimensional table.

emerging design In qualitative research, a strategy in which data collected early in the investigation influence the methods that are subsequently used and the kinds of data that are subsequently collected.

end matter In a research report, content that follows the final section or chapter (e.g., reference list, appendices).

endnote Brief note at the end of a written document that communicates further information about a particular statement within the document.

equivalent-forms reliability Extent to which two different versions of the same assessment instrument (e.g., "Form A" and "Form B") yield similar results.

ethnography Qualitative research design in which a researcher looks in depth at a cultural group in its natural setting.

experience-sampling method (ESM) Approach to data collection in which a researcher collects frequent and ongoing data about people as they live their normal, everyday lives.

experimental design Quantitative research design that enables conclusions about cause-and-effect relationships; involves both (a) manipulation of one or more variables to determine their possible effects on one or more other, dependent variables; and (b) control of additional variables that might otherwise have an impact on the dependent variables.

experimental group Group of people in a research study who are given a particular experimental intervention in order to observe the intervention's possible effect on one or more dependent variables; also known as a *treatment group.*

explanatory sequential design Two-phase mixed-methods design in which quantitative data collection is followed by the collection of qualitative information that can help clarify the meanings of the quantitative findings.

exploratory sequential design Two-phase mixed-methods design in which qualitative data are collected in an effort to inform the planning and implementation of subsequent quantitative data collection.

ex post facto design Quantitative research design in which a researcher identifies existing conditions or previously occurring events in people's lives and then collects data to determine whether a relationship exists between these factors and subsequent characteristics or behaviors.

external validity Extent to which the results of a research study are applicable to other contexts, especially in real-world situations; term more commonly used in quantitative research than in qualitative research.

extreme case sampling Type of purposive sampling that involves focusing on particular cases or individuals determined to be unusual and atypical, or to represent extreme examples; most often used in qualitative (rather than quantitative) research.

face validity Extent to which, on the surface, an assessment strategy *appears to yield* an accurate assessment of a particular characteristic.

factorial design Quantitative research design that examines the possible effects of two or more independent variables and possible ways they might interact in their influences.

figure Graphic illustration of certain concepts, phenomena, or statistical findings.

five-number summary Method of describing variability in which five numbers are reported: the lowest and highest numbers in the distribution, Quartile 1 (25th percentile), the median, and Quartile 3 (75th percentile).

focus group Small group of people who are assembled and asked to express their perspectives about a particular issue.

footnote Brief note at the bottom of a page that communicates further information about a particular statement on the page.

front matter In a research report, content that precedes the introductory text (e.g., title, abstract, table of contents).

gatekeeper In an ethnographic study, a person who can smooth the way for a researcher's entrance into a particular cultural group.

generalizability As a characteristic of a research study, the extent to which the study's findings can be applied to other people, situations, or contexts.

geometric mean Measure of central tendency appropriate for a data set reflecting growth or deceleration; calculated by multiplying all data points and then finding the *N*th root of the product.

grounded theory study Qualitative research design in which a researcher collects data relevant to a research problem and uses them to develop a new theory about a particular phenomenon.

Hawthorne effect Phenomenon in which participants in a research project change their behavior simply because they know they are being exposed to an experimental intervention.

independent variable Variable in a quantitative research study that is hypothesized to have an influence on one or more other variables; in many cases, is directly manipulated by the researcher.

inductive reasoning Reasoning process in which a researcher begins with multiple specific observations about a particular phenomenon and then makes broader generalizations about the phenomenon.

inferential statistics Statistics that enable inference-drawing about large populations from sample data.

informed consent Practice of both (a) informing participants (or, if applicable, their legal guardians) about the nature of a research study and (b) obtaining written permission for participation.

informed consent form Form that describes the nature of a research project and of one's participation in it; required for many studies involving human beings.

institutional animal care and use committee (IACUC) Committee at a research institution that scrutinizes proposals for studies involving nonhuman animals; its primary role is to determine and minimize the extent to which the studies might cause unnecessary suffering, distress, or death.

institutional review board (IRB) Committee at a research institution that scrutinizes proposals for studies involving human participants; its primary role is to ensure that the studies will not expose participants to undue harm or in some other way violate basic ethical standards.

instrumentation bias Characteristic of a measurement instrument or procedure that consistently and differentially slants the obtained results in a particular direction.

intangible phenomenon Phenomenon that has no clear-cut basis in the physical world (e.g., an opinion, feeling, or underlying personality characteristic).

interlibrary loan Multilibrary alliance through which one can gain access to books and additional sources at libraries other than one's own.

internal consistency reliability Extent to which all of the items or other tasks within a single assessment instrument yield similar results.

internal validity Extent to which a research design and its implementation enable defensible conclusions about cause-and-effect and other relationships among variables.

interquartile range Measure of variability reflecting the difference between Quartile 1 (25th percentile) and Quartile 3 (75th percentile).

interrater reliability Extent to which two or more individuals evaluating the same product or performance make identical or similar judgments; also known as *interrater agreement.*

interval estimate Statistically computed range within which a population parameter probably lies; also known as a *confidence interval.*

interval scale Numerical scale that reflects equal units of measurement, but with an arbitrary zero point (i.e., the number 0 does *not* indicate total absence of the characteristic being measured).

interview guide List of guiding, flexible questions to be used when conducting relatively unstructured individual interviews and focus-group sessions.

IQ score Standard score with a mean of 100 and a standard deviation of 15.

key informant In an ethnographic study or similar qualitative design, a person who can provide especially useful information and insights and possibly also facilitate contacts with other helpful individuals; also known as a *primary informant.*

keywords Words or short phrases used to locate relevant sources in a database or on the Internet.

kurtosis Parameter of a symmetrical distribution that reflects the degree to which the distribution is unusually pointy or flat, such that the percentages within each portion of the distribution are notably different from those in the normal distribution.

leptokurtic distribution Distribution in which a greater percentage of data points are clustered near the mean than is true for a normal distribution; in its graphic representation, it appears more "pointy" at the middle than a normal distribution does.

Likert scale Numerical continuum on which a particular characteristic might be judged and quantified; also known as a *rating scale.*

longitudinal mixed-methods design Research design in which both quantitative and qualitative data are collected for a single group of participants on two or more occasions over a lengthy time period.

longitudinal study Descriptive quantitative study in which participants are followed over a lengthy period, with data being collected at two or more times.

matched pairs Pairs of people who are quite similar with respect to certain characteristics that might possibly have an effect on a dependent variable (e.g., gender, age, grade point average); used as a strategy for controlling for possible confounding variables in a study.

mean Measure of central tendency based on the arithmetic (i.e., mathematical) average of all data points related to a particular variable; in everyday language, known simply as an *average.*

measurement Systematic method of assigning numerical values or categories to observations so that they can be analyzed and interpreted with some degree of objectivity.

median Measure of central tendency based on the exact midpoint of a distribution of data points related to a particular variable.

mediating variable Variable in a quantitative research study that is influenced by an independent variable and then, in turn, influences a dependent variable; can help explain why the independent variable has a certain effect on the dependent variable.

member checking In qualitative research, a process in which a researcher (a) asks participants to review interview transcripts in order to double-check their accuracy and/or (b) seeks participants' feedback and possible validation regarding findings; provides a means of incorporating additional checks and balances into data-collection and data-analysis processes.

memos In qualitative research, a researcher's notes-to-self about preliminary interpretations of data being collected.

meta-analysis Statistical procedure in which statistical analyses from many previous quantitative studies become the "data" to be analyzed as a whole; useful in determining the degree to which previous studies about a topic have yielded consistent results.

mixed-methods research Research that includes elements of both quantitative and qualitative research.

mode Measure of central tendency based on the most frequently occurring number in a data set related to a particular variable.

moderating variable Contextual variable in a quantitative research study that influences the nature and strength of a cause-and-effect relationship between two other variables.

multiphase iterative design Mixed-methods design that includes three or more phases; earlier phases inform conceptualization and implementation of subsequent phases.

multitrait–multimethod approach Strategy in which two or more different characteristics are each assessed using two or more different techniques; useful in determining the validity of each technique.

narrative inquiry In qualitative research, a research strategy in which a researcher gathers and analyzes stories and narratives from individuals who have experiences and other prior knowledge related to the events or phenomena under investigation.

negative correlation Relationship between two variables in which one variable tends to decrease as the other variable increases, reflecting an inverse relationship.

negatively skewed distribution Distribution in which data points lower than the median (midpoint) have greater variability than do data points higher than the median.

nominal scale Use of numbers or verbal labels to assign each piece of data to one of two or more categories.

nonparametric statistics Statistical procedures that are not based on any assumptions about the nature of the population from which one or more samples have been drawn; available only for relatively simple statistical questions (e.g., questions related to central tendency or correlation).

non-peer-reviewed research report Research report that has not been screened for quality or accuracy by one or more experts; also known as a *nonjuried* or *nonrefereed* report.

nonprobability sampling Sample selection process in which some members of a population have a greater chance of being selected than others.

normal distribution (normal curve) Theoretical pattern that many research variables approximate, in which most data points are in the middle range and only a few lie at the extremes; also known as a *bell curve.*

norm-referenced score Score that indicates where each individual is positioned in a distribution relative to other members of the person's norm group.

novelty effect Phenomenon in a research study in which participants' behavior changes not as the result of a specific experimental intervention, but rather as a result of the fact that the environment has changed in a noticeable way.

null hypothesis Hypothesis postulating that a statistically significant result is due entirely to chance.

observation study Descriptive quantitative study in which observations focus on specific, predetermined behaviors and are systematically quantified in some manner.

online database In academic research, an Internet-based resource containing general information about thousands of scholarly works; is easily searchable with the use of appropriate keywords or other relevant information (e.g., author, journal title, publication date).

open-access journal Scholarly journal that anyone can read on the Internet and download without cost.

open coding In qualitative data analysis, an inductive approach that entails reviewing collected data line by line to identify codes, rather than prescribing coding categories in advance.

operational definition Definition of a variable in a research study in terms of how it will be identified or measured (e.g., an operational definition of *intelligence* might be participants' performance on an IQ test).

ordinal scale Numerical scale in which assigned numbers reflect only the rank-ordering of various pieces of data with respect to a particular variable; does not legitimately allow addition or subtraction of two or more data points.

orienting experiences Activities conducted as a part of a community-based research project in an effort to help academic researchers learn about the community and its current issues and concerns (e.g., background readings, community tours).

parameter Numerical characteristic of a specific, entire population that summarizes a certain aspect of its distribution (e.g., its central tendency or variability).

parametric statistics Statistical procedures based on certain assumptions about the nature of the population from which one or more samples have been drawn; two common assumptions are (a) that the data reflect an interval or ratio scale and (b) that the data approximate a normal distribution.

participant observation In some ethnographic studies, an approach in which researchers immerse themselves in the daily lives of the people they are studying.

participatory action research Action-oriented approach to research that involves reciprocal collaboration with individuals who, in other circumstances, would be viewed only as the "subjects" of the research; its primary goal is to actively address a particular problem or issue affecting the collaborating individuals.

pdf *See* portable document format (pdf).

peer-reviewed research report Research report that has been judged by respected colleagues in one's field and deemed to be of sufficient quality and importance to warrant presentation or publication; also known as a *juried* or *refereed* report.

percentile rank (percentile) Norm-referenced score indicating the percentage of peers in a norm group getting a raw score less than or equal to a particular person's raw score.

periodical Publication for which new issues come out on a regular basis (e.g., a newspaper, magazine, or professional journal).

phenomenological study Qualitative research design in which a researcher tries to understand people's perceptions and perspectives relative to a particular situation.

phenomenology Philosophical approach to research in which the focus of investigation is on how people perceive and experience themselves and certain aspects of their world.

pilot study Brief exploratory investigation to determine the feasibility and validity of procedures, measurement instruments, or methods of analysis that might be useful in a subsequent, more in-depth research study.

placebo In an experimental study, a treatment that is presumed to have little or no effect on a dependent variable of interest.

plagiarism Misrepresentation of another person's work as being one's own.

platykurtic distribution Distribution in which a greater percentage of data points are located near the extremes than is true for a normal distribution; in its graphic representation, it appears flatter at the middle than a normal distribution does.

point estimate Single statistic that is used as a reasonable estimate of the corresponding population parameter.

portable document format (pdf) Electronic document that captures both the content and format of a printed document or static computer file.

positionality In qualitative research, a concept that encompasses a researcher's background characteristics and experiences, past and current social roles, and personal biases as they all relate to the setting, participants, and/or research topic.

positive correlation Relationship between two variables in which one variable tends to increase as the other variable increases.

positively skewed distribution Distribution in which data points higher than the median (midpoint) have greater variability than do data points lower than the median.

positivism Philosophical perspective based on the idea that with appropriate techniques, scientists can objectively uncover absolute facts about cause-and-effect relationships in the physical world and in human experience.

postpositivism Philosophical perspective based on the ideas that (a) progress toward an accurate understanding of a phenomenon is likely to be gradual and probabilistic and (b) a truly complete understanding of the phenomenon may ultimately be impossible.

power Degree to which a statistical test maximizes the likelihood of correctly rejecting an incorrect null hypothesis and thus minimizes the chances of making a Type II error.

practical significance Degree to which findings related to a research intervention are useful in real-world procedures and practices; often involves a cost–benefit analysis.

pragmatism Philosophical perspective based on the idea that absolute "truths" about certain phenomena *and* people's constructed beliefs about those phenomena are both legitimate objects of study; also known as *realism.*

pre-experimental design Primitive quantitative research design in which either (a) the independent "variable" has only one value (and hence doesn't vary) or (b) members of two or more treatment and/or control groups have not been selected randomly; useful only for forming tentative hypotheses that require further testing.

primary data Data that directly emerge or emanate from an unobservable phenomenon.

primary informant In an ethnographic study or similar qualitative design, a person who can provide especially useful information and insights and possibly also facilitate contacts with other helpful individuals; also known as a *key informant.*

probability sampling Sample selection process in which each member of a population has an equal chance of being chosen.

purpose statement One or more declarative sentences that clearly describe the primary goal(s) a researcher intends to achieve (in a research proposal) or intended to achieve (in a final report) by conducting a research study.

purposive sampling Sampling selection process in which participants or other units of study are chosen on the basis of how much and what types of information they can yield about the topic under investigation.

qualitative research Research yielding information that cannot be easily reduced to numbers; typically involves an in-depth examination of a complex phenomenon.

quantitative research Research yielding information that is inherently numerical in nature or can be easily reduced to numbers.

quasi-experimental design Quantitative research design in which an independent variable is manipulated in order to determine its possible effect on another (dependent) variable, but without total control of additional variables that might have an impact on the dependent variable.

range Measure of variability calculated by subtracting the lowest-value data point from the highest-value data point.

rating scale Numerical continuum on which a particular characteristic might be judged and quantified; also known as a *Likert scale.*

ratio scale Numerical scale that reflects equal units of measurement and a true zero point that indicates total absence of the characteristic being measured.

raw score Number of questions correctly answered or number of points accumulated that a research participant earns on a measurement instrument.

reactivity General phenomenon in which people's behavior changes once they know they are being observed.

realism Philosophical perspective based on the idea that absolute "truths" about certain phenomena *and* people's constructed beliefs about those phenomena are both legitimate objects of study; also known as *pragmatism.*

reflexivity In qualitative research, a researcher's conscious attempt to (a) identify personal, social, political, or philosophical biases that might influence data collection and interpretation and then (b) take steps to minimize such influences.

reliability Extent to which an assessment strategy yields consistent information about the characteristic(s) being assessed.

repeated-measures variable Variable that is assessed two or more times for every participant in a research study; when it is an independent variable in an experimental or quasi-experimental design, it is also known as a *within-subjects variable.*

research Process of systematically collecting, analyzing, and interpreting data in order to enhance understanding of a particular phenomenon.

research design General structure that guides data collection and analysis in order to address a research problem.

research hypothesis Proposed explanation for a phenomenon under investigation; may be put forth either before data collection has begun (in which case it is also known as an *a priori* hypothesis) or after some or all data have been collected and analyzed.

research methodology General approach a researcher takes in carrying out a research project (e.g., a particular quantitative or qualitative research design).

research problem General question or problem that drives a research project, with the ultimate goal being to get an answer or solution.

research tool Specific mechanism or strategy a researcher uses to collect, manipulate, or interpret data.

respondent validation Qualitative research strategy in which a researcher seeks validation from study participants about conclusions and interpretations drawn from collected data.

response bias Tendency for people to make inaccurate statements (often inadvertently) in their responses to questions in interviews or on questionnaires.

response rate Percentage of people agreeing to participate in a survey.

restorying In narrative inquiry research, an approach to data analysis wherein researchers first (a) analyze participants' recollections and other data sources in a search for key themes and then (b) reorder the data to craft rich narratives that provide a better understanding of the setting, context, time period, and/or sequence of events.

return rate Percentage of people completing and submitting questionnaires that they have received via mail or e-mail.

robustness In statistics, degree to which a statistical procedure yields generally valid results even when the data violate one or more basic assumptions on which the procedure rests.

rubric Two-dimensional table that includes (a) two or more characteristics on one dimension and (b) concrete criteria for rating them on the other dimension; useful in numerically measuring characteristics of a multifaceted performance or product.

sample Subset of a population of people, another biological species, or inanimate objects; data collected from this subset used to draw conclusions about the population from which it has come.

sampling bias Any factor that causes a researcher to inadvertently obtain a nonrepresentative sample of the population being studied.

scales Categories of measurement that ultimately dictate the statistical procedures (if any) that can be used in processing numerical data. *See also* interval scale, nominal scale, ordinal scale, ratio scale.

scatter plot Two-dimensional graphic representation of the relationship between two variables, with each dot on the graph representing a particular participant; also known as a *scattergram.*

scientific method General approach to learning about a phenomenon in which a researcher (a) identifies a currently unanswered question, (b) systematically plans and implements data collection relevant to the question, and then (c) analyzes and interprets the data in an effort to answer the question.

search engine Program that enables one to identify Internet websites relevant to a particular topic (e.g., Google, Bing, or Yahoo!).

secondary data Data obtained from one or more people's descriptions or interpretations of primary data related to an unobservable phenomenon.

semistructured interview Interview in which a researcher asks predetermined questions but also asks individually tailored follow-up questions in order to gain clarity regarding certain responses.

significance level Probability that a statistically significant result might be due to chance alone; when determined in advance as the criterion for rejecting a null hypothesis, it is also known as *alpha* (α).

single-case intervention research Quasi-experimental research study in which a single individual or single intact group is exposed to one or more specific interventions; reversal time-series, alternating-treatments, and multiple-baseline approaches might be used either individually or in combination to provide evidence of cause-and-effect relationships.

skewed distribution Nonsymmetrical distribution of a research variable in which data points on one side extend much further from the median (midpoint) than do data points on the other side.

snowball sampling In qualitative research, a method of identifying participants that involves asking existing participants to refer or recruit other individuals (e.g., colleagues or acquaintances) who also have experiences and knowledge relevant to the topic under study.

social desirability effect Tendency of some participants in a research study to intentionally or unintentionally misrepresent their personal characteristics and behaviors in order to make a favorable impression.

sociogram Graphic depiction of people's interpersonal links within a group.

sociometric matrix Two-dimensional table that represents how members of a group view one another with respect to a particular characteristic (e.g., likeability or trustworthiness).

standard deviation Measure of variability in which (1) each data point is squared, (2) the sum of the squares is calculated, (3) this sum is divided by the total number of data points, and (4) the square root of the obtained result is calculated; is a frequently used value in parametric statistics.

standard error of the mean Standard deviation for a distribution of sample means; indicates how much a sample mean is likely to vary from one sample to another when all samples are the same size and randomly drawn from the same population.

standardization Extent to which an assessment instrument or procedure involves similar content, format, administration procedures, and scoring criteria for all participants in a research study.

standard score Norm-referenced score that indicates how far a person's performance is from the mean of a distribution in terms of standard deviation units.

stanine Standard score with a mean of 5 and a standard deviation of 2; always reported as a whole number between 1 and 9.

statistic Number computed for a sample that has been drawn from a larger population; intended to provide an estimate of a parameter of the overall population.

statistical regression Phenomenon in which research participants who are identified as being at extreme ends of a distribution at one time tend to be closer to the mean at a subsequent time; is often the result of imprecise measurement of a characteristic at either or both times.

statistical significance Situation in which a statistical result meets or surpasses a predetermined significance level and hence the null hypothesis can be rejected.

statistics Academic discipline comprised of computational procedures for finding patterns in numerical data; alternatively, two or more numbers intended to provide estimates of population parameters based on data collected from a sample of that population.

structured interview Interview in which a researcher predetermines all of the questions to be asked.

subproblem In research, a component of a research problem that is researchable and potentially solvable in its own right.

substantial phenomenon Phenomenon that can be directly observed in the physical world.

survey research Descriptive study in which a large number of people are asked questions and their responses tabulated (and, if necessary, coded) in an effort to identify general patterns or trends in a certain population.

systematic review Research study in which results of many previous qualitative and/or mixed-methods research reports become the objects of investigation; potentially relevant reports are systematically analyzed for their quality, and contents of high-quality reports are coded in an effort to identify key ideas and themes.

table Two-dimensional matrix in which words and/or numbers are displayed for the purpose of summarizing certain ideas or showing certain patterns or relationships.

table of contents List of major sections and subsections in a written document.

table of random numbers Large two-dimensional list of numbers that can guide random selection of population members to be included in a research sample.

table of specifications Two-dimensional grid that lists both the specific topics and the specific behaviors that reflect achievement in a particular content domain, with individual cells indicating the relative importance of each topic–behavior combination.

test–retest reliability Extent to which a single assessment instrument yields the same results for the same people over a short time interval.

theoretical framework Theory-based perspective that provides a foundation for a research study; often leads the researcher to form one or more *a priori* hypotheses and to analyze and interpret data within the context of that perspective.

theoretical sampling In a grounded theory study, a strategy of choosing data sources that are most likely to help a researcher develop a theory of the phenomenon under investigation.

theory Integrated set of concepts and principles developed to explain a particular phenomenon.

thick description Qualitative research strategy in which an observed situation is described in enough detail that readers can construct some of their own interpretations.

three-interview series In phenomenological research, a method of interviewing that entails conducting three interviews with each participant; the first focuses on the participant's life history, the second focuses on the details of the participants' lived experiences, and the third encourages the participant to reflect on possible meanings of his or her experiences.

transferability Extent to which a research study's findings might be similar or applicable to other individuals, settings, and contexts; term most frequently used by qualitative researchers.

treatment In an experimental study, an experimenter-controlled intervention hypothesized to have an effect on one or more dependent variables.

treatment group Group of people in a research study who are given a particular experimental intervention in order to observe the intervention's possible effect on one or more dependent variables; also known as an *experimental group.*

triangulation Collection and comparison of multiple kinds of data, with the goal of finding consistencies or inconsistencies among them.

Type I error Error in statistical decision making in which a null hypothesis that is true is inappropriately rejected; also called an *alpha error.*

Type II error Error in statistical decision making in which an incorrect null hypothesis is inappropriately retained; also called a *beta error.*

Uniform Resource Locator (URL) Specific Internet address at which a particular resource or document can be found (e.g., "google.com" or "www.census.gov").

unit of analysis Specific entity being studied in a research project (e.g., individuals, organizations, cultural artifacts).

unobtrusive measure Means of assessing research participants' behaviors in such a way that the participants are unaware that they are being observed.

URL *See* Uniform Resource Locator (URL).

validity (in assessment) Extent to which an assessment strategy yields accurate information about the characteristic it is intended to assess and thereby enables justifiable inferences about that characteristic.

variable In a research project, any characteristic, experience, behavior, or outcome of interest that has two or more distinct values.

variance Measure of variability equal to the square of the standard deviation; is a frequently used value in parametric statistics.

within-subjects design Quantitative research design in which all participants are exposed to all experimental treatments and any control conditions; enables control of confounding individual-difference variables; also known as a *repeated-measures design.*

within-subjects variable Independent variable that is systematically varied for every participant in a quantitative research study, especially in an experimental or quasi-experimental design; sometimes known as a *repeated-measures variable*.

youth participatory action research (YPAR) Variation of action research that entails working with young people as co-researchers to investigate and take action on a particular problem or issue.

***z*-score** Standard score with a mean of 0 and a standard deviation of 1.

References

Abraham, E. P., Chain, E., Fletcher, C. M., Gardner, A. D., Heatley, N. G., Jennings, M. A., . . . Florey, H. W. (1941). Further observations on penicillin. *Lancet, 2,* 177–188.

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). Los Angeles, 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.

Ballenger, C. (2009). *Puzzling moments, teachable moments: Practicing teacher research in urban classrooms*. New York, NY: Teachers College Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York, NY: Freeman.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences, 13*(1), 1–14.

Bartholomew, D. J. (2004). *Measuring intelligence: Facts and fallacies.* Cambridge, England: Cambridge University Press.

Bazeley, P. (2010). Computer-assisted integration of mixed methods data sources and analyses. In A. Tashakkori & C. Teddlie (Eds.), *Mixed methods in social & behavioral research* (2nd ed., pp. 431–467). Los Angeles, CA: Sage.

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, A. B., Christoffel, K. K., Morales, M. A., Lopez, J. E., & Longjohn, M. (2011). Combatting childhood obesity through a neighborhood coalition: Community organizing for obesity prevention in Humboldt Park. In S. Whitman, A. M. Shah, & M. R. Benjamins (Eds.), *Urban health: Combating disparities with local data*. New York, NY: Oxford University Press.

Becker, H. S. (1970). Whose side are we on? In W. J. Filstead (Ed.), *Qualitative methodology* (pp. 15–26). Chicago, IL: Markham.

Bergman, M. M. (2010). Hermeneutic content analysis: Textual and audiovisual analyses within a mixed methods framework. In A. Tashakkori & C. Teddlie (Eds.), *Mixed methods in social & behavioral research* (2nd ed., pp. 379–396). Los Angeles, CA: Sage.

Bill & Melinda Gates Foundation. (2010). *Learning about teaching: Initial findings from the Measures of Effective Teaching project* [policy brief]. Retrieved from https://docs.gatesfoundation.org/documents/preliminary-finding-policy-brief.pdf

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.

Brown, A. (1992). Design experiments: Theoretical and methodological challenges in creating interventions in classroom settings. *The Journal of the Learning Sciences, 2*(2), 141–178.

Brown, T. M., & Rodriguez, L. F. (2009). *Youth in participatory action research: New directions in youth development, Number 123*. San Francisco, CA: Jossey Bass.

Brundage, A. (2013). *Going to the sources: A guide to historical research and writing* (5th ed.). Chichester, West Sussex, England: Wiley.

Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research, 6*(1), 97–113.

Bulgren, J. A., Lenz, B. K., Schumaker, J. B., Deshler, D. D., & Marquis, J. G. (2002). The use effectiveness of a comparison routine in diverse secondary content classrooms. *Journal of Educational Psychology, 94,* 356–371.

Buttorff, D. (2012). *The keys to motivation: Including student interest and goals.* Unpublished manuscript, Southern Oregon University, Ashland, Oregon.

Cahill, C., Rios-Moore, I., & Threatts, T. (2008). Different eyes/open eyes: Community-based participatory action research. In J. Cammarota, & M. Fine (Eds.). *Revolutionizing education: Youth participatory action research in motion* (pp. 89–124). New York, NY: Routledge.

Cammarota, J., & Fine, M. (Eds.). (2008). *Revolutionizing education: Youth participatory action research in motion*. New York, NY: Routledge.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81–105.

Campbell, D. T., & Russo, M. J. (2001). *Social measurement.* Los Angeles, CA: Sage.

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, IL: Rand McNally.

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline 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., . . . Sanders, A. G. (1940). Penicillin as a chemotherapeutic agent. *Lancet, 2,* 226.

Chandler, J. L. S. (2014). *Learning from the collusions, collisions, and contentions with White privilege experienced in the United States by White mothers of sons and daughters whose race is not White.* Doctoral dissertation, Cardinal Stritch University, Milwaukee, WI (available through the online database ProQuest Dissertations & Theses: Full text; publication number 3614469).

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). Los Angeles, 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.* Los Angeles, CA: Sage.

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). Los Angeles, CA: Sage.

Chatterjee, B. B., & Srivastava, A. K. (1982). A systematic method for drawing sociograms. *Perspectives in Psychological Researches, 5*(1), 1–6.

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.* Los Angeles, CA: Corwin Press.

Clandinin, D. J., & Connelly, F. M. (2000). *Narrative inquiry: Experience and story in qualitative research*. San Francisco, CA: Jossey Bass.

Clifford, J., & Marcus, G. E. (1986). *Writing culture: The poetics and politics of ethnography*. Berkeley, CA: University of California Press.

Cochran-Smith, M., & Lytle, S. L. (1993). *Inside/outside: Teacher research and knowledge.* New York, NY: 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.

Collins, A. (1992). Towards a design science of education. In E. Scanlon & T. O'Shea (Eds.), *New directions in educational technology* (pp. 15–22). New York, NY: Springer.

Cook, N. J., & Breedin, S. D. (1994). Constructing naive theories of motion on the fly. *Memory & Cognition, 22,* 474–493.

Corbin, J., & Strauss, A. C. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Los Angeles, CA: Sage.

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science, 19,* 51–57.

Creswell, J. W. (2010). Mapping the landscape of mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Mixed methods in social & behavioral research* (2nd ed., pp. 45–68). Los Angeles, CA: Sage.

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Los Angeles, CA: Sage.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Los Angeles, CA: Sage.

Creswell, J. W. (2015). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (5th ed.). Columbus, OH: Pearson.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Los Angeles, 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.)

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25,* 7–29.

Curran, W., Hague, E., & Gill, H. (2007). Practicing active learning: Introducing urban geography and engaging community in Pilsen, Chicago. In N. G. Hofman & H. Rosing (Eds.), *Pedagogies of praxis: Course-based action research in the social sciences* (pp. 79–93). Boston, MA: Anker.

Czechowski, K., Miranda, D., & Sylvestre, J. (2016). Like a rolling stone: A mixed-methods approach to linguistic analysis of Bob Dylan's lyrics. *Psychology of Aesthetics, Creativity, and the Arts, 10,* 99–113.

Dahlberg, B., Wittink, M. N., & Gallo, J. J. (2010). Funding and publishing integrated studies: Writing effective mixed methods manuscripts and grant proposals. In A. Tashakkori & C. Teddlie (Eds.), *Mixed methods in social & behavioral research* (2nd ed., pp. 775–802). Los Angeles, CA: Sage.

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.

DeCoster, J., Sparks, E. A., Sparks, J. C., Sparks, G. G., & Sparks, C. W. (2015). Opportunistic biases: Their origins, effects, and an integrated solution. *American Psychologist, 70,* 499–514.

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5–8.

Dewey, J. (1904). *The relation of theory to practice in education: The third NSSE yearbook* (pt. 1). Chicago, IL: University of Chicago Press.

Dewey, J. (1934). *Art as experience*. Toms River, NJ: Capricorn Books.

Dewey, J. (1938). *Experience and education*. New York, NY: Collier Books.

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.

Dyrness, A. (2011). *Mothers united: An immigrant struggle for socially just education*. Minneapolis, MN: University of Minnesota.

Eisner, E. W. (1998). *The enlightened eye: Qualitative inquiry and the enhancement of educational practice.* Columbus, OH: Merrill/Prentice Hall.

Elliott, D., & Stern, J. E. (Eds.). (1997). *Research ethics: A reader.* Hanover, NH: University Press of New England.

Ellis, C., Adams, T. E., & Bochner, A. P. (2011). Autoethnography: An overview. *Forum: Qualitative Social Research, 12*(1), Article 10.

Ellis, J. O. (1975). *Cognitive processes in logical problem-solving: The three-term series problem.* Unpublished doctoral dissertation, The Pennsylvania State University, University Park, PA.

Erickson, F. (1986). Qualitative methods in research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 119–161). New York, NY: MacMillan.

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.

Ferron, J. M., & Levin, J. R. (2014). Single-case permutation and randomization statistical tests: Present status, promising new developments. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case intervention research: Methodological and statistical advances* (pp. 153–183). Washington, DC: American Psychological Association.

Fidler, F., & Cumming, G. (2013). Effect size estimation and confidence intervals. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology: Vol. 2. Research methods in psychology* (2nd ed., pp. 142–163). Hoboken, NJ: Wiley.

Fine, M. (2008). An epilogue, of sorts. In J. Cammarota & M. Fine (Eds.), *Revolutionizing education: Youth participatory action research in motion*. New York, NY: Routledge.

Fleming, A. (1929). On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzæ*. *British Journal of Experimental Pathology, 10,* 226–236.

Forbes, M. L., Ormrod, J. E., Bernardi, J. D., Taylor, S. L., & Jackson, D. L. (1999, April). *Children's conceptions of space, as reflected in maps of their hometown.* Paper presented at the American Educational Research Association, Montreal, Quebec, Canada.

Fossey, D. (1983). *Gorillas in the mist.* Boston, MA: Houghton Mifflin.

Freeman, L. C. (2004). *The development of social network analysis: A study in the sociology of science.* North Charleston, SC: BookSurge, LLC.

Freire, P. (1970). *Pedagogy of the oppressed*. New York, NY: Basic Books.

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction* (8th ed.). Columbus, OH: Pearson.

Garson, G. D. (2013). *Grounded theory.* Asheboro, NC: Statistical Publishing Associates.

Geertz, C. (1973). *The interpretation of cultures*. New York, NY: Basic Books.

Gergen, K. J., Josselson, R., & Freeman, M. (2015). The promises of qualitative inquiry. *American Psychologist, 70,* 1–9.

Glaser, B. G. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory.* New York, NY: 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, IL: 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.

Gorard, S. (2010). Research design, as independent of methods. In A. Tashakkori & C. Teddlie (Eds.), *Mixed methods in social & behavioral research* (2nd ed., pp. 237–251). Los Angeles, 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.

Gould, S. J. (1985, June). The median isn't the message. *Discover, 6,* 40–42.

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.

Hacker, K. (2013). *Community-based participatory research*. Thousand Oaks, CA: Sage.

Hacker, K., Chu, J., Leung, C., Marra, R., Pirie, A., Brahimi, M., … Marlin, R. P. (2011). The impact of Immigration and Customs Enforcement on immigrant health: Perceptions of immigrants in Everett, Massachusetts, USA. *Social Science & Medicine, 73*(4), 586–594.

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains. *American Psychologist, 53,* 449–455.

Halpern, D. F. (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, NY: Erlbaum/Taylor & Francis.

Hämäläinen, R., & Oksanen, K. (2012). Challenge of supporting vocational learning: Empowering collaboration in a scripted 3D game—How does teachers' real-time orchestration make a difference? *Computers & Education, 59,* 281–293. https://doi.org/10.1016/j.compedu.2012.01.002

Harden, A., & Thomas, J. (2010). Mixed methods and systematic reviews: Examples and emerging issues. In A. Tashakkori & C. Teddlie (Eds.), *Mixed methods in social & behavioral research* (2nd ed., pp. 749–774.). Los Angeles, CA: Sage.

Harkness, S., Zylicz, P. O., Super, C. M., Welles-Nyström, B., Bermúdez, M. R., Bonichini, S., . . . Mavridis, C. J. (2011). Children's activities and their meanings for parents: A mixed-methods study in six Western cultures. *Journal of Family Psychology, 25,* 799–813.

Harwell, M. R., & Gatti, G. G. (2001). Rescaling ordinal data to interval data in educational research. *Review of Educational Research, 71,* 105–131.

He, M. F. (1999). A life-long inquiry forever flowing between China and Canada: Crafting a composite auto/biographic narrative method to represent three Chinese women teachers' cultural experiences. *Journal of Critical Inquiry Into Curriculum & Instruction, 1*(2), 5–29.

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, NY: Guilford Press.

Hess, C. (2016). *Associations between teacher interactional quality and student achievement: A classroom-level analysis of randomized and non-randomized teacher assignments in the Measures of Effective Teaching project.* Doctoral dissertation, University of Northern Colorado, Greeley, CO (soon to be available through the online database ProQuest Dissertations & Theses: Full text).

Hesse-Biber, S. N. (2010). *Mixed methods research: Merging theory with practice.* New York, NY: Guilford Press.

Hofman, N. G., & Rosing, H. (Eds.). (2007). *Pedagogies of praxis: Course-based action research in the social sciences*. Boston, MA: Anker.

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.

Irizarry, J. G. (2011). *The Latinization of U.S. schools: Successful teaching and learning in shifting cultural contexts*. Boulder, CO: Paradigm Press.

Jaccard, J., & Jacoby, J. (2010). *Theory construction and model-building skills.* New York, NY: Guilford Press.

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, CO.

Jacobson, R. M., Hanson, W. E., & Zhou, H. (2015). Canadian psychologists' test feedback training and practice: A national survey. *Canadian Psychology, 56,* 394–404.

Johnson, B. (2001). Toward a new classification of nonexperimental quantitative research. *Educational Researcher, 30*(2), 3–13.

Johnson, L. R. (2017). *Community-based qualitative research: Approaches for education and the social sciences*. Thousand Oaks, CA: Sage.

Johnson, L. R., Sokacich, D., Rutledge, E., & Rosario-Ramos, E. M. (2016, April). *Promoting legitimate knowledge and transformational resistance amongst young mothers: The Proyecto Atabey intergenerational mentoring program*. Paper presented at the American Educational Research Association, Washington, DC.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14–26.

Jones, S. (Ed.). (2017). *Writing and teaching to change the world: Connecting with our most vulnerable students*. New York, NY: Teachers College Press.

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., Blazevski, J. B., Bonney, C. R., . . . Kelly, K. L. (2007). Cognitive processing of self-report items in educational research: Do they think what we mean? *Educational Psychologist, 42,* 139–151.

Kellogg, R. T. (1994). *The psychology of writing.* New York, NY: 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, CO.

Kontos, S. (1999). Preschool teachers' talk, roles, and activity settings during free play. *Early Childhood Research Quarterly, 14*(3), 363–382.

Kosinski, M., Matz, S. C., Gosling, S. D., Popov, V., & Stillwell, D. (2015). Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. *American Psychologist, 70,* 543–556.

Kozinets, R. V. (2010). *Netnography: Doing ethnographic research online.* London, England: 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.

Krueger, R. A., & Casey, M. A. (2009). *Focus groups: A practical guide for applied research* (4th ed.). Los Angeles, CA: Sage.

Kuhn, D. (1995). Microgenetic study of change: What has it told us? *Psychological Science, 6,* 133–139.

Kuhn, D., Goh, W., Iordanou, K., & Shaenfield, D. (2008). Arguing on the computer: A microgenetic study of developing argument skills in a computer-supported environment. *Child Development, 79,* 1310–1328.

Kvale, S., & Brinkmann, S. (2009). *InterViews: Learning the craft of qualitative research interviewing* (2nd ed.). Los Angeles, CA: Sage.

Lara, L. G. (2009). *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.* Doctoral dissertation, University of Northern Colorado, Greeley, CO. (Available through the online database ProQuest Dissertations & Theses: Full Text; publication number 3397099)

Lather, P. (1986). Research as praxis. *Harvard Educational Review, 56*(3), 257–277.

Lather, P. (2006). Paradigm proliferation as a good thing to think with: Teaching research in education as a wild profusion. *International Journal of Qualitative Studies in Education, 19*(1), 35–57.

Lauer, J. M., & Asher, J. W. (1988). *Composition research: Empirical designs.* New York, NY: Oxford University Press.

Laursen, B., Bukowski, W. M., Aunola, K., & Nurmi, J.-E. (2007). Friendship moderates prospective associations between social isolation and adjustment problems in young children. *Child Development, 78,* 1395–1404.

Lawrence-Lightfoot, S., & Davis, J. H. (1997) *The art and science of portraiture*. San Francisco, CA: Jossey-Bass.

Ledgerwood, A. (2016). Introduction to the special section on improving research practices: Thinking deeply across the research cycle. *Perspectives on Psychological Science, 11,* 661–663.

Leung, C. Y. Y. (2012). *The immigration experiences, acculturation, and parenting of Chinese immigrant mothers.* Unpublished doctoral dissertation, University of Maryland, Baltimore County, MD.

Lewin, K. (1948). *Resolving social conflicts*. New York, NY: Harper & Row.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry.* Los Angeles, CA: Sage.

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.

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.

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, NY: Routledge.

Mackinnon, S. P., Sherry, S. B., Pratt, M. W., & Smith, M. M. (2014). Perfectionism, friendship intimacy, and depressive affect in transitioning university students: A longitudinal study using mixed methods. *Canadian Journal of Behavioural Science, 46,* 49–59.

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)

Mandler, J. M. (2007). On the origins of the conceptual system. *American Psychologist, 62,* 741–751.

Martin, V. B., & Gynnild, A. (Eds.). (2011). *Grounded theory: The philosophy, method, and work of Barney Glaser.* Boca Raton, FL: BrownWalker Press.

Maurois, A. (1959). *The life of Alexander Fleming: Discoverer of penicillin.* New York, NY: E. P. Dutton.

Maxwell, J. A., & Mittapalli, K. (2010). Realism as a stance for mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Mixed methods in social & behavioral research* (2nd ed., pp. 145–167). Los Angeles, CA: Sage.

Mayeux, L., Houser, J. J., & Dyches, K. D. (2011). Social acceptance and popularity: Two distinct forms of peer status. In A. H. N. Cillessen, D. Schwartz, & L. Mayeux (Eds.), *Popularity in the peer system* (pp. 79–102). New York, NY: Guilford Press.

McCallin, R. C. (1988). *Knowledge application orientation, cognitive structure, and achievement.* Unpublished doctoral dissertation, University of Northern Colorado, Greeley, CO.

McCaslin, M., Vega, R. I., Anderson, E. E., Calderon, C. N., & Labistre, A. M. (2011). Tabletalk: Navigating and negotiating in small-group learning. In D. M. McInerney, R. A. Walker, & G. A. D. Liem (Eds.), *Sociocultural theories of learning and motivation: Looking back, looking forward* (pp. 191–222). Charlotte, NC: Information Age.

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.

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.

Mehan, H. (1979). *Social organization in the classroom.* Cambridge, MA: Harvard University Press.

Mercier, H., Boudry, M., Paglieri, F., & Trouche, E. (2017). Natural-born arguers: How to make the best of our reasoning abilities. *Educational Psychologist, 52,* 1–16.

Merten, D. E. (2011). Being there awhile: An ethnographic perspective on popularity. In A. H. N. Cillessen, D. Schwartz, & L. Mayeux (Eds.), *Popularity in the peer system* (pp. 57–76). New York, NY: Guilford Press.

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, IL.

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.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Los Angeles, CA: Sage.

Mills, G. E. (2018). *Action research: A guide for the teacher researcher* (6th ed.). New York, NY: Pearson.

Mills, G. E., & Gay, L. R. (2016). *Educational research: Competencies for analysis and application* (11th ed.). Columbus, OH: Pearson.

Mishler, E. G. (1995). Models of narrative analysis: A typology. *Journal of Narrative and Life History, 5*(2), 87–123.

Mitchell, G. (2012). Revisiting truth or triviality: The external validity of research in the psychological laboratory. *Perspectives on Psychological Science, 7,* 109–117.

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, RI.

Morelli, G. A., & Rothbaum, F. (2007). Situating the child in context: Attachment relationships and self-regulation in different cultures. In S. Kitayama & D. Cohen (Eds.), *Handbook of cultural psychology* (pp. 500–527). New York, NY: Guilford Press.

Morrell, E. (2004). *Becoming critical researchers: Literacy and empowerment for urban youth*. New York, NY: Peter Lang.

Morrell, E. (2008). Six summers of YPAR: Learning, action, and change in urban education. In J. Cammarota & M. Fine (Eds.), *Revolutionizing education: Youth participatory action research in motion* (pp. 155–183). New York, NY: Routledge.

Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological Science, 25,* 1159–1168.

Muncey, T. (2010). *Creating autoethnographies*. Thousand Oaks, CA: Sage.

Munter, M., & Paradi, D. (2009). *Guide to PowerPoint.* Hoboken, NJ: Pearson.

Murphy, K. R., Myors, B., & Wolach, A. (2009). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests* (3rd ed.). New York, NY: Routledge.

Nelson, S. K., Kushlev, K., English, T., Dunn, E. W., & Lyubomirsky, S. (2013). In defense of parenthood: Children are associated with more joy than misery. *Psychological Science, 24,* 3–10.

Neuman, W. L. (2011). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Boston, MA: 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.

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.

Nieto, S. (2003). *What keeps teachers going?* New York, NY: Teachers College Press.

Nussbaum, E. M. (2008). Collaborative discourse, argumentation, and learning: Preface and literature review. *Contemporary Educational Psychology, 33,* 345–359.

Nygreen, K. (2013). *These kids: Identity*, *agency*, *and social justice at a last chance high school*. Chicago, IL: University of Chicago Press.

O'Cathain, A. (2010). Assessing the quality of mixed methods research: Toward a comprehensive framework. In A. Tashakkori & C. Teddlie (Eds.), *Mixed methods in social & behavioral research* (2nd ed., pp. 531–555). Los Angeles, CA: Sage.

Ollerenshaw, J., & Creswell, J. W. (2002). Narrative research: A comparison of two restorying data analysis approaches. *Qualitative Inquiry, 8*(1), 329–347.

Onwuegbuzie, A. J., & Combs, J. P. (2010). Emergent data analysis techniques in mixed methods research: A synthesis. In A. Tashakkori & C. Teddlie (Eds.), *Mixed methods in social & behavioral research* (2nd ed., pp. 397–430). Los Angeles, CA: Sage.

Onwuegbuzie, A. J., & Leech, N. L. (2005). Taking the "Q" out of research: Teaching research methodology courses without the divide between quantitative and qualitative paradigms. *Quality and Quantity, 39,* 267–296.

Onwuegbuzie, A. J., & Teddlie, C. (2003). A framework for analyzing data in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 351–383). Los Angeles, CA: Sage.

Ormrod, J. E. (1979). Cognitive processes in the solution of three-term series problems. *American Journal of Psychology, 92,* 235–255.

Ormrod, J. E. (2016). *Human learning* (7th ed.). Columbus, OH: 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., & Trahan, R. G. (1982). Can signs help visitors control their own behavior? *Trends, 19*(4), 25–27.

Pardo-Ballester, C., & Rodriguez, J. C. (2009). Using design-based research to guide the development of on-line instructional materials. In C. A. Chapelle, H. G. Jun, & I. Katz (Eds.). *Developing and evaluating language learning materials* (pp. 86–102). Ames, IA: Iowa State University.

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide.* Oxford, United Kingdom: Blackwell.

Phillion, J. (2002). Classroom stories of multicultural teaching and learning. Special issue on narrative multiculturalism. *Journal of Curriculum Studies, 34*(3), 281–300.

Polkinghorne, D. E. (1989). Phenomenological research methods. In R. S. Valle & S. Halling (Eds.), *Existential-phenomenological perspectives in psychology* (pp. 41–60). New York, NY: Plenum.

Rakison, D. H., & Oakes, L. M. (Eds.). (2003). *Early category and concept development: Making sense of the blooming, buzzing confusion.* Oxford, England: Oxford University Press.

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, NH.

Ravert, R. D., Gomez-Scott, J., & Donnellan, M. B. (2015). Equivalency of paper versus tablet computer survey data. *Educational Researcher, 44,* 308–310.

Reiser, B. J., Tabak, I., Sandoval, W. A., Smith, B. K., Steinmuller, F., & Leone, A. J. (2001). BGuILE: Strategic and conceptual scaffolds for scientific inquiry in biology classrooms. In S. M. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 263–305). Mahwah, NJ: Lawrence Erlbaum Associates.

Rief, L. (2014). *ReadWriteTeach: Choice and challenge in the reading-writing workshop*. Portsmouth, NJ: Heinemann.

Roediger, R. H. (2004). What should they be called? *APS Observer, 17*(4), 46–48. Retrieved from http://journal.sjdm.org/roediger.html

Rogelberg, S. G., & Luong, A. (1998). Nonresponse to mailed surveys: A review and guide. *Current Directions in Psychological Science, 7,* 60–65.

Rosing, H. (2007). Food for more than thought: Course-based action research on corner stores in Chicago. In N. G. Hofman & H. Rosing (Eds.), *Pedagogies of praxis: Course-based action research in the social sciences* (pp. 1–20). Bolton, MA: Anker Publishing Company, Inc.

Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). London, England: Sage.

Sales, B. D., & Folkman, S. (Eds.). (2000). *Ethics in research with human participants.* Washington, DC: American Psychological Association.

Sandoval, W. A., & Bell, P. (2004). Design-based research for studying learning in context: Introduction. *Educational Psychologist, 39*(4), 199–201.

Schram, T. H. (2006). *Conceptualizing and proposing qualitative research* (2nd ed.). Columbus, OH: 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, NY: 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.

Seidman, I. (2013). *Interviewing as qualitative research: A guide for researchers in education and the social sciences*. New York, NY: Teachers College Press.

Shadish, W. R. (2014). Statistical analyses of single-case designs: The shape of things to come. *Current Directions in Psychological Science, 23,* 139–146.

Shadish, W. R., Hedges, L. V., Pustejovsky, J. E., Rindskopf, D. M., Boyajian, J. G., & Sullivan, K. J. (2014). Analyzing single-case designs: *d, G,* hierarchical models, Bayesian estimators, generalized additive models, and the hopes and fears of researchers about analyses. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case intervention research: Methodological and statistical advances* (pp. 247–281). Washington, DC: American Psychological Association.

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, CO.

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, NY: Guilford Press.

Shank, G. D. (2006). *Qualitative research: A personal skills approach* (2nd ed.). Columbus, OH: Pearson.

Shavinina, L. V., & Ferrari, M. (2004). Extracognitive facets of developing high ability: Introduction to some important issues. In L. V. Shavinina & M. Ferrari (Eds.), *Beyond knowledge: Extracognitive aspects of developing high ability* (pp. 3–13). Mahwah, NJ: Erlbaum.

Sheehan, K. B. (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

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information, 22,* 63–75.

Shernoff, D. J., Kelly, S., Tonks, S. M., Anderson, B., Cavanagh, R. F., Sinha, S., & Abdi, B. (2016). Student engagement as a function of environmental complexity in high school classrooms. *Learning and Instruction, 43,* 52–60.

Silverman, D. (1993). *Interpreting qualitative data: Methods for analysing talk, text and interaction.* London, England: 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., & Ahlborg, 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, NY.

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.

Spradley, J. P. (1980). *Participant observation*. Belmont, CA: Wadsworth.

Steiner, E. (1988). *Methodology of theory building.* Sydney, Australia: Educology Research Associates.

Stenhouse, L. (1981). What counts as research? *British Journal of Educational Studies, 29*(2), 103–211.

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.* Los Angeles, CA: Sage.

Strunk, W., Jr., & White, E. B. (2009). *The elements of style* (50th anniversary ed.). New York, NY: Pearson/Longman.

Student (pseudonym for W. S. Gosset). (1908). The probable error of a mean. *Biometrika, 6,* 1–25.

Survey Research Center, Institute for Social Research at the University of Michigan. (1976). *Interviewer's manual* (Rev. ed.). Ann Arbor, MI: Author.

Tay, L., Parrigon, S., Huang, Q., & LeBreton, J. M. (2016). Graphical descriptives: A way to improve data transparency and methodological rigor in psychology. *Perspectives on Psychological Science, 11,* 692–701.

Teddlie, C., & Tashakkori, A. (2010). Overview of contemporary issues in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Mixed methods in social & behavioral research* (2nd ed., pp. 1–41). Los Angeles, CA: Sage.

Thakor, M. V., Borsuk, W., & Kalamas, M. (2004). Hotlists and Web browsing behavior—an empirical investigation. *Journal of Business Research, 57,* 776–786.

Thompson, B. (2008). *Foundations of behavioral statistics.* New York, NY: Guilford Press.

Thompson, K. R. (2006). Axiomatic theories of intentional systems: Methodology of theory construction. *Scientific Inquiry Journal, 7*(1), 13–24.

Thrailkill, N. J. (1996). *Imagery-evoking and attention-attracting material as facilitators of learning from a lecture.* Unpublished doctoral dissertation, University of Northern Colorado, Greeley, CO.

Tolentino, E. P., & Lawson, L. (2017). "Well, we're going to kindergarten, so we're gonna need business cards!": A story of preschool emergent readers and writers and the transformation of identity. *Journal of Early Childhood Literacy, 17*(1), 47–68.

Trahan, R. G. (1978). *Social science research: Rocky Mountain National Park* (Contract agreement PX 1520-8-A529). Greeley, CO: Author.

Tyler, K. M., Uqdah, A. L., Dillihunt, M. L., Beatty-Hazelbaker, R., Connor, T., Gadson, N., . . . Stevens, R. (2008). Cultural discontinuity: Toward a quantitative investigation of a major hypothesis in education. *Educational Researcher, 37,* 280–297.

Uziel, L. (2010). Rethinking social desirability scales: From impression management to interpersonally oriented self-control. *Perspectives on Psychological Science, 5,* 243–263.

von Stumm, S., Hell, B., & Chamorro-Premuzic, T. (2011). The hungry mind: Intellectual curiosity is the third pillar of academic performance. *Perspectives on Psychological Science, 6,* 574–588.

Wainer, H. (1984). How to display data badly. *American Statistician, 38,* 137–147.

Walton, D. N. (2003). *Ethical argumentation.* New York, NY: Lexington Books.

Ward, C., Bochner, S., & Furnham, A. (2001). *The psychology of culture shock* (2nd ed.). London, England: Routledge.

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.

Wheelan, C. (2013). *Naked statistics: Stripping the dread from the data.* New York, NY: Norton.

Winchester, H. P. M. (1999). Interviews and questionnaires as mixed methods in population geography: The case of lone fathers in Newcastle, Australia. *Professional Geographer, 51*(1), 60–67.

Witt, E., & Best, J. (April 21, 2008). *How different are people who don't respond to pollsters?* Washington, DC: Pew Research Center. Retrieved from http://pewresearch.org/pubs/807/

Wolcott, H. F. (1994). *Transforming qualitative data: Description, analysis, and interpretation.* Los Angeles, 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

Yow, V. R. (2015). *Recording oral history: A guide for the humanities and social sciences* (3rd ed.). Lanham, MD: Rowman & Littlefield.

YuliEfiyante, A., Nirbito, J. G., Wahjoedi, & Sunaryanto. (2016). The implementation of design-based research in the development of entrepreneurial learning for fisherwomen in the south coast of Malang Regency, Indonesia. *IOSR Journal of Humanities and Social Science, 21*(9), 37–45.

Zambo, D. (2003). *Uncovering the conceptual representations of students with reading disabilities.* Unpublished doctoral dissertation, Arizona State University, Tempe, AZ.

Zirkel, S., Garcia, J. A., & Murphy, M. C. (2015). Experience-sampling research methods and their potential for education research. *Educational Researcher, 44,* 7–16.

Index

A

ABAB design, 235n
Abbreviations, use of, 168
Abduction, 45
Abraham, E. P., 48
Abrams, E., 190, 438
Abscissa, 176
Abstract, reports, 417–418
Academic integrity, 416–417
Accidental sampling, 205
Acknowledgments, 417, 418
Action research, 31, 116
  community-based participatory research (CBPR), 323
  community-based qualitative research (CBQR), 323–324
  community-based research (CBR) models, 321–324
  course-based action research (CBAR), 322
  data collection and analysis, 327–328
  defined, 28–29, 315
  design-based research, 318–319
  design selection decisions, 324–326
  disseminating findings, 328–331
  participatory action research (PAR), 319–321
  sample research report, 331–335
  teacher research, 316–318
  youth participatory action research (YPAR), 319–321
Adams, T. E., 262
Advance organizer, 38, 99, 413
Ahlborg, G., 264
Allen, E. M., 165
Alpha (α), significance level, 364–369
Alpha coefficient, Cronbach, 191n
Alpha error, 365–369
Alternating-treatments design, 235, 244
Altheide, D. L., 281
American Psychological Association (APA)
  code of ethics, 140
  style guide, 91–92, 154–155, 419–421
Analysis of covariance (ANCOVA), 227, 368
Analysis of variance (ANOVA), 368
Analytical memos, 268, 387
Analytic terms, in qualitative data analysis, 391
Anderson, C. A., 119n
Anderson, E. E., 211
APA (American Psychological Association) style, 91–92, 154–155, 419–421
Appendices, writing of, 160, 421
Applied research, defined, 52
a priori hypothesis, 27, 67–68, 75, 264, 415
Argument analysis, 42
Arithmetic average, 350
Artifacts, 262
Asher, J. W., 258
Assessment
  defined, 120
  identifying validity and reliability issues, 133–135
  measurement scales, types of, 123–126
  reliability, 127, 130–133
  standardization, 132
  statistical error reduction, 367
  substantial and intangible phenomena, 120–123
  validity, 127, 128–130, 133, 224
Assumptions, 26–27, 30–31, 72, 75
Audiovisual materials
  case study designs, 261
  grounded theory study, 264
  interviews, technology tools for, 280–281
  recording observations, 273–274
Audit trail, 271
Aunola, K., 358
Authorship, sharing of, 433
Autoethnography, 262
Average, arithmetic, 350
Average deviation, 354–355, 356
Axial coding, 388

B

Bakari, R., 167–170
Ballenger, C., 317
Bandura, A., 69
Barab, S., 318
Bartholomew, D. J., 120n
Baseline data, 234
Basic research, defined, 52
Bazeley, P., 396
Beck, C. T., 428n
Becker, A. B., 321, 323
Becker, H. C., 373
Bell, P., 318
Bell curve, 347–348
Bergman, M. M., 266
Bernardi, J. D., 389
Best, J., 211
Beta error, 365–339
Bias, 30–31
  academic integrity and, 416
  confirmation bias, 47, 67, 270
  defined, 209
  in experimental designs, 249–251
  grounded theory study, 264
  hypothesis, stating of, 67
  identification of, 212–214
  instrumentation bias, 211
  Internet, use in data collection, 199
  interviews, cultural differences and, 277
  interviews, guidelines for, 187–188
  mixed-methods designs, dissertation example, 404
  open coding strategy, 382
  phenomenological study, 263
  positionality, 268
  purpose statement, writing of, 59
  qualitative research, credibility and transferability, 269–271, 281
  qualitative research, data analysis and, 280–381
  qualitative research, dissertation example, 286–287
  questionnaires, construction of, 192
  reports, writing of, 415
  researcher bias, 212, 392–393
  research proposal, writing of, 157
  response bias, 211–212
  sampling bias, 209–210
  social desirability effect, 128, 212
  telephone interviews, 182
Bibliographic database software, 94
Bibliography, 164n
Bill & Melinda Gates Foundation, 102
Biserial correlation, 359
Bochner, A. P., 262
Bochner, S., 277
Books, literature review, 84–86, 96
Borg, W. R., 188, 281
Borsuk, W., 296
Boudry, M., 42
Bracketing, 263
Brainstorming software, 64–65

Bransford, J. D., 45
Breedin, S. D., 45
Brinkmann, S., 279
Brown, A., 318
Brown, A. L., 45
Brown, T. M., 329
Bryman, A., 292
Bukowski, W. M., 358
Bulgren, J. A., 238
Bulleted lists, 163–164
Bulletin boards, online, 46–47
Bushman, B. J., 119n

C
Cahill, C., 320
Calderon, C. N., 211
Call numbers, 85–86
Cammarota, J., 329
Campbell, D. T., 130, 223, 224, 227n
Campbell, S. B., 246
Caracelli, V. J., 292
Case study, 266
  characteristics of, 116
  data organization strategies, 381
  dissertation example, 283–287
  multiple or collective case study, 261
  one-shot experimental case study, 228–229
  overview of, 260–261
  sampling decisions, 272
Casey, M. A., 280
Caspi, A., 70
Causal-comparative designs, 239n
Cause-and-effect relationships
  common pitfalls in thinking, 47
  control, importance of, 222–227
  correlational research and, 177–178
  correlation and, 361
  experimental designs and, 220–221, 246–247
  inferential statistical tests, summary table, 368
  longitudinal mixed-methods designs, 298
  qualitative studies, 260
  reports, writing of, 414–415
Central limit theorem, 362n
Central tendency, measures of, 349–353, 357
Cepeda, N. J., 248, 249, 250
Chain, E. B., 48
Chain sampling, 272
Chamorro-Premuzic, T., 358
Chandler, J., 273, 273n
Charmaz, K., 264, 387, 389
Chatterjee, B. B., 123
Cheah, C., 402
Checklist, descriptive studies, 183–187
Chicago Manual of Style, 91–92, 418
Chipman, S. F., 192
Chi-square goodness-of-fit test, 368
Chi-square test, 124
Christoffel, K. K., 321
Circular definition, 72
Citations
  footnotes and endnotes, 418
  literature review, writing of, 100–101
  reference lists, 418–421
  research proposal, writing of, 160
  review of, 164
  source identification, style guides, 91–92
  use in literature search, 91
Cizek, G. J., 100
Clandinin, D. J., 264
Clifford, J., 263
Cluster sampling, 203, 205
Coauthors, 433
Cochran-Smith, M., 316, 317
Cocking, R. R., 45
Coding and subcoding of data, tips for, 381–386, 393
  grounded theory study example, 387–389
  mixed-methods designs, 396–398
  mixed-methods designs, dissertation example, 403–406
Coding clean-up, 383
Coefficient of determination, 359
Coghill, R. D., 48
Cohort-sequential study, 179
Cole, D. B., 202, 213
Collective case study, 261
Collective coding, 328
Collins, A., 318
Collins, J., 236
Combs, J. P., 304
Community-based participatory research (CBPR), 323, 327–331
Community-based qualitative research (CBQR), 323–324, 327–331
Community-based research (CBR) models, 321–324, 327–331
Computer technology
  data organization and analysis tools, 391–392
  electronic discussion groups, 46–47
  electronic spreadsheets, 342–343
  Excel spreadsheets, uses for, 436–443
  experience-sampling method (ESM), 180
  Internet search engines, 90–91
  interview data collection, 188, 280–281
  library, efficient use of, 93–96
  library catalogs, use of, 84–86
  mind-mapping (brainstorming) software, 64–65
  mixed-methods data analysis tools, 398–399
  observations, recording of, 186–187
  online databases, 86–89
  presentation software, 431–432
  project management software, 144–146
  questionnaires, administration and data analysis, 193–194
  as research tool, 33, 34
  Skype interviews, 182
  SPSS (Statistical Package for the Social Sciences), 444–449
  statistical software packages, 370
  tables and figures, creation of, 413
Conceptual analysis exercises, 71, 126–127, 133–135, 245–346, 267–268, 299–301, 324–325
Conceptual density, 264
Conceptual framework, 66–67, 75, 130, 428
Conclusion, research reports, 38–39, 72–73, 429
Conclusions, inductive reasoning and, 43–44, 47
Conference presentations, 329, 430–434
Confidence interval, 364
Confidentiality, 138, 169, 269, 416
Confirmability, 269–271
Confirmation bias, 47, 67, 270
Confirmatory analysis, 368
Confounding variables
  controlling for, 223–227
  defined, 223
  dissertation example, 252, 253
  reports, writing of, 416
  statistical control for, 227
  tips for identifying, 246–247
Connelly, F. M., 264
Connelly, L. M., 246
Consent, informed, 136–138
  ethnography, 262
  Internet, use in data collection, 198, 249, 280–281
  interviews, guidelines for, 188, 278
  placebo use, 225
  qualitative research, data collection, 269
Constructivism, as philosophical orientation, 31
Construct validity, 129
Content analysis, 116, 265–266, 267, 381, 389–390
Content validity, 128–129
Contingency coefficient, 359
Continuous variables, 345
Contradictory evidence, importance of looking for, 47, 118, 270
Control group, 223, 225
  experimental design summary table, 243–244
  nonrandomized control-group pretest-posttest design, 233
  pre-experimental designs, 228–230
  true experimental designs, 230–232
Control-group posttest-only design, 231, 243
Control-group pretest-posttest design, 230, 243
Control-group time-series design, 234, 244
Convenience sampling, 205, 272, 272n
Convergent designs, 293–295
Cook, N. J., 45
Copyright, 101, 417
Corbin, J., 264, 264n, 387, 389
Core category, data coding, 388–389
Correlation
  common pitfalls in thinking, 47
  defined, 176
  statistical measures of, 358–361, 368
Correlational research, 176–178
  characteristics of, 116

vs. ex post facto designs, 239–240
longitudinal studies and, 178–179
Correlational study, defined, 176
Correlation coefficient, 54, 177, 358–361
Council of Science Editors (CSE), 92
Course-based action research (CBAR), 322
Cover term, in qualitative data analysis, 391
Cowan, N., 35n
Credibility, 117–118, 269–271, 303–304
Creswell, J. W., 188, 261, 265, 268, 275, 281, 291, 293n, 303, 387, 388, 396
Criteria for the admissibility of data, 111
Criterion validity, 129
Critical thinking, 41–42, 47
Cronbach's alpha coefficient, 191n
Cross-sectional study, 178–179
CSE (Council of Science Editors), 92
Cuca, J. M., 165
Cultural domains, in qualitative research, 391
Cultural knowledge, in qualitative research, 391
Culture
interviews, strategies for, 277
questionnaires, construction of, 192
Cumming, G., 369n
Curran, W., 322, 322n
CyberTracker, 187
Cyclical process, research as, 29–30
Czechowski, K., 294

D

Dahlberg, B., 165
Data
correlational research, 176–178
criteria for admissibility, 111
dynamics in, 339
interpretation of, 29–30
mixed-methods, 28–29
ordinal and interval data, 185
primary and secondary data, 109–110
qualitative research, 28
quantitative research, 28
transient nature of, 108
use of term, 28
Data analysis. *See also* Statistics
action research, 317, 327–328
computer tools for, 33, 34, 391–392
content analysis study example, 389–390
data analysis spiral, Creswell, 387, 388
ethnographic study example, 390–391
Excel spreadsheets, uses for, 442–443
grounded theory studies, 264, 387–389
iterative processes, qualitative methods, 259, 380–381
mixed-methods designs, 304, 395–399
mixed-methods designs, dissertation example, 309–310, 402–406
narrative inquiry, 265
qualitative approaches, 114
qualitative research, analysis strategies, 114, 381–387
qualitative research, credibility and transferability, 269–271
qualitative research, evaluation checklist, 400–402
qualitative research, planning checklist, 393–395
qualitative research, researcher-as-instrument, 392–393
qualitative research, sample decisions, 271–273
qualitative research, dissertation example, 283–287
questionnaires, 193–194
reports, writing of, 412–414, 429
research proposal, writing of, 158–160
SPSS (Statistical Package for the Social Sciences), 444–449
systematic reviews, 399–400
Databases, electronic, 53
library catalogs, use of, 84–86
literature review, 86–89
Data collection
action research, overview, 316–317, 327–328
action research, sample report, 334–335
audit trail, 271
baseline data, 234
case study designs, 261
coding and subcoding strategies, 381–386
computer tools for, 33, 34
content analysis designs, 265–266
credible and trustworthy data collection, 269–271
data admissibility criteria, 111
descriptive studies, 182–199
emerging design, qualitative research, 269
ethnography, 261–263
Excel spreadsheets, uses for, 438–442
experience-sampling method (ESM), 180
explanatory sequential designs, 296–297
exploratory sequential designs, 295–296
feasibility of, 57
grounded theory study, 264
Internet-based studies, 248–249
Internet use in descriptive studies, 198–199
interviews, guidelines for, 187–188
interviews, qualitative research, 274–281
interviews, technology tools for, 280–281
iterative processes, qualitative methods, 259, 298–299, 380–381
member checking, 271
mixed-methods designs, dissertation example, 307–308
multiphase iterative designs, 298–299
narrative inquiry, 264–265
observations, computerizing of, 186–187
participatory action research, 320
phenomenological study, 263
planning for, 107–108, 110–111
qualitative research, observations, 112–117, 273–274
qualitative research, overview, 268–269
qualitative research, sample decisions, 271–273
qualitative research, dissertation example, 283–287
quantitative methodologies, 112–117
reflexivity, 270
report, writing of, 412, 429
research proposal, writing of, 158
sociogram, 121–122
sociometric matrix, 122–123
survey research, 181–182
teacher (action) research, 316–317
triangulation, 270
Data dredging, 367
Data interpretation. *See also* Data analysis
computer tools for, 33, 34
dissertation example, 373–376
iterative nature of qualitative research, 259, 380–381
reports, writing of, 414–415, 429
research projects, selection of, 53–54
statistical analysis and interpretation, 371–373
subproblems, 62
Data point (datum), 28
Davis, J. H., 270
Davitz, J. R., 165
Davitz, L. L., 165
Deaver, C. M., 237, 238
Deciles, 354n
DeCoster, J., 367
Dedication page, in a research report, 417
Deductive logic, 42–43, 47
Defining terms, 71–72
Delimitations, 72–73, 75, 415
Dellve, L., 264
Dependability, 269
Dependent variables
conceptual analysis exercise, 71–73
defined, 68
experimental designs, overview, 221
experimental design summary table, 243–244
ex post facto designs, 239–240
factorial designs, 241–242, 245
statistical error reduction, 367
Descriptive research
bias, identification of, 212–214
checklists, rating scales, and rubrics, 183–186
computerizing observations, 186–187
correlational research, 176–178
data collection, planning for, 182–199
data interpretation, 214
defined, 174
developmental designs, 178–179
dissertation example, 215–218
experience-sampling method (ESM), 180
instrumentation bias, 211
Internet, use in data collection, 198–199
interviews, guidelines for, 187–188
nonprobability sampling, 205–206

Descriptive research (*continued*)
observation studies, 175–176
probability sampling, 200–205
questionnaires, administration and data analysis, 193–194
questionnaires, construction of, 188–193
questionnaires, maximizing return rate, 194–198
response bias, 211–212
sample population choices, 199–207
sample size, decisions about, 207–209
sampling bias, 209–210
social desirability effect, 212
survey research, 181–182
very large populations, 206–207
Descriptive statistics, 33–35, 344, 349
measures of association: correlation, 358–361
measures of central tendency, 349–353
measures of variability, 353–357
SPSS (Statistical Package for the Social Sciences), 446–447
Deshler, D. D., 238
Design-based research, 116, 245–246, 299, 318–319
Design-Based Research Collective, 318
Developmental research, 116, 178–179
Deviation, average, 354–355, 356
Deviation, standard, 355
Dewey, J., 264, 316
Dewey decimal (DD) classification system, 85–86
Digital Object Identifier (DOI), 91, 420
Discrepant voices, as a strategy for enhancing credibility, 118
Discrete variables, 345
Discriminant sampling, 272
Distributions
normal and non-normal, 346–348
polymodal, 352
population mean estimates, 362–363
Do, S. L., 264
Documentation. *See also* Reports
case study designs, 261
content analysis designs, 265–266
ethnography designs, 262
memos, data collection and, 268
narrative inquiry, 265
observations, recording of, 273–274
qualitative research, rigorous subjectivity, 392–393
qualitative research, sample decisions, 271–273
DOI (Digital Object Identifier), 91, 420
Domain analysis, in qualitative data analysis, 391
Donellan, M. B., 193
Double-blind experiment, 225, 236, 295
Dowson, M., 211
Dunn, E. W., 177
Dyches, K. D., 65
Dyrness, A., 320, 329

E
Ecological validity, 119n
Editing. *See also* Writing proposals and reports
common editing marks, 161
literature review, 101–102
purpose statement, 60–61
questionnaires, construction of, 192
research proposal, revision of, 160–166
word-processing software, 41
writing first section of proposal, 76–77
Effect size (ES), 369
Eisenhardt, M., 281
Eisner, E. W., 268, 275, 278, 281
Eklöf, M., 264
Electronic databases, 53, 94
Electronic discussion groups, 46–47
Electronic planners, 144–146
Electronic questionnaires, 193–194
Electronic spreadsheets, 342–343, 392, 436–443
Elliott, D., 433
Ellis, C., 262
Ellis, J. O., 294, 295
Emerging design, 269
Emotions, common pitfalls in thinking, 47
End matter, 417, 418–421
Endnotes, 418
English, T., 177
Environment
confounding variables, controlling for, 223
ethnography designs, 261–263
Epistemology, as philosophical orientation, 30n
Epoché, 263, 308
Equivalent-forms reliability, 131
Equivalent time-sample design, 227n
Erickson, F., 270
ESM (experience-sampling method), 180, 226n, 234
Estimates of population parameters, 344–345
Ethics
academic integrity, 416–417
authorship, sharing of, 433
codes of, 140
honesty with colleagues, 139
institutional review board (IRB), 139–140
Internet, use in data collection, 198, 249
mixed-methods designs, 304–305
placebo use, 225
privacy rights, 138
protection from harm, 135–136
qualitative research, data collection, 269
voluntary and informed participation, 136–138
Ethnography, 266
characteristics of, 116
data analysis example, 390–391
overview of, 261–263
sampling decisions, 272
Excel spreadsheets, uses for, 436–443
Exceptions, seeking of, 270, 278
Experience-sampling method (ESM), 180, 226n, 234
Experimental designs
bias, identification of, 249–251
cause-and-effect relationships, 220, 246–247
control, importance of, 222–227
defined, 220
design strategies, 230–232
dissertation example, 251–254
double-blind experiment, 225
ex post facto designs, 239–240
factorial designs, 240–242, 245
identifying quantitative designs, 245–246
Internet-based studies, 248–249
meta-analysis, 248
overview, 220–221
pre-experimental designs, 228–230
quasi-experimental designs, 233–239
random selection, 225–226
summary table of, 243–244
Experimental group, 225
Experimental research, characteristics of, 116
Experts, advice from, 55–56
on assessment strategies, 130
research proposal, writing of, 164
Explanatory sequential designs, 296–297, 397
Exploratory sequential designs, 295–296
Ex post facto designs, 117, 239–240
bias, identification of, 250–251
factorial design combinations, 241–242, 245
summary table of designs, 243–244
External validity, 119, 250–251
Extreme case sampling, 272

F
Face-to-face interviews, 181–182. *See also* Interviews
Face validity, 128
Factor analysis, 368
Factorial designs, 240–242, 244, 245
Fair use, 101
Faust, K., 123
Feasibility, evaluating plans, 142–143
Federal Digital System (FDsys), 89
Ferguson, D. L., 245, 340
Ferrari, M., 358
Ferron, J. M., 341n
Fidler, F, 369n
Figures, use of, 38, 413, 417
Fine, M., 321, 329
Fisher's exact test, 368
Fiske, D. W., 130
Five-number summary, 354–355, 356
Flanagan, D. P., 358
Fleming, A., 48, 234
Florey, H. W., 48
Focus groups, 274–275, 279, 280–281
Folkman, S., 136
Folk terms, in qualitative data analysis, 391

Follow-up studies, value of, 118
Footnotes, 40, 418
Forbes, M. L., 389
Formal research, defined, 24–25
Fossey, D., 28
Freeman, L. C., 123
Freeman, M., 31
Free-recall tasks, 221
Freeware
  bibliographic software, 94
  mind-mapping (brainstorming) software, 64–65
  project management software, 144, 146
  spreadsheet software, 342
  statistical software packages, 370
Freire, P., 319
Front matter, 417–418
Furnham, A., 277

G
Gall, J. P., 188, 281
Gall, M. D., 188, 281
Gallo, J. J., 165
Gallop, R., 261
Garcia, J. A., 180
Garson, G. D., 389
Gatekeeper, 262
Gay, L. R., 207
Geertz, C., 263
Gender differences, questionnaire construction, 192
Generalizability, 119, 250
Generalizations, pitfalls in thinking, 47
Geometric mean, 351–352
Gergen, K. J., 31
Gestalt, 36
Geva, D., 246
Gill, H., 322
Glaser, B. G., 264, 281, 387, 389
Goh, W., 178n
Going native, 262n
Gomez-Scott, J., 193
Good, R., 401
Goodall, J., 28
Google Books, 89, 96
Google Scholar, 53, 89
Gorard, S., 29
Gosling, S. D., 198, 205
Gosset, W. S., 368
Gould, S. J., 353
Graham, W. F., 292
Grammar, 39, 41, 162–163
Graphics
  electronic spreadsheets, 342
  figures and tables, use of, 38
  line graph, 340
  simple linear sequence, 339
  tables and figures, use of, 413
  two-dimensional table, 339–340
  word-processing software, 40
Greene, J. C., 292
Grounded theory study, 263–264, 267
  characteristics of, 117
  data analysis example, 387–389
  sampling decisions, 272
Group dynamics, 279
Guba, E. G., 269
Gynnild, A., 389

H
Hacker, K., 322, 323, 328, 331
Hague, E., 322
Hallström, I., 266
Halpern, D. F., 42
Hämäläinen, R., 319
Hanson, W. E., 296
Harden, A., 399
Harkness, S., 397–398
Hawthorne effect, 223
He, M. F., 265
Heck, A., 236, 237
Heine, S. J., 192
Hell, B., 358
Hess, C., 102
Hesse-Biber, S. N., 264, 299, 305
Historical records, grounded theory study, 264
Historical research, characteristics of, 117
Hofman, N. G., 322
Holmbeck, G. N., 70
Holmes, C. J., 88
Houser, J. J., 65
Howe, K., 281
Huang, Q., 340
Huberman, A. M., 386
Human mind
  collaboration with others, 46–48
  common pitfalls in thinking, 47
  critical thinking, 41–42
  deductive logic, 42–43
  inductive reasoning, 43–44
  scientific method, 44–45
  theory building, 45–46
Human subjects. *See also* Sampling
  control and experimental groups, 225
  ethnography, 261–263
  grounded theory study, 264
  Hawthorne effect, 223
  institutional review board (IRB), 139–140
  Internet-based studies, 249
  interviews, guidelines for, 188
  matched pairs, 226
  mixed-methods designs, sampling decisions, 302–303
  novelty effect, 223
  privacy rights, 138
  protection from harm, 135–136
  qualitative research, data collection, 269
  qualitative research, sampling decisions, 271–273
  random selection, 225–226
  voluntary and informed participation, 136–138
Hypotheses, 27
  common pitfalls in thinking, 47
  mixed-methods designs, use of, 292, 301–302
  null hypothesis, 67–68
  statement of, 67–68
  statistics for testing hypotheses, 364–369
  writing first section of proposal, 75
Hypothesis testing, 42

I
Idiographic research, 260–261
Implicit patterns, 261
Included term, in qualitative data analysis, 391
Independent variables
  conceptual analysis exercise, 71–73
  defined, 68
  experimental designs, overview, 221, 243–244
  ex post facto designs, 239–240
  factorial designs, 240–242, 245
Inductive reasoning, 43–44, 47, 119
Inferential statistics, 34–35, 344, 361
  meta-analysis, 369–370
  parametric statistics, 349
  population parameters, estimation of, 361–364
  SPSS (Statistical Package for the Social Sciences), 447–449
  summary table of, 368
  testing hypotheses, 364–369
Informants, primary (key), 262
Informed consent, 136–138
  ethnography, 262
  Internet, use in data collection, 198, 249, 280–281
  interviews, guidelines for, 188, 278
  placebo use, 225
  qualitative research, data collection, 269
Institute for Social Research, University of Michigan, 206
Institutional animal care and use committee (IACUC), 140
Institutional review board (IRB), 139–140
  interview guide, use of, 275
  mixed-methods designs, 304–305
  multiphase iterative designs, 299
  participatory action research, 321
  qualitative research, data collection, 269
Instrumentation bias, 211
Insubstantial phenomena, 120n
Intangible phenomena, 120–123, 128, 131
Interlibrary loan, 96
Internal consistency reliability, 131, 191
Internal validity
  control, importance of, 222–227
  credibility and, 117–118
  cross-sectional studies, 179
  defined, 117, 222
  potential threats to, 224

International alphabets and characters, in word-processing software, 40
Internet
conducting experiments on, 248–249
data collection, descriptive studies, 198–199
ethnographic studies on, 261, 261n
interview data collection tools, 280–281
reference list citations, 420
searches on, 90–91
source identification, URLs and DOIs, 91, 420
writing style resources, 411
Interpretation of findings
computer tools for, 33, 34
correlational studies, 177–178
descriptive research data, 214
mixed-methods designs, 304
qualitative research, credibility and transferability, 269–271
research proposal, writing of, 158–160
subproblems, 62
Interquartile range, 354–355, 356
Interrater reliability
defined, 131, 383
mixed-methods designs, dissertation example, 403
observation studies, 175
reports, writing of, 413
strategies for, 385
validity, threats to, 224
Interval data, 185, 346, 349
Interval estimates, 363–364
Interval scales, 124–125
Intervening variables, 69
Interviewer's Manual (1976), 206, 211
Interview guide, 275–277
Interviews. *See also* Mixed-methods research; Questionnaires; Sampling
bias, identification of, 212–214
case study designs, 261
dissertation examples, 215–218, 283–287
grounded theory study, 264
guidelines for productive interviews, 275–277
member checking, 271
phenomenological study, 263
qualitative research, sample decisions, 271–273
qualitative research, strategies for, 274–281
quantitative study, interview guidelines, 187–188
sample coding strategy, 383–385
survey research, 181–182
three-interview series, 263
Iordanou, K., 178n
IQ scores, 357
Irizarry, J. G., 320, 329
Iterative process, research as, 29–30, 259
action research, data collection, 327
design-based research, 318–319
multiphase iterative designs, 298–299

J
Jaccard, J., 35, 45, 264, 264n
Jackson, D. L., 38, 43, 245, 389, 421–422, 444
Jacobson, R. M., 296, 303
Jacoby, J., 35, 45, 264, 264n
Jernigan, T. L., 88
John, O. P., 198
Johnson, B., 239n
Johnson, J. M., 281
Johnson, L. R., 321, 323, 328
Johnson, R. B., 30
Jones, B. D., 186
Jones, S., 317
Josselson, R., 30
Journals, 49
action research, disseminating findings, 328–331
Google Scholar, 53
identifying a research problem, 55
library collections, 84–85
manuscript submissions, 432–433
online databases, 86–89
open-access journals, 88
publishing in, 56
reflective journals, 262
reviewer critiques, responding to, 433–434
teacher research, publication of, 317
JSTOR, 88

K
Kahn, P. G. K., 339
Kalamas, M., 296
Karabenick, S. A., 192
Kellogg, R. T., 37
Kendall coefficient of concordance, 359
Kendall's tau correlation, 359
Key informants, 262
Keywords, 83, 90–91
Kim-Cohen, J., 70
Kime, N., 264
Kinnick, V., 251–254
Koch, R. S., 48
Kontos, S., 175
Kosinski, M., 205
Kozinets, R. V., 261
Krantz, J., 198–199
Kraut, R., 198, 261n, 265n
Krueger, R. A., 280
Kruskal-Wallis test, 368
Kuder-Richardson Formula 42, 191n
Kuhn, D., 178n
Kurtosis, 348
Kushlev, K., 177
Kvale, S., 279

L
Labistre, A. M., 211
Laboratory studies, value of, 118
Language
effective communication guidelines, 37
international alphabets and characters, 40
questionnaires, construction of, 189–190
report writing, clear and coherent writing, 424–426
of research proposal, 153
as research tool, 35–41
Lara, L., 306–310
Latent variables, 120n
Lather, P., 7, 319
Lauer, J. M., 258
Laursen, B., 358
Lawrence-Lightfoot, S., 270
Lawson, L., 317
Leavenworth, P., 56
LeBreton, J. M., 340
Ledgerwood, A., 367
Leech, N. L., 29
Lenz, B. K., 238
Leptokurtic distribution, 348
Leung, C. Y. Y., 73–74, 278, 383–385, 402–406
Levin, J. R., 341n
Lewin, K., 316
Liberman, N., 189
Library
guidelines for efficient use, 93–96
interlibrary loan, 96
library catalog, use of, 84–86
as research tool, 32–33, 53
Library of Congress (LC) classification system, 85–86
Likert scales, 183–186
Limitations, 72–73
data interpretation, dissertation example, 375–376
reports, writing of, 415, 416
statistical analysis of data, 372
writing first section of proposal, 75
Lincoln, Y. S., 269
Lindsay, J. J., 119n
Line graph, 340
Line of regression, 176, 177
Lippa, R. A., 54
Lipsey, M. W., 366n
List servers, 46–47
Literature review
citations, use of, 91
computer tools for, 33, 34
dissertation example of, 102–104
evaluating research of others, 96–97
Excel spreadsheets for tracking, 436–438
fine-tuning research problem, 77–78
ground theory study, 264
guidelines for writing, 98–102
identifying a research problem, 55
identifying sources, style guides and, 91–92
Internet search engines, 90–91
knowing when to quit, 97
library, guidelines for efficient use, 93–96
library catalogs, use of, 84–86
meta-analysis, 369

mixed-methods designs, 302
online databases, 86–89
planning for, 92–93
reference librarians, consulting with, 89
reference lists, use of, 91
report, writing of, 428–429
role of, 82–83
seeking advice and feedback, 102
strategies for, overview, 83
Loftus, A., 191
Loftus, E. F., 245
Logic, deductive, 42–43, 47
Logistics planning, qualitative research, 282–283
Longitudinal mixed-methods designs, 297–298
Longitudinal study, 178–179
Longjohn, M., 321
Lopez, J. E., 321
Lundeberg, M., 192
Lundqvist, A., 266
Luong, A., 194, 197, 210, 212
Lytle, S. L., 316, 317
Lyumbomirsky, S., 177

M

Mackinnon, S. P., 297–298
Malthus, R., 351
Mandler, J. M., 270
Manifest variables, 120n
Mann-Whitney *U*, 368
MANOVA (Multiple analysis of variance), 368
Manuscripts. *See* Journals; Reports; Writing proposals and reports
Marcus, G. E., 263
Marquis, J. G., 238
Martin, V. B., 389
Masland, R., 44
Matched pairs, 226
Maturation, study validity and, 224
Matz, S. C., 205
Maurois, A., 48
Maxwell, J. A., 31
Mayeux, L., 65
McCallin, R. C., 137, 242, 301
McCaslin, M., 211
McCloskey, M., 45
McCrea, S. M., 189
McGee, M. A., 360
McGibbon, E., 261
McGrew, K. S., 358
McGue, M., 100, 433
McInerney, D. M., 211
McLoughlin, W. J., 165
Mean
arithmetic mean, 350–352
geometric mean, 351–352
population mean estimates, 362–363
standard error of the mean, 362–363
statistical mean, 125
Measurement
defined, 120
identifying scales of measurement, 126–127
interval scales, 124–125
nominal scales, 123–124
ordinal scale, 124
ratio scales, 125–126
summary table, scale types, 126
tools for, 33
Measures of central tendency, 349–353, 357
Measures of variability, 350, 357
Median, statistics, 350–352
Mediating variables, 69, 70, 178, 368
conceptual analysis exercise, 71–73
Mehan, H., 261
Member checking, 271
Memos, 268, 270, 387
Mercier, H., 42
Merten, D. E., 261
Message boards, online, 46–47
Meta-analysis, 248, 369–370
Methodological memos, 268
Methodology. *See also* Descriptive research; Planning and design, research
action research, sample report, 333
common methods, summary table, 116–117
defined, 32
mixed-methods designs, dissertation example, 307
vs. planning, 107–108
qualitative design decisions, 259–260
quantitative and qualitative approaches, 1112–117
research proposal, writing of, 157–158, 165
Methods, writing reports, 412
Microform (microfilm, microfiche), 84–85
Microgenetic studies, 178n
Microsoft Excel, use of, 436–443
Middleton, M., 190, 438
Milch-Reich, S., 246
Miles, M. B., 386, 401
Mills, G. E., 207, 316
Miltenberger, R. G., 237, 238
Mind, human
collaboration with others, 46–48
common pitfalls in thinking, 47
critical thinking, 41–42
deductive logic, 42–43
inductive reasoning, 43–44
scientific method, 44–45
theory building, 45–46
Mind-mapping software, 64–65
Miranda, D., 294
Mishler, E. G., 265
Mitchell, K. J., 373–376
Mittapalli, K., 31
Mixed-methods research
applications of, 258, 292–293
characteristics of, 291–292
in content analysis designs, 266
convergent designs, 293–295
data analysis, software tools, 398–399
data analysis and interpretation, 395–398
decision to use, checklist for, 305–306
defined, 28–29
dissertation examples, 306–310, 402–406
ethical considerations, 304–305
evaluation checklist, 400–402
explanatory sequential designs, 296–297
exploratory sequential designs, 295–296
identification of, 299–301
literature review, 302
longitudinal mixed-methods designs, 297–298
multiphase iterative designs, 298–299
notation of qual or quant portions, 293
planning and design, 114–117, 301–304
sample decisions, 302–303
MLA (Modern Language Association) style, 91–92
Mode, statistical, 124, 349–352
Moderating variables, 70, 178, 242, 297, 368
conceptual analysis exercise, 71–73
Modern Language Association (MLA) style, 91–92
Moffitt, T. E., 70
Mohan, L., 192
Morales, M. A., 321
Morelli, G. A., 277
Morrell, E., 320
Mueller, P. A., 37
Multiphase iterative designs, 298–299
Multiple analysis of variance (MANOVA), 368
Multiple-baseline design, 235–237, 238, 244
Multiple-case intervention research, 238
Multiple-case study, 261, 283–287
Multiple correlation, 359
Multiple linear regression, 368
Multitrait-multimethod approach, 130
Muncey, T., 262
Munter, M., 431
Murphy, K. R., 366n
Murphy, M. C., 180
Myors, B., 366n

N

Narrative inquiry, 117, 264–265, 267
NCE scores, 357n
Negative correlation, 358
Negatively skewed distribution, 348
Nelson, S. K., 177
Neuman, W. L., 55, 197, 275
Nichol, A. A. M., 431
Nicholls, M. E. R., 191
Nieto, S., 317
Nirbito, J. G., 319
Nominal data, 346
Nominal scales, 123–124
Non-normal distributions, 346–348
Nonparametric statistics, 348–349, 367
correlational statistics, summary table, 359
inferential statistics, summary table, 368

Non-peer-reviewed research reports, 49
Nonprobability sampling, 205–206
Nonrandomized control-group pretest-posttest design, 233, 243
Normal distribution (normal curve), 124n, 346–348, 362–363
Norm group, 356–357
Norm-referenced scores, 355–357
Novelty effect, 223
Null hypothesis, 67–68, 364
  statistics for testing hypotheses, 364–369
Nurmi, J.-E., 358
Nussbaum, E. M., 42
Nygreen, K., 321, 329

O
Oakes, L. M., 270
Oakley Browne, M. A., 360
Objectivity
  common pitfalls in thinking, 47
  report writing and, 425
Observations
  case study designs, 261
  data collection, computer use in, 186–187
  grounded theory study, 264
  qualitative research, sample decisions, 271–273
  recording of, 273–274
Observation study, 117, 175–176
O'Cathain, A., 303
Odds ratio, 368
Ogive curve, 351, 352
Oksanen, K., 319
Okubo, M., 191
Ollerenshaw, J., 265
One-group pretest-posttest design, 229, 243
One-shot experimental case study, 228–229
Online databases, 86–89
Online discussion forums, 46–47
Online Writing Lab (OWL), Purdue University, 411
Onwuegbuzie, A. J., 29, 31, 304, 396, 396n
Open-access journals, 88
Open coding, 382, 388
Operational definition, 72
Oppenheimer, D. M., 37
Ordinal data, 185, 346, 349
Ordinal scale, 124
Ordinate, 176
Organization, research reports, 421–423
Organization of qualitative data, 381–383, 391–392
Orienting experiences, 323
Ormrod, J. E., 35, 135n, 37, 45, 47, 186, 190, 202, 213, 242, 295n, 301, 389, 438
Ormrod, R. K., 118, 242, 246, 301
Orr, C. A., 191
Outliers, as a strategy for enhancing credibility, 118, 278

P
Paglieri, F., 42
Pajares, F., 69
Palmer, J. C., 245
Paradi, D., 431
Paradigm, as philosophical orientation, 30n
Parallel structure, 163–164
Parameters, 344–345
Parametric statistics, 348–349, 367
  correlational statistics, summary table, 359
  inferential statistics, summary table, 368
Pardo-Ballester, C., 319
Parrigon, S., 340
Partial correlation, 227, 359
Participant observation, 262
Participants, use of term, 231, 231n
Participatory action research (PAR)
  community-based participatory research (CBPR), 323
  community-based qualitative research (CBQR), 323–324
  data collection and analysis, 327–328
  defined, 315
  design selection decisions, 324–326
  disseminating findings, 328–331
  general methodology, 319–321
  research report example, 331–335
  youth participatory action research (YPAR), 319–321
Pashler, H., 248
Path analysis, 368
pdf (portable document format), 53
Pearson product-moment correlations *(r)*, 125, 358–359, 368
Peer review, conference presentations, 430–432
Peer-reviewed research reports, 49, 85, 90
Pelham, W. E., Jr., 246
Percentages, use in data analyses, 124, 347–348
Percentile ranks (percentiles), 124, 348, 354n
Periodicals, 84–85
Permissions, fair use and, 101
Personal space, 277
Peter, E., 261
Peterson, L., 236
Petticrew, M., 399
Pexman, P. M., 431
p-hacking, 367
Phenomenological study, 117, 180, 263, 266
Phenomenology, as philosophical orientation, 31, 263
Phi coefficient, 359
Phillion, J., 265
Philosophical orientation, 30–31
Pilot study, 143
  data coding scheme, testing of, 382–383
  dissertation example, 252
  interviews, guidelines for, 188
  observation studies, 175–176
  questionnaires, construction of, 191–192
Placebo, 225
Plagiarism
  defined, 100
  literature review, writing of, 100–101
  literature search, tips for, 95
  research proposal, writing of, 160
Planning and design, research, 27–28
  assessment reliability, 127, 130–135
  assessment strategies, 120–126
  assessment validity, 127, 128–130, 133–135
  computer tools for, 33, 34
  data, nature and role of, 108–11
  descriptive studies, data collection, 182–199
  enhancing credibility of findings, 117–118
  enhancing generalizability of findings, 119
  ethics considerations, 135–141
  feasibility, judging of, 142–143
  general approach to, 107–108
  institutional review board (IRB), 139–140
  interviews, guidelines for, 187–188
  pilot studies, 143
  planning *vs.* methodology, 107–108
  plan of attack, tips and tools for, 144–146
  quantitative and qualitative methodologies, 112–117
  questionnaires, construction of, 188–193
  research design, defined, 106
  unit of analysis, 107
Plano Clark, V. L., 291, 303
Platykurtic distribution, 348
Point biserial correlation, 359
Point estimates, 363–364
Polkinghorne, D. E., 263
Polymodal distribution, 352
Pompea, S. M., 339
Popov, V., 205
Population. *See also* Sampling
  estimates of population parameters, 344–345
  statistics, estimating population parameters, 361–364
Portable document format (pdf), 53
Positionality, 268
Positive correlation, 358
Positively skewed distribution, 348
Positivism, as philosophical orientation, 30
Poster sessions, 430–432
Postpositivism, as philosophical orientation, 30
Posttests
  control-group posttest-only design, 231
  control-group pretest-posttest design, 230
  experimental design summary table, 243
  nonrandomized control-group pretest-posttest design, 233
  one-group pretest-posttest design, 229
Pousette, A., 264
Power, statistical tests, 366–369
Power hierarchies, 272–273, 324
PowerPoint presentations, 431–432
Practical significance, 371–372

Pragmatism, as philosophical orientation, 31
Pratt, M. W., 297–298
Praxis, as philosophical orientation, 31, 319, 322
Predictions, 352–353
Predictive validity, 129
Pre-experimental designs, 228–230, 243
Preliminary pages, 417–418
Presentation software, 431–432
Pretests, 226
  control-group pretest-posttest design, 230
  experimental design summary table, 243
  nonrandomized control-group pretest-posttest design, 233
  one-group pretest-posttest design, 229
Primary data, 109–110
Primary informants, 262, 272
Privacy rights, 138, 169, 269, 416
Probabilistic reasoning, 42
Probability
  inferential statistical tests, summary table, 368
  statistically significant, defined, 364
  statistics for testing hypotheses, 364–369
  Type I and II errors, 365–369
Probability sampling, 200–205, 272, 302
Problems, research, 25–26
  action research, teacher research, 316–317
  concepts and variables, identification of, 68–73
  conceptual analysis exercise, 71–73
  content analysis designs, 265–266
  data interpretation and, 386–387
  delimitations, 72–74
  discussing with others, 77–78
  finding research projects, 52–59
  fine-tuning, guidelines for, 77–78
  hypothesis, stating of, 67–68
  importance of study, 73, 74, 75
  limitations, 72–73
  mixed-methods designs, 301–302
  purpose statement, writing of, 56–61
  qualitative design decisions, 259–260
  report, writing of, 412, 428
  research proposal, writing of, 157, 165
  subproblems, 61–65
  theoretical or conceptual framework for, 66–67
  writing first section of proposal, 73–77
Process, quantitative and qualitative approaches, 113–114
Professional journals, 49
  action research, disseminating findings, 328–331
  identifying a research problem, 55
  library collections, 84–85
  manuscript submissions, 432–433
  online databases, 86–89
  open-access journals, 88
  publishing in, 56
  reviewer critiques, responding to, 433–434
  teacher research, publication of, 317
Professional organizations
  codes of ethics, 140
  conference presentations, 56, 430–432
Program resources and materials, 329
Project management software, 144–146
Proof, data interpretation and, 414–415
Proofreading, 41, 162–163
Proportional stratified sampling, 202–203, 205
Proposal, research
  characteristics of, 152–153
  common weaknesses in, 165
  as contract, 166–167
  evaluation checklist, 166
  first draft, revision of, 160–166
  first draft, writing of, 156–160
  first section, writing of, 73–77
  formatting, headings and subheadings, 154–155
  importance of, 151–152
  organization of, 153–154
  purpose statement, 152–153
  proposal example, 167–170
ProQuest Dissertations & Theses Global, 53, 96, 411
ProQuest Historical Newspapers, 88
Protection from harm, 198
Protective factors, as moderating variables, 70
Pseudo-subproblems, 62–63
PsychINFO, 87–88
Psychological Research on the Net, 198–199
Public forums, 329–331
PubMed, 89
Punctuation, 39, 162–163
Purdue University Online Writing Lab (OWL), 411
Purpose statement, 37
  conceptual analysis exercise, 71–73
  quantitative and qualitative approaches, 113
  research proposal, writing of, 157, 165
  writing first section of proposal, 73–77
  writing guidelines, 56–61
Purposive sampling, 206, 272, 302–303

Q

Qualitative research. *See also* Descriptive research; Mixed-methods research; Planning and design, research
  advantages of, 260
  assessment reliability calculations, 131–132
  case study, 229, 260–261
  characteristics of, 258–259
  content analysis, 265–266, 389–390
  credible and trustworthy data collection, 269–271
  data analysis, planning checklist, 393–395
  data analysis strategies, 381–387
  data collection, overview of, 268–269
  defined, 28, 258
  design selection decisions, 266–268
  dissertation example, 283–287
  emerging design, 269
  enhancing credibility of findings, 117–118
  enhancing generalizability of findings, 119
  ethnography, 261–263, 390–391
  evaluation checklist, 400–402
  evaluation criteria, 281
  grounded theory study, 263–264, 387–389
  interviews, strategies for, 274–281
  iterative nature of data collection and analysis, 259, 380–381
  logistics planning, 282–283
  memos, use of, 268
  mixed-methods design, dissertation example, 402–406
  narrative inquiry, 264–265
  observations, tips for, 273–274
  phenomenological study, 263
  planning and design, 112–117
  positionality, 268
  researcher-as-instrument, 392–393
  research problems and methodology, 259–260
  sample selection decisions, 271–273
  software tools for data organization and analysis, 391–392
  summary table, 266–267
  systematic reviews, 399–400
Quantitative data. *See* Statistics
Quantitative research. *See also* Descriptive research; Mixed-methods research; Planning and design, research
  assessment reliability calculations, 131–132
  defined, 28
  enhancing credibility of findings, 117–118
  enhancing generalizability of findings, 119
  identifying design type, 245–246
  interviews, guidelines for, 187–188
  planning and design, 112–117
  variables, identification of, 68–70
Quartile, 354–355, 356
Quasi-experimental designs, 233, 243–244, 250–251
Quasi-experimental research, 117
Questionnaires, 182. *See also* Interviews; Mixed-methods research; Sampling
  administration and data analysis, 193–194
  bias, identification of, 212–214
  construction of, 188–193
  dissertation example, 215–218
  Internet, use in data collection, 198–199
  interviews, guidelines for, 187–188
  return rate, maximizing, 194–198
Questions, survey research, 181–182
Quota sampling, 205–206
Quotations, use of, 101

R
Rakison, D. H., 270
Ramirez, L., 215–218
Random numbers table, 200–201
Random sample, estimating population parameters, 361–364
Random selection, 200–205, 225–226
Range, as a measure of variability, 354–355, 356
Rank-order correlation, 124
Rank-order data, 124
Rating scales, 125
  descriptive studies, planning for, 183–186
  interviews, guidelines for, 187–188
  ordinal and interval data, 185
  validity of, 128
Ratio data, 346
Ratio scales, 125–126
Ravert, R. D., 193
Raw scores, 348, 355–357
Reactivity, 118
Realism, as philosophical orientation, 31
Real-world contexts, as means of enhancing credibility and generalizability, 118, 119
Reasoning skills, 41–42, 47
Reference librarians, 89
Reference lists, 418–421
  literature review, writing of, 100–101
  research proposal, writing of, 160
  review of, 164
  source identification, style guides, 91–92
  use in literature search, 91
Reflective journals, 262, 270
Reflective memos, 268, 270
Reflexivity, 270
Regression, statistical, 224, 368
Reiser, B. J., 319
Reliability, 130–133
  of assessment strategies, 127
  correlation coefficients and, 360
  defined, 131
  equivalent-forms reliability, 131
  identifying issues with, 133–135
  internal consistency reliability, 131
  interrater reliability, 131, 224, 383, 403, 413
  qualitative research, credibility and transferability, 269–271
  questionnaires, construction of, 191
  statistical error reduction, 367
  test-retest reliability, 131
Reliability coefficients, 191n
Repeated-measures design, 231n
Repeated-measures variables, 180n, 226–227
Replication studies, value of, 119, 248
Reports. *See also* Writing proposals and reports
  action research, disseminating findings, 328–331
  action research, report example, 331–335
  authorship, sharing of, 433
  bias, acknowledgment of, 393
  computer tools for, 33, 34, 39–41
  conference presentations, 430–432
  critiquing checklist, 427–429
  journal articles, 432–433
  meta-analysis of, 248, 369–370
  peer- *vs.* non-peer reviewed, 49
  preliminary pages, 417–418
  quantitative and qualitative approaches, 114
  reviewer critiques, responding to, 433–434
  revision, need for, 427
  systematic reviews of, 399–400
  writing guidelines, 37–41
Representative samples, 119, 361–364
Research
  applied research, 52
  assumptions, 26–27
  basic research, 52
  characteristics of, 24–30, 24n
  example paper, penicillin discovery, 48
  misconceptions about, 23–24
  philosophical assumptions, 30–31
  uses of term, 23, 24n
  value of, 48–49
Research cycle, 25–30
Researcher-as-instrument, 392–393
Researcher bias, 212. *See also* Bias
Research hypotheses, 27
Research methodology. *See also* Descriptive research; Planning and design, research
  action research, report example, 333
  common methods, summary table, 116–117
  defined, 32
  mixed-methods designs, dissertation example, 307
  vs. planning, 107–108
  qualitative design decisions, 259–260
  quantitative and qualitative approaches, 112–117
  research proposal, writing of, 157–158, 165
Research problem/question. *See* Problems, research
Research reports. *See* Reports
Research tools. *See also* Tools
  computer technology, 33, 34
  defined, 32
  discipline-specific tools, identifying, 49–50
  electronic discussion groups, 46–47
  human mind, 41–48
  language and writing, 35–41
  libraries, 32–33
  measurement, 33
  statistics, 33–35
  word-processing software, 39–41
Respondent validation, 118
Response bias, 211–212
Response rates, 181–182, 210–211
Restorying, 265
Results, report writing, 412–414
Return rate, 182, 194–198
Reversal time-series design, 234–235, 244
Rigorous subjectivity, 392–393
Rios-Moore, I., 320
Risk factors, as moderating variables, 70
Roberts, H., 399
Robustness, in statistical analyses, 349
Rodriguez, J. C., 319
Rodriguez, L. F., 329
Roediger, H. R., 231n
Rogelberg, S. G., 194, 197, 210, 212
Rohrer, D., 248
Rosales, Ruiz, J., 245, 340
Rosario-Ramos, E. M., 323
Rosing, H., 322
Rothbaum, F., 277
Rubrics, 184
Russo, M. J., 130
Rutledge, E., 323

S
Saldaña, J., 383, 386
Sales, B. D., 136
Sampling
  convenience sampling, 272, 272n
  defined, 199
  descriptive research decisions, 199–207
  determining sufficient sample size, 207–209
  discriminant sampling, 272
  enhancing generalizability of findings, 119
  grounded theory study, 264
  matched pairs, 226
  mixed-methods designs, 302–303
  nonprobability sampling, 205–206
  phenomenological study, 263
  probability sampling, 200–205
  purposive sampling, 206, 272
  qualitative research decisions, 271–273
  random selection, 225–226
  report, writing of, 429
  sampling bias, 209–210
  size of, statistical errors and, 366
  snowball (chain) sampling, 272
  statistics, estimating population parameters, 361–364
  survey research, 181–182
  theoretical sampling, 272
  very large populations, 206–207
Sandoval, W. A., 318
Saunders, M. G., 44
Scales, measurement
  defined, 123
  descriptive studies, planning for, 183–186
  identifying scales of measurement, 126–127
  interval scales, 124–125
  Likert scales, 183–186
  nominal scales, 123–124
  ordinal scale, 124
  ratio scales, 125–126
  summary table, scale types, 126
Scatter plot (scattergram), 176–177
Schallert, D. L., 264

Schram, T. H., 264, 271–272
Schumaker, J. B., 238
Schuman, H., 277
Schunk, D. H., 69
Schwab, R. S., 44
Schwarz, N., 211, 212
Scientific method, 44–45
Scott, K. M., 360
Scott-Jones, D., 136
Search engines, Internet, 90–91
Secondary data, 109–110
Seidman, I., 263
SEM (structural equation modeling), 178, 227, 368
Semantic relationship, in qualitative data analysis, 391
Semistructured interviews, 181–182
Shadish, W. R., 341n
Shaenfield, D., 178n
Shaklee, J. M., 183–184, 185
Shanahan, T., 37
Shank, G. D., 275, 277, 278
Shavinina, L. V., 358
Sheehan, K. B., 194
Shenton, A. K., 269
Sherman, S. J., 189
Shernoff, D. J., 180
Sherry, S. B., 297–298
Sigma (Σ), 350
Significance, practical, 371–372
Significance level, 364–369
Significance of study, 75, 157
Silverman, D., 44, 274, 275
Simple linear regression, 368
Simple linear sequence, 339
Simple random sampling, 202, 205
Simple time-series design, 233–234, 244
Single-case design, 237n
Single-case experiment, 237n
Single-case intervention research, 237–239
Skagert, K., 264
Skewed distributions, 348
Skype interviews, 182, 280–281
Smith, M. M., 297–298
Smith, R., 283–287
Snowball sampling, 272
Social action, as philosophical orientation, 31
Social desirability effect, 128, 212
Social media, convenience sampling, 205
Sociogram, 121–122
Sociometric matrix, 122–123
Software. *See* Computer technology
Sokacich, D., 323
Solomon, H., 66
Solomon, R. I., 230–231
Solomon four-group design, 230–231, 241, 243
Sowell, E. R., 88
Sparks, C. W., 367
Sparks, E. A., 367
Sparks, G. G., 367
Sparks, J. C., 367
Spearman rank order correlation (rho), 359
Specifications, table of, 129–130
Spelling, 39, 41, 162–163
Spradley, J. P., 391
Spreadsheets, electronic, 342–343
SPSS (Statistical Package for the Social Sciences), 444–449
Squire, K., 318
Srivastava, A. K., 123
Srivastava, S., 198
Standard deviation, 125, 355
  normal distribution curve, 347–348
  population mean estimates, 362–363
Standard error of the mean, 362–363
Standardization, 132
Standard scores, 124n, 356–357
Stanine scores, 357
Stanley, J. C., 223, 224, 227n
Static group comparison, 229–230, 243
Statistical Package for the Social Sciences (SPSS), 444–449
Statistical regression, 224
Statistical significance, 364–369
Statistics
  average deviation, 354–355, 356
  Chi-square test, 124
  confidence interval, 364
  continuous variables, 345
  control for confounding variables, 227
  correlation coefficient, 54
  data, dynamics in, 339
  data interpretation, 371–373
  data interpretation, dissertation example, 373–376
  data sets, organizing of, 337–343
  defined, 337, 344
  descriptive statistics, 33–34, 344, 349, 359
  discrete variables, 345
  effect size, 369
  electronic spreadsheets, 342–343, 442–443
  as estimates of population parameters, 344–345
  five-number summary, 354–355, 356
  functions of, 344–345
  geometric mean, 351–352
  inferential statistics, 34–35, 344, 361, 368
  interquartile range, 354–355, 356
  interval data, 124–125, 346
  interval estimates, 363–364
  line of regression, 176, 177
  measures of association: correlation, 358–361
  measures of central tendency, 349–353, 357
  measures of variability, 353–357
  meta-analysis, 248, 369–370
  mode, 124
  nominal data, 124, 346
  normal and non-normal distributions, 346–348
  norm-referenced scores, 355–357
  ordinal data, 124, 346
  parametric and nonparametric statistics, 348–349
  percentile rank, 124
  point estimates, 363–364
  population parameters, estimating of, 361–364
  range, 354–355, 356
  rank-order correlation, 124
  ratio data, 346
  reports, writing of, 413–414
  as research tool, 33–35
  scatter plot (scattergram), 176–177
  significance level (alpha (α)), 364–369
  single- and multi-group data, 345
  software packages for, 370
  SPSS (Statistical Package for the Social Sciences), 444–449
  standard deviation, 355
  standard error of the mean, 362–363
  standard score, 356–357
  stanines, 357
  structural equation modeling (SEM), 178
  table of random numbers, 200–201
  testing hypotheses, 364–369
  Type I and II errors, 365–369
  variability, measures of, 357
  variance, 355–356
  z-scores, 356–357
Stenhouse, L., 316
Stern, J. E., 433
Stevens, S. S., 123
Stillwell, D., 205
Story line, in qualitative data coding, 389
Stratified random sampling, 202, 205
Strauss, A., 264
Strauss, A. C., 264, 264n, 387, 389
Stricker, J. M., 237, 238
Structural equation modeling (SEM), 178, 227, 368
Structured interviews, 181–182
Strunk, W., Jr., 434
Student (W. S. Gosset), 368
Student's *t*-test, 368
Style guides, 91–92
  bibliographic software, 94
  reference lists, 419–421
  research proposal, writing of, 156–160
  for research report writing, 410–411
Subcodes, in qualitative data analysis, 381–386, 393
  grounded theory study example, 387–389
  mixed-methods designs, 396–398
  mixed-methods designs, dissertation example, 403–406
Subjectivity, rigorous, 392–393
Subjects, use of term, 231, 231n

Subproblems. *See also* Problems, research
characteristics of, 62–63
defined, 61
identification of, 63–65
report, writing of, 428
writing first section of proposal, 74–75
Substantial phenomena, 120
Summary, writing reports, 416, 425
Sunaryanto, 319
Survey research, 181–182. *See also* Sampling
bias, identification of, 212–214
characteristics of, 117
dissertation example, 215–218
Internet, use in data collection, 198–199
interval data, 125
interview guidelines, 187–188
questionnaires, administration and data analysis, 193–194
questionnaires, construction of, 188–193
questionnaires, maximizing return rate, 194–198
Survey Research Center, University of Michigan, 206, 211
Surveys. *See* Interviews
Sylvestre, J., 294
Symposia, 329
Systematic reviews, 399–400
Systematic sampling, 203–204, 205

T

Table of contents, 417, 418
Table of random numbers, 200–201
Table of specifications, 129–130
Tables, use of, 38, 40, 413, 417
Tashakkori, A., 292n
Tay, L., 340
Taylor, A., 70
Taylor, S. L., 389
Teacher research, 316–318
Teddlie, C., 292n, 396, 396n
Telephone interviews, 181–182
Terms, consistency of use, 161–162
Terms, defining of, 71–72, 412
Testing, validity and, 224
Test-retest reliability, 131
Thakor, M. V., 296
Theoretical framework, 66–67
report, writing of, 428
writing first section of proposal, 75
Theoretical sampling, 272
Theory
defined, 45
grounded theory study, 263–264
theory building, 45–46
Thick description, 118, 413
Thomas, J., 399
Thompson, B., 360
Thompson, P. M., 88
Thrailkill, N., 100, 221, 226–227, 232
Threatts, T., 320
Three-interview series, 263
Time-series designs
alternating-treatments design, 235
control-group time-series design, 234
multiple-baseline design, 235–237
reversal time-series design, 234–235
simple time-series design, 233–234
single-case intervention research, 237–239
summary table, 244
Toga, A. W., 88
Tolentino, E. P., 317
Tools. *See also* Research tools
data organization and analysis, 391–392
electronic discussion groups, 46–47
electronic spreadsheets, 342–343
Excel spreadsheets, uses for, 436–443
for finding research projects, 52–54
Internet search engines, 90–91
interview data collection, 280–281
library, efficient use of, 93–96
library catalogs, use of, 84–86
mind-mapping (brainstorming) software, 64–65
mixed-methods data analysis tools, 398–399
online databases, 86–89
presentation software, 431–432
project management software, 144–146
questionnaires, administration and data analysis, 193–194
SPSS (Statistical Package for the Social Sciences), 444–449
statistical software packages, 370
tables and figures, creation of, 413
transcription software, 280–281
video-conferencing software, 182, 280–281
voice-recognition software, 188, 280–281
word-processing software, 39–41
Tracking changes, word-processing software, 40
Trahan, R. G., 118, 183, 184, 246
Transcription software, 280–281
Transferability, data collection guidelines, 269–271
Treatment, defined, 220
Treatment group, 225
Triangulation, 117–118, 129n, 130, 270, 293–295
Triserial correlation, 359
Trope, Y., 189
Trouche, E., 42
True experimental designs, summary table, 243
t-test for a correlation coefficient, 368
Two-factor designs, 240–241, 244
Tyler, K. M., 277
Type I error, 365–369
Type II error, 365–369

U

Ubuntu, 36–37
Uniform Resource Locator (URL), 91, 420
Unit of analysis, 107
Unobtrusive measures, 118
URL (Uniform Resource Locator), 91, 420
U.S. Government Printing Office, 89
U-shaped relationship, 360
U.S. National Library of Medicine, 89
Uzeil, L., 212

V

Validation, respondent, 118
Validity
of assessment strategy, 127, 128–130, 133
bias and, 249–250
confounding variables, controlling for, 223–227
construct validity, 129
content validity, 128–129
control, importance of, 222–227
correlation coefficient and, 360
criterion validity, 129
defined, 128
ecological validity, 119n
external validity, 119
face validity, 128
identifying issues with, 133–135
internal validity, 117–118, 179, 222
potential threats to, 224
predictive validity, 129
qualitative research, credibility and transferability, 269–271
questionnaires, construction of, 192
sample and, 199
Solomon four-group design, 231
statistical error reduction, 367
Vanderwood, M., 358
Variability, measures of, 350, 353–357
Variability, sampling decisions, 273
Variables
conceptual analysis exercise, 71–73
confounding variables, 223–227, 246–247
continuous variables, 345
control, importance of, 222–227
correlational research, 176–178
defined, 68
dependent and independent, 68, 221
discrete variables, 345
experimental design, overview, 220–221, 243–244
ex post facto designs, 239–240
factorial designs, 240–242, 245
identification of, 68–70
inferential statistical tests, summary table, 368
latent variables, 120n
manifest variables, 120n
mediating variables, 69, 178
moderating variables, 70, 178, 242

pre-experimental designs, 228–230
quantitative and qualitative methodologies, 112–117
repeated-measures variables, 180n, 226–227
true experimental designs, 230–232
within-subjects variable, 226–227
writing first section of proposal, 75
Variance, 355–356
Vazire, S., 198
Vega, R. I., 211
Verbal reasoning, 42
Verb tenses, report writing, 425
Video-conferencing software, 182, 280–281
Voice-recognition software, 188, 280–281
Volunteer subjects. *See* Human subjects; Sampling
von Stumm, S., 358
Vul, E., 248

W
Wagner, E. D., 242, 301
Wainer, H., 413
Waivers, 139
Walton, D. N., 45
Ward, C., 277
Wasserman, S., 123
Web of Science, 88–89
Wells, J. E., 360
Wennick, A., 266
Whajoedi, 319
Wheelan, C., 212
White, E. B., 434
Wikipedia, 90
Wilcoxon signed-rank test, 368
Winchester, H. P. M., 296
Within-subjects design, 231–232, 243
Within-subjects variable, 226–227
Witt, E., 211
Wittink, M. N., 165
Wixted, J. T., 248
Wolach, A., 366n
Wolcott, H. F., 392, 425
Wong, P. T. P., 165
Word clouds, 398
Word-processing software, 39–41
WorldCat, 89
Worldview, as philosophical orientation, 30n
Writing proposals and reports
academic integrity, maintaining, 416–417
action research, dissertation example, 331–335
appendix content, 421
authorship, sharing of, 433
bibliographic database software, 94
clarity and coherence, guidelines for, 424–426
critiquing final reports, checklist for, 427–429
data interpretation, 414–415
data interpretation, dissertation example, 373–376
data presentation and analysis, 412–414
descriptive research, dissertation example, 215–218
drafts, 39
effective communication guidelines, 37–41
end matter, 417, 418–421
essential elements, 411–412
experimental study, dissertation example, 251–254
footnotes and endnotes, 418
front matter, 417–418
guidelines for, 37–41
identifying weaknesses of study, 415
literature review, guidelines for, 98–102
literature review, dissertation example, 102–104
mixed-methods study, dissertation example, 306–310, 402–406
organizing a report, 421–423
preliminary pages, 417–418
purpose statement, 56–61
qualitative study, dissertation example, 283–287
questionnaires, construction of, 189–193
reference lists, 418–421
research problem or questions, 412
research proposal, common weaknesses, 165
research proposal, first drafts, 156–160
research proposal, first section, 73–77
research proposal, revision of, 160–166
research proposal, sample document, 167–170
revision, planning for, 427
study design and data collection methods, 412
style decisions, 410–411
style guides, 91–92
summary and connections, 416
word-processing software, 39–41
writing schedule, 426–427

Y
Youth participatory action research (YPAR), 319–321
YuliEfiyanti, A., 319

Z
Zambo, D., 39, 275–277, 422–423
Zeith, T. Z., 358
Zhou, H., 296
Zirkel, S., 180
z-score, 356–357