The Scholar-Practitioner's Guide to Research Design

The Scholar-Practitioner's Guide to Research Design

EDITORS

Gary J. Burkholder, PhD
Walden University

Kimberley A. Cox, PhD
Walden University

Linda M. Crawford, PhD
Walden University

LAUREATE PUBLISHING

Published by
Laureate Publishing, Inc.
650 S. Exeter Street
Baltimore, MD 21202
www.laureate.net

Director, Laureate Publishing: Anne E. Jacobs
Managing Editor: Jason Jones
Development Editor: Annie Pezalla, PhD
Cover and Interior Design: Stephen Druding
Editorial Services: Absolute Service, Inc.
Production Services: Absolute Service, Inc.

ISBN-13: 978-1-62458-030-7 (Paperback edition)
ISBN-13: 978-1-62458-031-4 (VitalSource edition)
ISBN-13: 978-1-62458-032-1 (Kindle edition)
ISBN-13: 978-1-62458-033-8 (Apple edition)
ISBN-13: 978-1-62458-034-5 (Nook edition)
ISBN-13: 978-1-62458-035-2 (Adobe Digital Edition)

First Edition

14 15 16 17 18 / 10 9 8 7 6 5 4 3 2 1

Contributors

Patricia M. Burbank, DNSc, RN, FAAN
Professor and Associate Dean for Academic Affairs
College of Nursing
University of Rhode Island
Kingston, Rhode Island

Gary J. Burkholder, PhD
Senior Research Scholar and Faculty
Schools of Psychology and Health Sciences
Walden University
Minneapolis, Minnesota
Vice President of Research and Accountability
Laureate Education
Baltimore, Maryland

Kimberley A. Cox, PhD
Faculty
School of Psychology
Walden University
Minneapolis, Minnesota

Linda M. Crawford, PhD
Faculty
School of Education
Walden University
Minneapolis, Minnesota
Director of Quality Assurance
Laureate Education
Baltimore, Maryland

Paula Dawidowicz, PhD
PhD Education Research Coordinator
Walden University
Minneapolis, Minnesota

John H. Hitchcock, PhD
Director and Associate Professor of Instructional Systems Technology
Center for Evaluation and Education Policy
School of Education
Indiana University
Bloomington, Indiana

James Jackson, EdD
Faculty
School of Education
Walden University
Minneapolis, Minnesota

Laura Knight Lynn, PhD
Executive Director
Center for Research Quality
Walden University
Minneapolis, Minnesota

Louis Milanesi, PhD
Executive Director of Student Affairs
Walden University
Minneapolis, Minnesota

Bonnie K. Nastasi, PhD
Professor
Department of Psychology
School of Science and Engineering
Tulane University
New Orleans, Louisiana

Annie Pezalla, PhD
Associate Director of Curriculum and Assessment
Center for Research Quality
Walden University
Minneapolis, Minnesota

Kurt Schoch, EdD
Faculty
School of Education
Walden University
Minneapolis, Minnesota

Mary S. Stewart, PhD
Research Associate
Center for Evaluation and Education Policy
School of Education
Indiana University
Bloomington, Indiana

Reviewers

Aimee Ferraro, PhD, MPH
Faculty
School of Health Sciences
Walden University
Minneapolis, Minnesota

Gerald Giraud, PhD
Vice President for Academic Affairs and Chief Academic Officer
Northwest College
Powell, Wyoming

William C. Schulz III, PhD
Faculty
School of Management
Director of Academic Initiatives
Walden University
Minneapolis, Minnesota

Elizabeth Adams St. Pierre, PhD
Professor
Department of Educational Theory and Practice
University of Georgia
Athens, Georgia

Contents

Foreword

In the agricultural age, success required land. In the industrial age, success required capital. In the information age, success requires knowledge. With the right knowledge, you can acquire land and capital. With the right knowledge, you can become more effective at whatever you choose to do, both personally and professionally. However, information alone is not knowledge. The Internet is filled with information, but it yields very little by way of knowledge. Knowledge comes from synthesizing research findings, observations from practice, insights from experts, interpretation based on theory, and ongoing learning.

The Scholar-Practitioner's Guide to Research Design will advise you on what you need to do to make a contribution to knowledge that is worthy of an advanced degree. This is what it means to be awarded an advanced degree—that you have made an original contribution to knowledge. This may sound daunting and it may evoke anxiety and doubt. One source of anxiety you may experience that is common among graduate students is worry about whether the results will turn out the way you hope they will. Abandon that notion right now. Your task—indeed, your responsibility—is not to produce desired results. Your task—indeed, your responsibility—is to find out the answer to your research question. There is no right or wrong answer. There is nothing to prove. The world is the way it is. Your task—indeed, your responsibility—is to contribute to our understanding of how the world works. In so doing, you will be engaged in inquiry, learning, discovery, and deeper questioning. Bring to your inquiry an open mind. Be prepared to follow the data wherever they take you, to see what your research reveals, and to make sense of what you find. Forget about proving something. Focus on discovery.

Your task—indeed, your responsibility—is to formulate a researchable question, identify methods that will help you pursue an answer to that question, systematically gather the appropriate data, document what you do every step along the way, organize and interpret your data, and report the findings. Part 1 of *The Scholar-Practitioner's Guide to Research Design* will advise you on how to complete each of these inquiry elements. Quality is judged by the relevance and significance of your question, the extent to which your proposed research design will help you answer the question, careful and systematic implementation of your design to gather valid and reliable data, and thoughtful analysis. Nowhere in that process are you obligated to prove a certain result. Your obligation is to engage in serious and thoughtful inquiry, to interpret your findings with attention to both the strengths and weaknesses of your methods and data, and to place your findings in the larger context of knowledge within your field. In so doing, you will acknowledge those who have contributed knowledge before you and you will end by proposing inquiry questions for future generation of knowledge. This book will guide you on how to fulfill each of the steps along the way, so that you make a contribution to knowledge and (in recognition of that contribution) attain your advanced degree, thereby joining the worldwide community of scholar-practitioners. In the book's Introduction, the editors define the unique niche and important role of a scholar-practitioner with the research community.

Developing a Mindset of Inquiry

Although the short-term purpose of your research is to fulfill the requirements for an advanced degree, the larger and longer term purpose is to develop a mindset of inquiry. Effectiveness as a scholar-practitioner involves ongoing inquiry, continuous learning, and application of research skills to your endeavors. Let me, then, preview for you eight things you are going to have the opportunity to learn as a new researcher that will serve you well for the rest of your life. In the spirit of Stephen Covey's best-selling *The 7 Habits of Highly Effective People*, these are *The 8 Habits of Highly Effective Scholar-Practitioners*.

The 8 Habits of Highly Effective Scholar-Practitioners

1. Savvy and discriminating research consumption. Regardless of whether you ever again conduct original research after attaining your advanced degree, you cannot help but be a consumer of research in this knowledge age. The challenge is to become a *sophisticated* consumer of research. We are inundated constantly by research findings. A lot of it is junk, much of it is opinion disguised as research, a lot of it is biased, and much of it is fabricated. However, a lot of it provides new insights and knowledge that will make you more effective in whatever you do. The trick is to know how to tell the difference. The scholar-practitioner mindset is to engage new research findings by asking questions that may require digging deeper: What design produced those findings? Which data were gathered and how were the data analyzed? How do these findings fit with other research in the field? Within what theoretical framework was the inquiry conducted? What are the practice implications of the findings, if any? A scholar-practitioner mindset of inquiry means that you will not accept news headlines about research findings at face value, but rather that you will look behind those headlines, using your research knowledge and skills to determine the significance and relevance of the findings to your own work and life. Chapters 1 and 9 address how understanding the language and processes of research will help you become a better producer, as well as a more astute and critical consumer of research.

2. Asking meaningful and important questions. One of the first things you will do is work on formulating your research question. In everyday life, we are sloppy about how we frame questions. Learning to ask researchable questions with potentially actionable answers—the hallmark of the scholar-practitioner—is a skill that will serve you well in your life. You will listen more astutely to how people frame their inquiries, the questions they ask, the format of those questions, and whether the questions are posed in ways that could possibly be answered. A sophisticated scholar-practitioner is a sophisticated question asker. It all begins with good questions. Chapter 9, Critiquing and Synthesizing the Literature, emphasizes that learning to identify and assess the research questions of other researchers will help you become more skilled and savvy about how to formulate your own inquiry questions. The special topics covered in Part 2 of *The Scholar-Practitioner's Guide to Research Design* illustrate ways in which different approaches to and frameworks for conducting research—grounded theory, phenomenology, survey research, case studies, program evaluation, and action research—are distinct, in part, because they ask different questions and reflect diverse inquiry traditions.

3. Critical thinking skills. Scholar-practitioners are critical thinkers. Presumably, to reach the stage of working on an advanced degree, you already have some critical thinking skills. You may even think you are a rigorous thinker. Get ready to go deeper, not just to design and complete your research, but in all that you do. The stakes can be quite high. Philosopher Hannah Arendt devoted much of her life to trying to understand how the Holocaust could have occurred. Having experienced totalitarianism as a young woman living in Germany prior to World War II, then having fled as conditions worsened, she came to believe that thinking rigorously in public deliberations and acting democratically were intertwined. Totalitarianism is built on and sustained by deceit and thought control. To resist efforts by the powerful to deceive and control thinking, Arendt (1968) believed that people needed to *practice* thinking. Toward that end, she developed eight exercises in political thought. She wrote that "experience in thinking . . . can be won, like all experience in doing something, only through practice, through exercises" (p. 4). From this point of view, I invite you to consider your research inquiry as an opportunity to practice thinking, to do what Arendt hoped her exercises in political thought would do—namely, help us "to gain experience in *how* to think." Her exercises "do not contain prescriptions on what to think or which truths to hold" but rather on the act and process of thinking. For example, Arendt thought it important to help people think conceptually,

> to discover the real origins of original concepts in order to distill from them anew their original spirit which has so sadly evaporated from the very keywords of political language—such as freedom and justice, authority and reason, responsibility and virtue, power and glory—leaving behind empty shells (pp. 14–15).

Your research will involve examining the history of key concepts: how they had been applied in the past and their relevance for today. That is an exercise in critical thinking, as will be all aspects of your inquiry. Indeed, as Chapter 9 emphasizes, understanding all the lessons from the first part of this book—the terminology of research, the philosophical underpinnings of research, the roles of theory, the various designs of research (qualitative, quantitative, and mixed methods), and the considerations of quality and ethics—will help you become a more critical consumer of information and, in turn, a better writer and thinker overall.

4. Advanced interviewing skills. Many of you are likely to use interviewing as one of your inquiry methods. Becoming a skilled interviewer will enhance the validity and trustworthiness of the data you collect through interviewing. However, becoming a skilled interviewer will also make you a more effective practitioner in the knowledge age. Every interaction with other human beings offers opportunities to inquire into their worldviews, perspectives, experiences, opinions, feelings, and behaviors. Knowing how to ask inviting and engaging questions and how to listen attentively, probe astutely, and follow up on what you hear—these are the elements of effective interviewing and they will serve you well not only in conducting research, but also in your relationships with family members, significant others, colleagues, and everyone you encounter on your life journey.

I find interviewing people to be invigorating and stimulating—the opportunity for a short period of time to enter another person's world. I am personally convinced that to be a good interviewer, you must like doing it. This means being interested in what people have to say. You must, yourself, believe that the thoughts and experiences of the people being interviewed are worth knowing. In short, you must have the utmost respect for these persons who are willing to share with you some of their time to help you understand their world. If you include interviewing in your research, this will be an opportunity not only to deepen your skills, but also to reflect on how you approach interactions of all kinds with other people. How good are you at asking questions? How good are you at listening? How skilled are you at making people comfortable enough to tell you their stories? How confident are you as an interviewer in our modern interview society (a core aspect of the knowledge age)? Scholar-practitioners engage in ongoing learning and the development of key inquiry skills, such as in-depth interviewing. Chapter 17 provides new researchers with a detailed overview of the interviewing essentials that will foster success in the collection of data.

5. Astute observational skills. Every interview is also an observation. With our heads buried in our devices these days, we miss a lot of what goes on around us. All of science is based on—indeed, depends on—astute, focused, deep, and discerning observation. The great thinkers have all been great observers: Aristotle, Copernicus, Newton, Darwin, Einstein, Jung, Goodall, and whoever have been the pioneers in your own specialized scholar-practitioner field. Skilled interviewing involves more than just looking. It involves *seeing*. It involves knowing what to look for. It requires distinguishing signal from noise, astute pattern recognition, and an openness to be a witness to whatever is unfolding before you. In our everyday lives, we walk around in a fog, engrossed in our own thoughts, operating on preconceived ideas, immersed in selective perception, operating on biases, and generally being oblivious. To move into an observational mindset means to lift the fog, stop the internal noise and gaze outward into the intriguing panorama of complex reality, open the mind to new possibilities and to seeing things in new ways—in short, to become a skilled scholar-practitioner observer. Skill at observation, as with interviewing skills, will serve you well not only in conducting research, but also in all aspects of your life. Chapter 10, Writing the Research Proposal, offers guidance on how to incorporate observations into research design.

6. Rigorous meaning-making skills. All data have to be interpreted. Data do not speak for themselves. Making sense of data is a quintessential inquiry competence and skill. One aspect of analysis and interpretation, as with data collection itself, whether through interviewing or observation, involves distinguishing signal from noise. What is significant? What are the patterns and themes that are worth your attention? What do they mean? How do those meanings inform your understanding of the world and, through enhanced understanding, your actions in the world and *your* practice as a scholar-practitioner? One of the pitfalls of analysis is letting your preconceptions, selective perceptions, biases, and preferred ways of seeing things dominate your data interpretation. Rigorous analysis requires rigorous critical thinking. This includes triangulation: looking across different kinds of data (interviews, observations, documents) and diverse interpretive frameworks (previous research findings, alternative theories, competing conceptual frameworks) to figure out what interpretation best fits the data—not your hopes for the data, but the actual data. Rigorous analysis requires a disciplined search for data that does not fit the dominant pattern, a search for alternative explanations, and a willingness to play devil's advocate in your own research. As with the other skills I have enumerated, becoming better at making sense of data will serve you well, not only in your research, but also in deciding the value of all kinds of findings and propositions that you will encounter as you are bombarded from all sides with allegations about how the world is. Astute sense-making, skilled interpretation, and rigorous meaning-making are what will move you from understanding to action in the knowledge age. The Introduction and Chapters 1 and 2 encourage you to reflect upon your worldview and your orientation to research as a core part of the meaning-making exercise. A scholar-practitioner is a reflective practitioner.

7. Integrating theory and practice. Kurt Lewin (1890–1947), a pioneer of applied social psychology, famously observed that "there is nothing as practical as a good theory." Knowing how to integrate theory and practice is at the core of being an effective scholar-practitioner. The purpose of theory is to explain how the world functions. The purpose of practice is to make a difference in the world. Clearly, knowing how the world functions is essential for knowing how to make a difference in the world. That is what is meant by *practical theory*. In much of academia, theory and practice are separate; there are courses on theory and other courses on practice. The scholar-practitioner, then, faces the challenge of drawing on theory to inform practice and using lessons from practice to deepen and enrich theory. Here is an example of this concept drawn from brain research.

Brain research reveals that the human brain is where theory and practice are integrated. The evidence from social and behavioral science shows that in order to make decisions when we are faced with complex choices in new situations, we fall back on a set of rules and standard operating procedures that predetermine what we will do and that effectively short-circuit genuine situational adaptability. The evidence illustrates that we are running on preprogrammed tapes most of the time. That has always been the function of social norms, rules of thumb, and scientific paradigms. Faced with a new situation, we turn to old and comfortable patterns (sometimes consciously, sometimes unconsciously). This may help explain why so many researchers find that the approaches in which they are trained and with which they are most comfortable *just happen* to be particularly appropriate in each new inquiry in which they engage—time after time after time.

Several interdisciplinary fields of inquiry have provided insights into how we manage decision making in the face of complexity. Decision sciences have been identifying decision heuristics that cut through the messy, confusing, and overwhelming chaos of the real world, so that we can avoid analysis paralysis and take action. We rely on routine "heuristics"—rules of thumb, standard operating procedures, practiced behaviors, and selective perception that can be distorting the possibilities.

- The *confidence heuristic* ensures that the *amount* of information we obtain increases our confidence in our judgments, regardless of the accuracy or redundancy of the information.
- The *representativeness heuristic* guides our sense-making in new situations by focusing our attention on those aspects of the situation that are most familiar and similar to our previous experiences. We make a new problem or situation *representative* of things we already know, selectively ignoring information and evidence that is unfamiliar or that does not fit our preconceptions developed through past experiences.
- The *availability heuristic* operates to make sense of new situations by bringing readily to mind things that happened to us recently or information that we come across frequently.

What intersecting and overlapping fields of research on decision making and brain functioning reveal is that we cannot systematically consider every possible variable in a situation. Findings from cognitive science, decision science, and contingency theory triangulate to reveal that what makes thought and action possible is viewing the real world through some kind of framework for making sense of situations—a framework that tells us what factors deserve priority in our sense-making. This already and inevitably happens unconsciously. It is how our brains work. The issue is whether and how we can become more consciously and intentionally deliberative about how we engage in sense-making and situation recognition. We can do this through ongoing and in-depth reflective practice to become *reflective theory–based practitioners*. We can do this by systematically evaluating our decision-making patterns, engaging in ongoing learning, and deconstructing our design tendencies and methodological decision-making.

Let me reiterate the overall point here. We are already integrating theory and practice in our brains, because to practice (or act) at all, we make sense of what to do by interpreting the situations we face through some set of theoretical screens and constructs that operate through decision heuristics, cognitive algorithms, paradigm parameters, and mental shortcuts. The issue, then, is not whether we integrate theory and practice in our brains. The issue is whether we become more intentional and deliberative about how we do so. Toward that end, your inquiry journey will provide an opportunity to exercise greater intentionality and deliberativeness in integrating theory and practice in your own decision-making processes, especially including decisions you make about your research question, design, analysis, and reporting. What you learn about your own patterns and tendencies will have influence far beyond the research you conduct for your advanced degree. As noted earlier, but well worth reiterating: A scholar-practitioner is a reflective practitioner. Chapter 2, Philosophical Foundations and the Role of Theory in Research, and Chapter 3, Conceptual and Theoretical Frameworks in Research, are especially helpful for developing an understanding of theory–practice integration, while Chapter 15, Action Research, demonstrates the practical application of the theory–practice connection.

8. Systematic evaluative thinking. The eighth and final thing that you are going to have the opportunity to explore as a new researcher (and that will serve you in good stead for the rest of your life) is systematic evaluative thinking. Chapter 16 includes detailed guidance on conducting program evaluation, but whether or not you undertake a formal program evaluation for your research, you will still, inevitably, be engaged in evaluative thinking. This means that, as discussed in Chapter 7, you will have to become explicit about criteria you will use to judge quality, validity, relevance, credibility, and utility of research findings, and the processes that generated those findings. This applies not only to your own research, which will be evaluated by you, your research committee, and others who read your findings, but also to your review of the research of others. As you engage in a literature review, you will be using systematic evaluative thinking to determine which published research is particularly relevant and meaningful for your own inquiry. As you review methodological possibilities and analysis options, you will be engaged in systematic evaluative thinking. This will require you to get better at generating evaluative criteria and applying those criteria to make judgments about what is significant, meaningful, and important. Pay attention to the criteria you use and the implications of applying those criteria. As with critical thinking and integrating theory and practice, systematic evaluative thinking is a skill you can improve with practice, but only if you pay attention, reflect on your thinking patterns, and learn about what you are studying and, perhaps even more important, about yourself.

On Being Present for the Whole Journey

In closing, I would like to share with you a cautionary tale to inform your inquiry journey:

> Master Halcolm was renowned as a wise sage. Students came to him for advice and counsel. When a group of students about to begin research projects came to see Halcolm, he asked them, "What did you learn in your readings this week?"
>
> "We learned that a journey of 1000 miles begins with the first step," replied the students.
>
> "Ah, yes, the importance of beginnings," smiled Halcolm.
>
> "Yet, I am puzzled," said a student. "Yesterday, I read that there are 1000 beginnings for every ending."
>
> "Ah, yes, the importance of seeing anything through to the end," confirmed Halcolm.
>
> "Which is more important—to begin or to end?" asked a student.
>
> "Two great self-deceptions are asserted by the world's self-congratulators: that the hardest and most important step is the first and that the greatest and most resplendent step is the last. While every journey must have a first and last step, my experience is that what ultimately determines the nature and enduring value of the journey are the steps in between. Each step has its own value and importance. Be present for the whole journey, learners that you are. Be present for the whole journey." (Patton, 2015, pp. 734–735)

Michael Quinn Patton
Saint Paul, Minnesota

REFERENCES

Arendt, H. (1968). *Between past and future: Eight exercises in political thought*. New York, NY: Viking Press.

Patton, M. Q. (2015). *Qualitative research and evaluation methods* (4th ed.). Los Angeles, CA: Sage.

Acknowledgments

We are indebted to many people for their support and intellectual contributions during the creation of *The Scholar-Practitioner's Guide to Research Design*. First, a great thank you to our families and friends for their support and encouragement during the many hours of writing and editing. We also extend our deepest appreciation to the chapter contributors. We sought colleagues across institutions who not only had deep expertise in research, but also significant experience working with scholar-practitioners, for whom this book was written. Our colleagues worked with us to refine the vision for the book, so that it would be maximally useful for those who are just learning about research and who will continue to advanced practice. Many chapter contributors also sought expertise from other colleagues, who provided them with valuable feedback on content and tone. On behalf of the chapter contributors, we thank these colleagues for their insights. Several reviewers also provided feedback that helped us to ensure the book was positioned appropriately for the scholar-practitioner audience. We used these reviewers' insights to ensure that our book was inclusive of multiple perspectives—perspectives that are indeed important for early researchers.

We thank Jason Jones, who served as the book's Managing Editor. He recognized the importance of creating a research design handbook for scholar-practitioners and guided us gracefully through its development and production. We also thank Anne Jacobs, Laureate Publishing's Director, who oversaw the project and provided us with support, advice, and guidance as the book took shape. She has been a strong proponent of our work and the utility that *The Scholar-Practitioner's Guide to Research Design* will have for students. Our gratitude also goes to our colleagues with whom we have had many hours of dialogue on teaching, the value of research, and service on doctoral dissertation committees. We have grown tremendously as researchers through these dialogues.

Finally, we thank all the students whom we have taught and mentored over the years. When we started out in our academic careers, we did not imagine that our life's work would be about preparing both academic researchers and scholar-practitioners. Through their honest struggles and successes as budding researchers, our students have taught us much about what they need in order to understand the research process. We conceptualized *The Scholar-Practitioner's Guide to Research Design* to meet their specific needs. This book is our way of saying thank you to all of our students for what they have taught us about teaching research.

Gary J. Burkholder, PhD
Kimberley A. Cox, PhD
Linda M. Crawford, PhD

Introduction

One of the hallmarks of graduate education is research. We have been prepared, at some level, to do research since our earliest years of schooling. Over time, our understanding of the meaning of *research* changes and is different depending on the context. For example, a student in junior high school conducts research by going to the Internet (and perhaps the library) for information to write a paper about a famous person; in the process, that student learns that research is associated with collecting and synthesizing information. At the baccalaureate and master's degree levels, students may write research papers for courses that involve collecting information from an expanded set of resources that integrates and summarizes knowledge in a particular area. Perhaps there is also the goal of applying that knowledge to the solution of practical problems. At the doctoral level, students are typically expected to produce original research that adds new knowledge to their discipline of study or use research to solve an applied problem in a novel way.

Research at all levels tends to come with some amount of anxiety. For example, in our earliest undergraduate education courses, we may ask ourselves, "Do I have enough sources to support a five-page paper?" However, that anxiety may be most pronounced at the doctoral level. Many students enter their doctoral programs being told that they will have to write the equivalent of a book and that the research process will take several years. To complicate matters, they also see in their programs of study the requirement for a number of research design and statistics courses, which may raise issues of math and other kinds of analytic performance anxiety. There are also many students who would like to have a doctoral degree, but who never intend on becoming professional researchers or academics. Perhaps they wonder if their programs will require less statistics and other kinds of analysis and design training, given that they plan on continuing to work in their chosen fields of study.

The large number of working professionals who seek doctoral degrees reflects the increasingly pragmatic focus on education worldwide. In the new knowledge economy of the 21st century, employability has surfaced as a key goal of higher education. No longer focused solely on creating new academics, universities are creating programs designed for people intending to work outside of academia. In our experience with distance education, people who otherwise would not be attracted to a doctoral program now have flexible options to achieve a degree that can bring with it professional prestige, higher salaries, and the potential to expand their careers. These opportunities have created a growing demand for education by a large class of students known as scholar-practitioners.

Tenkasi and Hay (2008) defined the scholar-practitioner as one who "ha[s] one foot in each of the worlds of academia and practice and are pointedly interested in advancing the causes of both theory and practice" (p. 49). Hay (2004) provided an example from management in that scholar-practitioners "promise to be boundary-spanners who move between the worlds of academic and business in order to generate two outcomes—new theoretical knowledge and organizational results" (p. 1). Kram, Wasserman, and Yip (2012) used the definition from Tenkasi and Hay, among others, to define the scholar-practitioner as "one who actively engages in developing new knowledge and applying it in practice" (p. 334). Most definitions of scholar-practitioner carry implicit messages about the role of research: (a) scholar-practitioners are likely not positioning themselves as academics, although scholarship is a key part of their identity, and (b) research is important to the scholar-practitioner who seeks to understand and apply theory to solve practical problems arising in the professions.

Thus, the scholar-practitioner has a responsibility to apply the scientific method, gather appropriate evidence situated in the knowledge within the discipline, and use theory to solve important problems in practice. One term used frequently in the health care professions is *evidence-based practice*, whose goal is "the use of the best available evidence together with a clinician's expertise and a patient's values and preferences in making health care decisions" (Agency for Healthcare Research and Quality, 2015, p. 1). The concept of evidence-based practice is easily expanded to most other professions because it involves using scientific evidence, theory, professional expertise, and situational characteristics to achieve results in patients, businesses, schools, and the many other settings in which scholar-practitioners operate. This is not to say that scholar-practitioners are not interested in theory; one of the primary goals of any research is to understand the role of theory in the discipline and to use evidence and application to advance theory. Here, we recognize that many scholar-practitioners have as their goal the application of scientific discovery and theory to problems in practice. It is these scholar-practitioners for whom this book was primarily written.

The Purpose of the Book

In some ways, doing research is like learning a foreign language: It takes dedicated practice to learn the vocabulary; to acquire listening, reading, writing, and speaking skills; and to become proficient in the grammar. It also takes a lot of patience. Our goal in writing *The Scholar-Practitioner's Guide to Research Design* was to provide a straightforward and easy-to-understand text for students conducting research in the social, behavioral, health, management, and education sciences (for ease, when we refer to *social sciences* going forward, we would consider that inclusive of these discipline areas) that would help them learn this new language of research. It was written with a specific orientation to the scholar-practitioner who is new at research, with that scholar-practitioner being someone at either the master's or doctoral degree level. We wrote this book with the knowledge that there is a paucity of texts on the market that simultaneously present the basic concepts of research for scholar-practitioners and establish a foundation upon which to build a deeper understanding of the issues related to research methods relevant to their particular disciplines and interests.

Our intent was not to present detailed analysis of the various epistemologies and traditions of research methods associated with different choices of approaches to inquiry or design. Rather, our goal was to provide you (the scholar-practitioner) with an overview of the main topics associated with research to help you establish a foundational understanding of research methods, then to follow this information with references to readings that we consider to be essential for further exploration. You should consider this book as an entry into the much richer world of research and all of its interpretations, as there are many ways that research has evolved. However, it is important to understand the fundamentals before delving deeper, because such understanding will allow you to seek the most appropriate expression of research in your own work as a scholar-practitioner.

Organization of the Book

The Scholar-Practitioner's Guide to Research Design is divided into two parts. Part 1 addresses the foundational concepts of research design and methods. This is the essential information that every scholar-practitioner needs to know in order to be reasonably conversant in research methods, to be a critical consumer of research findings, and to begin the process of writing a research proposal. This information also forms the foundation upon which you can build your own understanding of theory and research in your discipline and use that new understanding to engage in the research that is most meaningful to answer the research questions of interest. The topics also provide sufficient understanding of research endeavors to avoid some common misconceptions students have about research. For example, in our experience, we find that many still say they choose qualitative research because it is easier or they choose mixed methods research because they were told that these are the only studies that journals will publish. This book strives to address such misconceptions.

The presentation of the content in Part 1 provides a logical flow of the issues relevant to research; however, although we advise reading Chapters 1 and 2 first, the remaining chapters can be read and studied in no particular order. Conversely, we strongly recommend that students master the concepts presented in Part 1 before moving on to Part 2; as such, we encourage that all chapters of Part 1 be read prior to progressing to Chapter 10, given that Chapter 10 integrates information presented in the previous chapters in its exploration of writing the research proposal.

Part 2 provides a more detailed exploration of a number of research designs and methodologies that—based on our combined experience of over 60 years of working with scholar-practitioners—we have found doctoral students tend to choose. As the chapters in Part 2 make clear, these are the designs and methodologies that are likely to be grounded in practice and that naturally address the real-world problems that scholar-practitioners want and need to solve.

The information conveyed in Part 2 will prove useful for students who are trying to decide whether a particular design most appropriately addresses their research question, as well as for those who have decided on a design and want to know what resources will help them achieve the requisite depth of understanding of that design or methodology. Although our assumption is that you will have read Part 1 first, the book is sufficiently flexible that you could focus on the particular Part 2 chapters of interest to you and those that are relevant to the likely research designs to be considered for the study.

Following is a brief overview of the chapters in *The Scholar-Practitioner's Guide to Research Design*.

Part 1: Foundations in Research Design

- **Chapter 1, Introduction to Research**, explores key research vocabulary and the foundations of research. The chapter provides the new researcher with the essential building blocks of the language of research upon which to build fluency. The key steps of the research process—which necessarily link to subsequent chapters in Part 1—are also outlined.
- **Chapter 2, Philosophical Foundations and the Role of Theory in Research,** provides an introduction to the philosophy of science and to theory and explores how each influences the research conducted by social scientists. The chapter focuses on some of the key issues related to science, including science as a means of generating knowledge; explanations of ontology and epistemology and their bearing on knowledge and the choice of research paradigm and design; positivist and postpositivist orientations to understanding truth; and the nature of causality in the social sciences.
- **Chapter 3, Conceptual and Theoretical Frameworks in Research**, examines conceptual and theoretical frameworks and the roles that they play in guiding the research study and grounding it in the current literature in the discipline. The chapter explores the definitions of conceptual and theoretical frameworks used by researchers and provides a definition that will be used to guide the remaining chapters of this book. Also included is practical guidance on how to develop a conceptual framework that guides the development of the research proposal.
- **Chapter 4, Quantitative Research Designs**, provides a foundational overview of quantitative research designs and explores their strengths and limitations. The chapter explains important design-related terms and concepts (including *variable*, *operational definition*, *levels of measurement*, and *types of variables*). The concept of sampling and common sampling strategies are also discussed.
- **Chapter 5, Qualitative Research Designs**, introduces the readers to qualitative research designs. The chapter reviews the purposes of qualitative research, examines definitions used by other experts in qualitative research design, and explores the qualitative designs most frequently chosen by scholar-practitioner researchers. The underlying philosophy of sampling in qualitative research is described, with guidance provided on sample sizes for different kinds of studies, and the concept of

trustworthiness in qualitative studies is discussed. A description of the various roles of the researcher in qualitative designs is also included.

- **Chapter 6, Mixed Methods Designs and Approaches**, provides an introduction to mixed methods research, which is the paradigm that tends to suffer the most misunderstanding. The chapter provides a clear overview of what mixed methods research entails and the reasons that researchers might choose this particular strategy. Included is a description of the philosophical underpinnings of mixed methods research and the different types of mixed methods research designs, as well as strategies for integrating quantitative and qualitative methods when interpreting the results of mixed methods studies.
- **Chapter 7, Quality Considerations**, describes issues related to quality in qualitative and quantitative research. The chapter provides an overview of the importance of quality of data sources and methods, as well as of data collection and analysis efforts. The concepts of reliability and validity in experimental research designs are discussed, with a particular focus on internal and external validity; statistical conclusion and construct validity; and trustworthiness (the equivalent to validity and reliability in quantitative research).
- **Chapter 8, Ethical Considerations**, provides a detailed overview of the ethical issues of most relevance to student researchers, including a description of the historical evolution of ethical codes and a comprehensive listing of professional and international codes of ethics. Topics discussed include assessing risks and benefits to participation, the function and composition of institutional review boards, resources for formal training in research ethics, informed consent, and methods for protecting privacy of data. In addition, several topics of contemporary importance are addressed—the use of deception in experimental research, conducting research with vulnerable populations, and the ethical implications associated with Internet research.
- **Chapter 9, Critiquing and Synthesizing the Literature**, offers practical strategies for critiquing literature, as well as for synthesizing it. Important considerations related to each skill are presented, including common mistakes and strategies for organizing the reading and writing processes. The chapter also provides both good and bad examples of the critique and synthesis of research.
- **Chapter 10, Writing the Research Proposal**, coalesces the information from the previous chapters into a practical guide to understanding and writing the key sections of a research proposal. Topics discussed include the problem statement, the purpose statement, research questions, and the theoretical and conceptual framework that comprise the rationale. Tools to help ensure alignment (or consistency) across the key sections of the research proposal are also provided.

Part 2: Special Topics in Applied Research Design

- **Chapter 11, Grounded Theory**, offers a thorough overview of the concepts associated with grounded theory approach and analysis. The chapter introduces the origins and philosophical underpinnings of grounded theory, gives examples of how grounded theory has been used within several disciplines, and provides information on appropriate sample size data collection and analysis. Constant comparative analysis, a feature unique to grounded theory, is also described.

- **Chapter 12, Phenomenology**, provides a detailed overview of an approach that is both a method and a methodology and that has a rich philosophical tradition. The chapter examines the origins and philosophical underpinnings of phenomenology and gives examples of the use of phenomenology across a variety of disciplines and the research questions that are most appropriate. Issues related to sample size, the role of the researcher (which is rather unique to phenomenological data collection and analysis), and issues related to quality are also discussed.
- **Chapter 13, Survey Research**, offers an excellent overview of the key issues facing those who choose to create their own instruments. The chapter defines survey research, provides a brief historical overview of the evolution of the methodology, and gives examples of survey research that represent multiple disciplines in the social sciences. Resources for locating and gaining permission to use existing instruments and a description of the process of creating an instrument are also included.
- **Chapter 14, Case Study Research**, explores using case study as an approach to understanding a phenomenon. The chapter defines case study and describes the origins and philosophical underpinnings of the approach. Discipline-specific applications of case study methodology are provided, as well as a description of the appropriate research questions addressed by case studies. In addition, methodology considerations (including sample size, data collection, and data analysis) are discussed and some misconceptions about case study research are clarified.
- **Chapter 15, Action Research**, discusses this practical approach to research (which involves researchers and key stakeholders as partners) and compares it with traditional experimental research. The chapter provides an overview of what action research is (and is not), its historical origins, and its philosophical underpinnings. The unique roles of the action researcher, as well as the role that stakeholders play in executing the research project, are described. In addition, various aspects of the methodology (including data collection and analysis, sample size determination, and ways to establish rigor) are explained.
- **Chapter 16, Program Evaluation**, examines the use of this strategy in research, describing the foundational role of stakeholders in program evaluation and the logic model of evaluation and explaining the key differences between process and formative evaluation and between outcome and summative evaluation. The chapter also provides examples of program evaluation across the disciplines and explores methodological aspects of program evaluation, such as sample size, data collection and analysis, and issues of quality. In addition, a number of web-based resources are included for those who want to pursue program evaluation as a research strategy.

- **Chapter 17, Interviewing Essentials for New Researchers**, offers helpful insights into the key issues that arise when conducting interviews and provides the new researcher with some practical tips for success. Several types of interviews commonly used in research—including structured, semi-structured, and unstructured—are discussed and the differences between individual and focus group interviews are explained. The influence of power and culture dynamics in interviews are also explored.

As you read, keep in mind that *The Scholar-Practitioner's Guide to Research Design* text is designed for you—it is meant to be a flexible resource to help you gain the confidence to conduct scholarly research, whether as an academic or as a scholar-practitioner. Use the key resources provided in each chapter as a guide for further exploration. Bring any questions you have to colleagues, your professors, or your dissertation/thesis committee; each of these people can help you to achieve the depth of understanding of basic research principles that you will need to complete a successful research study.

Lastly, remember that research is a scientific enterprise, but it is also a social enterprise; scientists learn from each other and slowly accumulate a knowledge base in a discipline. Kuhn (1962) wrote that "normal science, the puzzle-solving activity we have just examined, is a highly cumulative enterprise, eminently successful in its aim, the steady extension of the scope and precision of scientific knowledge" (p. 52). But Hacking (1983) also acknowledged that "there can be heapings up of knowledge without there being any unity of science to which they all add up" (p. 55). Each discipline has a rich store of knowledge, some organized through theories and some not. The important work you do as a researcher steadily extends the scope of knowledge. You—as does every researcher—bring a fresh perspective to the discipline. Engaging that perspective in research helps to bring more clarity (or further questions) to specific topics of importance in your chosen field of study. We sincerely hope that *The Scholar-Practitioner's Guide to Research Design* empowers you to probe the depths of understanding of the research topics of most interest to you and to the discipline and that it encourages you to craft a research study that addresses the important questions that need answering.

Gary J. Burkholder, PhD
Kimberley A. Cox, PhD
Linda M. Crawford, PhD

REFERENCES

Agency for Healthcare Research and Quality. (2015). *Evidence-Based Decisionmaking*. Retrieved from http://www.ahrq.gov/professionals/prevention-chronic-care/decision/index.html

Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge, United Kingdom: Cambridge University Press.

Hay, G. W. (2004). Executive PhDs as a solution to the perceived relevance gap between theory and practice: A framework of theory-practice linkages for the study of the executive doctoral scholar-practitioner. *International Journal of Organisational Behavior*, *7*(2), 375–393.

Kram, K. E., Wasserman, I. C, & Yip, J. (2012). Metaphors of identity and professional practice: Learning from the scholar-practitioner. *The Journal of Applied Behavioral Science*, *48*(3), 304–341. doi: 10.1177/0021886312439097

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

Tenkasi, R. V., & Hay, G. W. (2008). Following the second legacy of Aristotle: The scholar-practitioner as an epistemic technician. In A. B. Shani, S. A. Mohrman, W. A. Pasmore, B. N. Stymne, & N. Adler (Eds.), *Handbook of collaborative management research* (pp. 49–72). Thousand Oaks, CA: Sage.

1

Introduction to Research

Gary J. Burkholder, PhD

Introduction and Overview

Research is a hallmark of higher education. In the 19th century, research became an integral part of the university ethos. In the United States specifically, that role assumed a pragmatic function as "American reformers further transformed higher education by stressing the relationship between the university and society through the concept of service and direct links with industry and agriculture" (Altbach, 2011, p. 17). Earlier models of higher education included significant focus on research. As higher education evolved, and particularly in the United States, where a more pragmatic approach to higher education was realized, it became apparent that higher education needed to be more directly connected with the needs of society. For example, these needs included teaching and research in industry and agriculture. Even today, there is an apparent worldwide shift in focus to employment and employability training, which has opened new paths to professional doctorate degrees oriented toward scholar-practitioners who value the application of theory and research to practical problems. Thus, the role of research, and science, continues to be an important part of higher education at the doctoral level.

Science has played a significant role as a means of generating new knowledge. Historians of science can trace the origins of a scientific method to the time of Aristotle and earlier. Aristotle was probably one of the first who formalized an approach to knowledge generation that involved a process of inquiry and analysis in which general principles can be formulated from what is observed (induction), and from these general principles, hypotheses can be derived that can be used to test those principles (deduction). The results of the tests of the hypotheses can then be used to inform theory. The cycle of induction and deduction is what is responsible for the creation of new knowledge (Gauch, 2003). In the 18th and 19th centuries, rationalism, defined as the belief that knowledge can be created through internal reflection and logic, was superseded by empiricism, the belief that what is known is that which is discoverable by the senses and ultimately measurable. It is empiricism that has dominated contemporary scientific research.

The scientific method has provided a powerful approach toward knowledge generation and dissemination. The scientific method has led to the discoveries that have revolutionized technology and the ways we use it, our understanding of the nature of the universe, and the eradication of diseases. These discoveries are based on what can be directly seen in the microscope and the telescope and what can be measured and strictly controlled. Those discoveries came from the natural sciences, such as chemistry and biology, colloquially referenced as the *hard sciences*. In the social or *soft sciences*, by comparison, applying typical scientific standards is more challenging. For example, in the social sciences, it is difficult to isolate one factor that is suspected to cause a certain kind of behavior and measure its impact. People behave in expected and unexpected ways that may change in different contexts and times. This variation creates a major challenge in trying to explain and predict human behavior. Scientists try to isolate behavior and devise research studies to understand the effects of that behavior; the results are then interpreted by scientists and reviewed by the wider community of scientists as well as the public. Despite measures to isolate variables and validate survey instruments and assessments and thus bring objectivity into the process of discovery, interpreting human behavior is ultimately subjective and raises questions about what we can truly know and understand about behavior. It also raises questions about scientific interpretations of those behaviors. Therefore, it is not surprising that people have varying levels of confidence in the scientific enterprise as well as its explanations based on research.

Funk and Rainie (2015) compared the beliefs of scientists from the American Association for the Advancement of Science and those from a random sample of the American public and found some encouraging results: Overall, 79% of the adults reported that science has made life easier for most people, and 71% felt that investments in science have paid off in the long run. However, other findings provide a more nuanced examination of attitudes toward science. For example, 88% of scientists believed that it is safe to eat genetically modified foods, whereas only 37% of the adults surveyed believed they are safe. The same, but smaller, differences were found with evolution: Overall, 98% of scientists versus 65% of adults believed that humans have evolved over time. With vaccines, 86% of scientists versus 68% of adults believed that vaccination of children should be required. With climate change, 87% of scientists versus 50% of the public believed it was due to human activity. Compare this with the finding that the public believes that only 57% of scientists agree that climate change has been caused by humans. The differences reflected here are quite significant and suggest a public misperception of what scientists have come to understand through research. Additionally, Pew researchers found that perceptions of science and its contribution to society have decreased among scientists and the American public, both down by about 11 points. This report shows that in many ways, the perceptions of the role of science are decreasing over time, and on some key issues, significant differences in belief exist between scientists and the nonscience public over the role of science in understanding those issues.

Given the rich historical origins of scientific research, the important role that research plays in society, and the views of the value of scientific research that often differ between scientists and the public, it is important for new social science researchers to be well grounded in the science that underlies their specific disciplines as well as to understand the strengths and limitations of scientific research. Due to the changing landscape of social science research, it is important for you to understand the foundations and language of research as well as its methods. With this understanding, you can enter confidently and knowledgably into the dialogues that characterize scholarship in the disciplines, in the debates that occur over what is scientific *truth*, and be prepared to engage in ethically and socially responsible research.

The purpose of this first chapter is to provide an overview of essential terminology of research. The language of research is found in many of the research texts you will read as well as the hundreds of peer-reviewed and non-peer-reviewed research manuscripts you will examine as part of the process of learning the state of scholarship in your discipline, so understanding the language will help you to become a better producer of as well as a critical consumer of research. Many authors use terminology related to research differently and, in some cases, incorrectly. The goal is not to take a position on whether individual definitions are correct or incorrect; rather, I will define the terminology used in this text to help provide clarity in how terms are used in order to minimize confusion.

The chapter begins with an introduction to research approaches (qualitative, quantitative, and mixed methods) as well as distinctions among research approach, design, and methodology. Each of these will be carefully defined to ensure that there is no confusion during the reading of this text. Finally, the chapter ends with an overview of a description of the key phases of research, the components of which are described in various chapters throughout this text.

Introduction to Foundational Principles of Research

> Science is a social field of forces, struggles, and relationships that is defined at every moment by the relations of power among the protagonists. Scientific choices are guided by taken-for-granted assumptions, interactive with practices, as to what constitutes real and important problems, valid methods, and authentic knowledge. (Bourdieu, 1991, p. 3)

Bourdieu's writing underscores that science is a quintessentially social endeavor. Kuhn (2012) also referred to science as a social enterprise. Like any social group, social scientists have developed ways of speaking about how they engage with research. Given that multiple disciplines exist, each with varying perspectives on reality and differences in interpretation of which problems are important, it is not surprising to learn that each discipline has language and practices specific to its own unique perspectives. Methods that guide research in anthropology, for example, may be different from those that guide research in public health; experimental designs may be a more important standard in some disciplines than others. Thus, not all scientists necessarily adopt the same terminology or ways of conducting research. What I focus on in this chapter are some basic definitions of terms that are common across disciplines to establish a common language used in this text. The first important concept is *scientific method*.

Scientific Method

The *Oxford English Dictionary* (2014) provides the following definition for *scientific method*:

> A method of observation or procedure based on scientific ideas or methods; ***spec.*** an empirical method that has underlain the development of natural science since the 17th cent.
>
> It is now commonly represented as ideally comprising some or all of (a) systematic observation, measurement, and experimentation, (b) induction and the formulation of hypotheses, (c) the making of deductions from the hypotheses, (d) the experimental testing of the deductions, and (if necessary) (e) the modification of the hypotheses; though there are great differences in practice in the way the scientific method is employed in different disciplines (e.g., paleontology relies on induction more than does chemistry, because past events cannot be repeated experimentally). (para. 1)

The scientific method has characterized research in the natural sciences since the 17th century; thus, its basic tenets have become well established. When people think of the scientific method, it usually brings back memories of high school or college chemistry lab. The researcher carefully measures quantities of substances to introduce into the test tube in which the chemical reaction takes place, carefully controls all environmental variables, and interprets the results that are expected to confirm theory. This image of the scientific method, consistent with the definition above, has guided research in the social sciences since their beginnings (the laboratory view of science has been more of an influence in some disciplines than others). Particularly in the foundation years in which the procedures in the social sciences were being established, research typically emulated the methods that were showing success in the natural sciences.

However, in the social sciences, the research settings are often not laboratories, but rather located in real-world contexts involving individuals and groups. Several researchers have challenged the conventional definition of scientific method (e.g., Bauer, 1992; Windschitl, Thompson, & Braaten, 2008) precisely because specific methods, methodologies, and designs have evolved within scientific disciplines. Each discipline trains and orients researchers to the ways questions are asked, theory is approached, and data are collected and analyzed that may be specific to that discipline (Bauer, 1992). Windschitl et al. (2008) proposed one alternative to the scientific method, *model-based inquiry*, which espouses five key components of scientific inquiry.

- *Testable:* Knowledge advances by continuing to pose questions based on scientific theories. Questions can be answered using discipline-specific ways of collecting and analyzing data that lead to further questions.
- *Revisable:* Theories can be revised based on emerging evidence, and the revision leads to further questions that can be tested.
- *Explanatory:* The goal of science is to provide understanding of cause-and-effect relationships among variables and concepts.
- *Conjectural:* “Causal accounts often involve theoretical or observable process that can only be inferred from empirical observation (data)

and that scientific argument aims to persuade others that explanations based on these inferences account adequately for the observations" (Windschitl et al., 2008, p. 4), which essentially describes the role of abduction in science.

- *Generative:* Specific scientific models are not the end goal but rather create opportunities to continue to refine, support, or disconfirm. The knowledge achieved through science is constantly evolving.

The Windschitl et al. (2008) model appears to focus more on the process of scientific discovery. One can acknowledge differences among scientific disciplines and accept that there cannot be a single approach to data observation, collection, and analysis. Therefore, it is possible to extract from various definitions the characteristics of the *method* that make it scientific.

- *The scientific method requires systematic observation*. Scientists depend on careful, systematic observation to ensure data are as free from error as possible. Ideally, data would be 100% accurate, and the scales used to collect data would measure exactly what is intended; for example, IQ scores would 100% reflect actual intelligence. This is true whether the data are numeric data, as might be collected in a survey, or textual data, as might be collected in interviews.
- *The scientific method involves theory development and subsequent testing of emerging theories*. The goal of science is to generate theory that can then be tested from hypotheses derived from the theory (deduction) or to examine repeated occurrences of events from which a testable theory can be derived (induction). Abduction, the process of providing the best possible explanation for phenomenon given the data, is a slightly different approach; for example, abduction characterizes processes used in grounded theory analysis and exploratory factor analysis (Haig, 2005). Regardless of where the researcher starts—for example, one of the strengths of qualitative research is its deep exploration of phenomena that can be used to lead to theory generation or to better understand application of existing theory—a key focus of the scientific method is to generate and regenerate theory (while recognizing differences in the role of theory across disciplines).
- *The scientific method, as applied to the social sciences, seeks to understand, explain, and predict behavior. This implies understanding cause-and-effect relationships*. The goal of science is to describe what can be seen, use the description to generate theory to explain behavior, and analyze data to test the extent to which the theory consistently explains behavior across time and place.
- *The scientific method seeks objectivity, to the extent possible constrained by its methods*. In survey research, where people provide numerical responses to questions, the researcher is relatively distant from the analysis. Numbers are submitted to statistical tests and yield results that are or are not statistically significant. In interview research, by comparison, the researcher or research team serves as both the instrument for data collection as well as the means of analysis. The researcher

> as instrument introduces subjectivity that is recognized and appreciated as an important source of variation that adds richness to interpretation of textual data. Regardless of the type of data, interpretation of results is always situated in the context of personal beliefs and worldview, of discipline mores and norms, and within the larger sociopolitical landscape. Even in cases where subjectivity is acknowledged, there are methods for bracketing, which means minimizing or putting to the side the role of researcher bias and experience. The scientific method provides the means, regardless of research approach, to appropriately account for subjectivity, particularly in qualitative research, and to provide interpretation that is as free of bias as possible.

Scientific research is one of the fundamental ways knowledge is added to the discipline and theory is confirmed or disconfirmed. For the most part, those pursuing master's degrees are expected to develop research skills that are used to apply existing knowledge and may not necessarily result in new knowledge. Those pursuing doctoral degrees, however, are expected to add new knowledge to the discipline (or perhaps a novel application of existing knowledge). Your goal may be purely academic; that is, you may strive to be a university professor and to conduct scholarly basic and applied research. Or you might be interested in a career that is more oriented toward the scholar-practitioner who is interested in the application of theory and evidence to solve local problems. In either case, you will use the scientific method, and the goal of the scientific method in both examples is the same: to generate new knowledge.

Paradigm

Kuhn (2012), in his book on scientific revolutions, described progress in science as the accumulation of knowledge that is periodically interrupted by theories or new ways of framing experience that result in a significant shift in the way we think about old questions and new ones. What emerges is a new paradigm. A paradigm is an organizing principle for new ideas and new frameworks for describing phenomena. The classic example of a paradigm shift is the change from a simple mechanical view of the structure of matter to a view framed by Einstein's theory of general relativity.

The choice of the term *paradigm* in the social sciences has been rather muddled. It seems to be used by various authors both to refer to the philosophical orientations, such as positivism and constructivism (e.g., see Guba & Lincoln, 1994; Houghton, Hunter, & Meskell, 2012), and to describe qualitative, quantitative, and mixed methods research strategies (e.g., see Denzin, 2009; Johnson & Onwuegbuzie, 2004). Reynolds (2007) indicated that many scientific theories are paradigms as well. Houghton et al. (2012) defined paradigm (citing Guba & Lincoln, 1994) as "a set of basic beliefs or a frame of reference that explains how individuals perceive the nature of the world and their places in it" (p. 34). Part of the reason for the variety in definitions may be because of the close connection between theory and the philosophy underlying theory. In other instances, it seems that the definitions may be too limiting, such as the reference to theories as paradigms.

The term *paradigm* will be used in this book to refer to the constellation of assumptions and orientations that guide researchers to adopt particular approaches to inquiry and their methods. Mittwede (2012) summarized the concept of paradigm clearly. He wrote,

> How is research to be conceived? Moreover, how ought reality be perceived? How do we know what we know, and to what extent can we be sure of it? These are the types of issues that are germane to the subject of research paradigms, for every paradigm is based upon presuppositions regarding reality and how it may be understood; thus, these paradigms are essentially matrices of deeply held assumptions. (p. 23)

Given the philosophical nature of this use of the word *paradigm*, authors will often also refer to paradigms as *philosophical orientations*, which is commonly used in the literature to reference the philosophical assumptions underlying the researcher's worldview and his or her research processes. The paradigms, or philosophical orientations, that guide social sciences research will be explored in more detail in Chapter 2.

Approaches to Inquiry

I refer to qualitative, quantitative, and mixed methods as *approaches*, similar to how Creswell (2013) employs the term. Qualitative approaches to inquiry are primarily for deep explanation of a phenomenon and are often used to develop theory. They are generally inductive, and the data from which the qualitative researcher works are principally textual, or narrative. These data include transcripts of interviews, passages from books, videography materials, and other sources. The qualitative approach has historically been associated with fields such as anthropology, education, and sociology. Other sciences, such as psychology, were primarily characterized by quantitative methods because of their evolution from the practices in the natural sciences.

Qualitative research gained popularity in other sciences, including psychology and public health, in the 1960s and 1970s. The rising recognition of communities and populations marginalized or ignored in research, such as those with race/ethnicity other than White; lesbian, gay, bisexual, and transgender communities; and women, brought a new imperative for an interpretive and explanatory approach to research in order to understand the experiences of members of these groups and how those experiences map to existing social science theories (e.g., see Smith, 1999, regarding research on indigenous peoples; Kitzinger, 1987, for one view of the marginalization of minority groups from traditional social science research; Lather, 1994, for views on feminism in research; Stanfield & Rutledge, 1993, for issues of race and ethnicity in research). Today, qualitative approaches to inquiry are increasing in other sciences as well, such as management and public health. The qualitative approach to inquiry is explained in more detail in Chapter 4.

The quantitative approach is primarily deductive: Hypotheses are generated, data are collected, and hypotheses are tested to see if the data provide support for theory. The quantitative approach is based on the model adopted in the natural sciences at the start of the scientific revolution; quantitative approaches tend to be associated with positivist epistemologies (described further in Chapter 2). Its hallmarks are the control, to the extent possible and desired, of variables and the collection of numeric data from respondents directly or through observations. These data are then analyzed and reported. Quantitative approaches are frequently used to test theory. The quantitative approach is described in more detail in Chapter 5.

Mixed methods research, the third major approach to inquiry, combines both qualitative and quantitative approaches in a holistic way. Mixed methods as an approach emerged in the 1970s and 1980s. It is a key approach to program evaluation research in which quantitative survey data are combined with interview and artifact text data to understand the efficacy of programs. Mixed methods as an approach to inquiry have been examined from philosophical, theoretical, and pragmatic perspectives; for examples, see Teddlie and Tashakkori (2008). Mixed methods approaches are rigorous in terms of justifying and integrating the various philosophies and theories representing both quantitative and qualitative approaches. Additionally, they tend to take longer to complete (and actual funded studies can be costly compared to purely qualitative and quantitative studies). For these reasons, you should carefully consider whether a mixed methods study is right for you; this book will help you with this decision (see Chapter 6).

Design

Within each of the three approaches to inquiry—qualitative, quantitative, and mixed methods—researchers have a number of choices of specific types of designs. Trochim (2006) defined research design as "the glue that holds the research project together. A design is used to structure the research, to show how all of the major parts of the research project work together to try to address the central research questions" (para. 1). For example, in qualitative research, there are several designs from which to choose, including case study, ethnographic, and grounded theory, among others. Note that particularly in qualitative research, designs such as case study can comprise multiple designs. For example, a case study may include quantitative surveys as well as grounded theory interviews. In quantitative studies, researchers can choose from experimental designs, such as randomized controlled trials; quasi-experimental designs, in which random assignment to groups does not happen; or observational studies. There are also nonexperimental designs that include correlational studies, which examine the natural relationships among variables with no researcher manipulation. The design also includes the fundamental research question(s) guiding the study and the hypotheses where appropriate. Last, the design informs many, if not all, parts of the study methods, such as participant recruitment, instrumentation selection, data collection, and analysis. The many facets of a study design emphasize the importance of thoroughly understanding the traditions associated with the choice of approach and design.

Methods

I reserve the term *methods* to describe the various components involved in executing the study: the design as previously described, population and sampling, how data are collected, and how they will be analyzed. In qualitative research, participant selection criteria need to be described because populations and

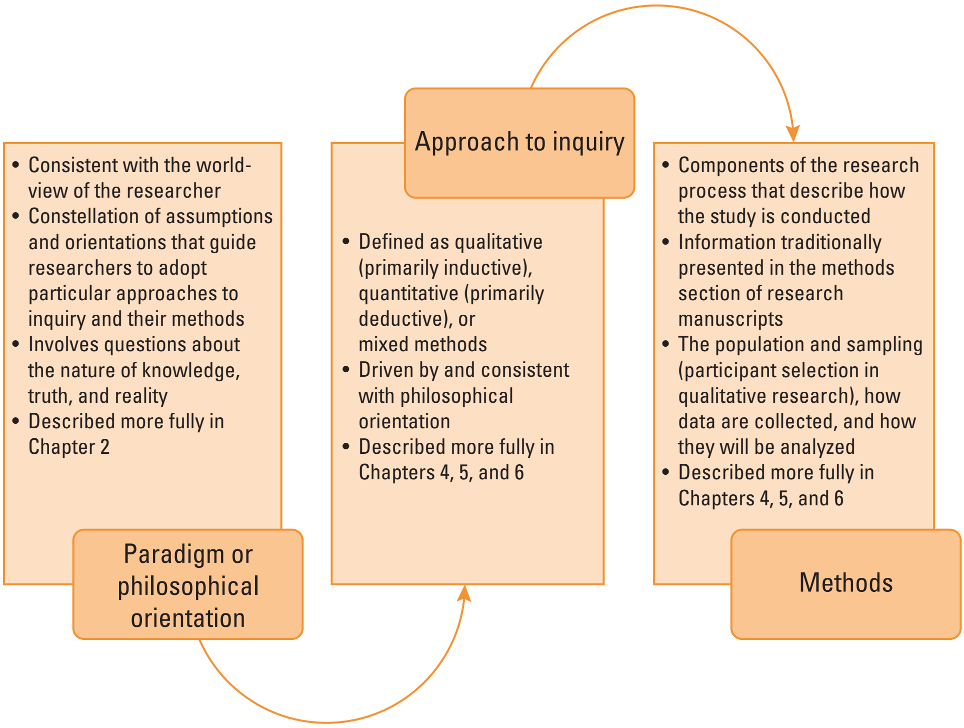


FIGURE 1.1 The relationships among paradigm, approach to inquiry, and methods.

samples are typically less meaningful in qualitative research. For quantitative research, population and sampling require identification of the population and how the members of the population will be selected; this is the sampling strategy. For both approaches, population and sampling include explanation of decisions about how participants will be included or excluded from the study. Scientists also need to justify the sample size that is required to achieve the level of confidence in the results. Data collection involves detailed description of how and when data will be collected as well as of the instrumentation. For example, instrumentation can be observation forms, interview protocols, tests, and surveys. Also required are participant procedures, specifically how data are being collected in ways that are ethical and honor choices of participants. Data analysis involves describing how data will be cleaned, coded, and analyzed. Together, these details convey to members of the broader scientific community, scientists within the specific discipline, and practitioners who read the studies how the research was (or will be) executed. These components of the research proposal define the methods for the study.

Figure 1.1 provides a graphical representation of the relationship between paradigm/philosophical orientation, approach, and methods and how each is related to and derives from the another.

Stages of the Research Process

Many master's-level students in research-oriented programs do replication studies because the goal is to learn how to do research. In doctoral programs, the goal is higher: to conduct independent research. One of the purposes of this text is to serve as a guide to the essential information needed to craft a scientific research proposal that will lead to an empirical research study. This section provides a basic framework to guide you in completing the research proposal. Use this framework as a cognitive map to guide you in mastering the components required for each of the stages of research presented in subsequent chapters.

1. *Determine an area of interest*. The first step in the research process is to identify an area that you want to research. This first idea can emerge from several sources. It can result from scholarly research, or it may be a problem in the setting in which you work that needs to be addressed. It can also be an idea that you have had that you would like to explore further. Ideas for great research originate for many reasons, and what is important is that any research idea is pursued systematically to determine if it is research worthy.
2. *Reflect upon your worldview and orientation to research*. It is important as scholars to be able to articulate your view of the world. Is your interest in prediction, or is it more about explanation and interpretation of experience? Are you more comfortable with distance and objectivity, or do you prefer to have direct contact with your participants and share in the discovery of meaning of their experiences? Do you believe that there is one "Truth," or is "truth" a relative concept that has meaning only in the context of individual or collective experience? This reflection is important for you to articulate because how you read the literature and how you design your study will be driven in part by your views of truth and reality. The importance of worldview and the ways that philosophical perspectives influence research are described in Chapter 2.
3. *Conduct a thorough literature search*. The systematic search is your strategy for finding the relevant literature that may exist in multiple sources. Depending on the nature of the research, this can include peer-reviewed research manuscripts, doctoral dissertations, books, and non-peer-reviewed publications such as magazines and newspapers. It might also include video footage that you have collected. Data come in numerous forms. Sometimes, finding the literature and other artifacts can be challenging, and this is particularly true for areas of exploration that have not had the benefit of in-depth scholarship. You may, for example, find several artifacts from the popular culture but very little from peer-reviewed scholarship. Your librarian can be very helpful in navigating the various sources of literature. More information about literature searches is provided in Chapter 9.
4. *Conducting and writing the literature review*. The literature review involves thorough reading of the books, manuscripts, and other documents, as well as listening to audiovisual artifacts, that you collect. This careful review allows you to understand the accumulation of knowledge in a particular area as well as the limits of that knowledge. The review also provides insight into the dominant theories in the discipline used to explain the phenomenon of interest. Thus, you are reviewing for what scientists know as well as what they do not know in the particular area of scholarship. The literature review is useful to determine the different research

paradigms, approaches to inquiry, designs, and methodologies that have been used to explore the topic. All are critical for a thorough understanding of the topic. More about the literature review is provided in Chapter 9.

5. *Determine the research problem*. Once the literature has been reviewed, the research problem needs to be identified. This essentially is a statement about the gap, or what still needs to be answered, in your area of scholarship. This is a critical step. Without a clearly articulated problem statement, the research will likely not make a contribution to new knowledge in the field. Chapter 10 provides more information on the research problem.
6. *Determine the purpose of the study*. Once you have determined the problem and you have determined that you have a legitimate research study that adds new knowledge, you need to articulate the purpose of the study. The purpose of the study provides your readers with a concise statement of the goal of the study that includes information on the design, population and sample, setting, and key variables and concepts being examined. Chapter 10 provides more information on the purpose of the study.
7. *Determine the research question*. This is an important step because the research question helps define the scope of the study, guides the development of the problem statement (i.e., what gap in the literature your research is filling), and helps influence the appropriate research design. Each study should have one guiding research question (some may have two or three, but if there are more, it may be that the scope of the study is too large and should be examined). For example, a research question guiding a phenomenological study might be “What are the lived experiences of veterans returning from war and reintegrating into their communities?” Chapter 10 provides more information on the research questions.
8. *Develop the conceptual and/or theoretical framework*. Concepts and theories have a reflexive role in your research: Your understanding of concepts and theories in the discipline related to the topic of interest emerges from the literature review. In conjunction with your worldview, the concepts and theories then contribute to the overall conceptual/theoretical framework that guides your research study. The conceptual framework provides a logical structure to the research study that connects research questions, methodology, and design in a way that is internally consistent. The conceptual and theoretical frameworks are described in detail in Chapter 3.
9. *Determine the appropriate research approach and design*. Based on the research question and the framework, you will need to decide on which approach is most relevant for the study. Review of the literature will be a primary source of information guiding the design. First, the literature informs the paradigms used previously to explore the topic. Second, from published, peer-reviewed studies, you can see the types of designs that have been used in the past to address the problem. It may be that you explore literature in other disciplines that suggest designs that have been applied but that may be novel in your home discipline. The research design and the methodology need to be consistent with the problem, purpose, and research questions. Chapters 4, 5, and 6 provide more detailed information on research designs.

10. *Develop the methods*. Once you have determined the correct design, the methods need to be articulated. To paraphrase from the previous section in this chapter: The methods include from whom you will collect the data, how the data will be collected, what instrumentation will be used to collect the data, and how the data will be analyzed. Chapters 4, 5, and 6 provide information on methods associated with qualitative, quantitative, and mixed methods approaches. Chapter 10 provides more information on presenting the methods in a research proposal.
11. *Align paradigm, approach, and methods*. Internal alignment and consistency are critical for a logical proposal. You, and members of your research committee, must always be aware that changes in any part of the overall proposal can result in inconsistencies among the various sections. For this reason, I encourage the use of an alignment tool, presented in Chapter 10, that helps everyone to ensure that changes in one aspect of the proposal are considered across all components of the research proposal.
12. *Gain approval for the study*. All research must undergo a process of approval at the sponsoring institution. Universities, hospitals, and other institutions that sponsor research and receive federal funds have an institutional research board (IRB) that serves to review and approve research consistent with federal requirements. The responsibility for the IRB is to ensure the ethical conduct of research. More about ethics in research will be presented in Chapter 8.

Summary

The chapter began with an exploration of *scientific method* and the key factors that characterize a scientific approach to understanding phenomena. Following these explanations, I positioned the key terms related to research that will be used throughout the book. Even though some of the terms are used differently and use of these terms can vary across academic programs and within and across disciplines, what is important is that you understand how terms are being used in this book so as to minimize confusion and maximize clarity of principles. In the end, you decide on the way you will use terms; any use of terminology should be clearly presented in the research proposal and backed with appropriate evidence so that it is absolutely clear to your readers. You are the scientist. Your task is to design and conduct original research whose paradigm, approach to inquiry, design, and methods are internally consistent and logical as well as to design a research study that provides for ethical treatment of participants. Any decision you make will require collecting the necessary evidence. As long as you back up your claims with solid evidence, it will be difficult for others to question those claims.

Chapter 2 provides a more detailed discussion of the philosophical underpinnings of research. In that chapter, Dr. Patricia Burbank joins me in describing the research paradigms (or philosophical orientations), as well as some of the persistent yet important issues in science, such as the nature of causality. In addition, the important role of theory in research is explored, and the continuum of inquiry is introduced. The continuum of inquiry concept can be used to guide researchers in making key decisions about the approaches to inquiry that make the most sense given the state of knowledge in the discipline.

KEY SOURCES

Houghton, C., Hunter, A., & Meskell, P. (2012). Linking aims, paradigm and method in nursing research. *Nurse Researcher*, *20*(2), 34–39.

Kuhn, T. (2012). *The structure of scientific revolutions: 50th Anniversary edition*. Urbana, IL: University of Chicago Press.

Patton, M. L. (2014). *Proposing empirical research: A guide to the fundamentals* (5th ed.). Glendale, CA: Pyrczak.

REFERENCES

Altbach, P. G. (2011). Patterns of higher education development. In P. G. Altbach, P. J. Bumport, & R. O. Berdahl (Eds.), *American higher education in the twenty-first century: Social, political, and economic challenges* (pp. 15–37). Baltimore, MD: Johns Hopkins University Press.

Bauer, H. (1992). *Science literacy and the myth of the scientific method*. Urbana, IL: University of Illinois Press.

Bourdieu, P. (1991). The peculiar history of scientific reason. *Sociological Forum*, *6*(1), 3–26.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: Sage.

Denzin, N. K. (2009). The elephant in the living room: Or extending the conversation about the politics of evidence. *Qualitative Research*, *9*, 139–160. doi:10.1177/1468794108098034

Funk, C., & Rainie, L. (2015). *Americans, politics and science issues*. Retrieved from http://www.pewinternet.org/2015/07/01/americans-politics-and-science-issues/

Gauch, H. G. (2003). *Scientific method in practice*. Cambridge, United Kingdom: Cambridge University Press.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 163–194). Thousand Oaks, CA: Sage.

Haig, B. D. (2005). An abductive theory of scientific method. *Psychological Methods*, *10*(4), 371–388.

Houghton, C., Hunter, A., & Meskell, P. (2012). Linking aims, paradigm and method in nursing research. *Nurse Researcher*, *20*(2), 34–39.

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

Kitzinger, C. (1987). *The social construction of lesbianism*. Thousand Oaks, CA: Sage.

Kuhn, T. (2012). *The structure of scientific revolutions: 50th Anniversary edition*. Urbana, IL: University of Chicago Press.

Lather, P. (1994). Critical inquiry in qualitative research: Feminist and poststructural perspectives. In B. F. Crabtree, W. L. Miller, R. B. Addison, V. J. Gilchrist, & A. J. Kuzel (Eds.), *Exploring collaborative research in primary care* (pp. 103–114). Thousand Oaks, CA: Sage.

Mittwede, S. K. (2012). Research paradigms and their use and importance in theological inquiry and education. *Journal of Education & Christian Belief*, *16*, 23–40.

Reynolds, R. D. (2007). *A primer in theory construction*. Boston, MA: Allyn & Bacon.

Scientific method. (2014). In *Oxford English Dictionary*. Retrieved from http://www.oed.com.ezp.waldenulibrary.org/view/Entry/383323?redirectedFrom=scientific+method&

Smith, L. T. (1999). *Decolonizing methodologies: Research and indigenous peoples*. London, United Kingdom: Zed Books.

Stanfield, J. H., II, & Rutledge, M. (1993). *Race and ethnicity in research methods*. Newbury Park, CA: Sage.

Teddlie, C., & Tashakkori, A. (2008). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks, CA: Sage.

Trochim, W. M. (2006). *The research methods knowledge base* (2nd ed.). Cincinnati, OH: Atomic Dog. Retrieved from http://www.socialresearchmethods.net/kb/

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education*, *92*, 941–967. doi:10.1002/sce.20259

2

Philosophical Foundations and the Role of Theory in Research

Gary J. Burkholder, PhD, and Patricia M. Burbank, DNSc, RN, FAAN

Introduction

Philosophy is the branch of study associated with understanding the fundamental nature of existence and reality. The *Oxford English Dictionary* (2006) defines *philosophy* as "the study of the fundamental nature of knowledge, reality, and existence, and the basis and limits of human understanding" (para. 7). Buckingham and Burnham (2015) wrote, "Philosophy is not so much about coming up with the answers to fundamental questions as it is about the process of trying to find these answers, using reasoning rather than accepting without question conventional views or traditional authority" (p. 13). The focus on reasoning has characterized knowledge generation since the time of Socrates, and so it is not surprising that philosophy has had such a strong influence on science and the scientific method.

Science is a body of knowledge that comes from systematic observation. When considering the term *science*, many think immediately back to their natural science classes, such as biology or chemistry, in secondary school. These disciplines involve knowledge generation through careful control, observation, and analysis, usually in laboratory settings. This method of knowledge generation can be somewhat generalized to other fields, such as those in the management, psychological, educational, and health sciences, that characterize a body of knowledge emerging from careful observation and analysis. The term *traditional science* is sometimes used to describe the scientific method that first evolved to explain physical phenomena and has been broadly applied through primarily quantitative approaches to explain human behavior. The traditional science method typically involves the generation or utilization of theory to explain phenomena in the natural world. Theories are then tested using careful methods of observation, data collection, and analysis.

It is apparent, and history shows, that there is an important role for the merging of these two disciplines, philosophy and science. A *philosophy of science* addresses fundamental questions about the nature of truth and the underlying assumptions of theories used to describe natural phenomena. In particular, philosophers of social science are interested in a number of implications of this intersection of science and philosophy because what happens during an interaction between chemicals in a laboratory appears to be fundamentally different from how animals and humans interact. The differences and similarities between the physical sciences and social sciences have been an ongoing debate among philosophers of science.

Some philosophers of science believe that the social and natural sciences are essentially the same and that their methods should closely correspond. They believe that social phenomena can be reduced to physical entities, which can be directly observed and are governed by physical laws. This reflects a reductionist view of science. Others, who hold the view that the social and natural sciences are inherently different, take the position that social phenomena involve explanations based on intention or motivation. Still others believe that the social sciences are radically different from the natural sciences and require an interpretive methodology that seeks to discover the *meaning* of human action.

These differences outline the range of perspectives that underlie decisions about how to answer research questions that arise in the social sciences. What is your perspective? If you believe that social sciences are similar to natural sciences, then traditional science methods using quantitative approaches are appropriate for the discovery of new knowledge in the social sciences. On the other hand, if you believe that social sciences are inherently different from natural sciences and that knowledge about human beings can advance only through the discovery of meaning through understanding intentions and motivations, then the traditional science methods will not work to uncover these meanings. Some refer to these as *nontraditional* science approaches that include interpretive and hermeneutic methods to explain and understand phenomena.

Different research methods would be employed in traditional and nontraditional approaches, and yet both approaches are valuable in generating a full range of knowledge in the social sciences. These differences raise a number of important questions, including those involving (a) the differences between what we can know in the natural sciences versus what we can know in the social sciences; (b) whether there can be laws that govern human behavior, much like the laws that govern physical phenomena; and (c) the nature of causality. It is these questions that this chapter will begin to address.

The purpose of this chapter is to provide a brief introduction to terminology and concepts associated with the philosophy of science that are most directly applicable to your research. From a purely practical standpoint, it might seem that philosophy does not have anything to do with conducting research. For example, you could imagine conducting a study, collecting and analyzing data, and reporting on those data without considering the philosophical or theoretical aspects of the underlying science. However, as will become clearer in this chapter, your choices of theory, research approaches, and research designs come with sets of assumptions about the nature of truth. More practically for you as researcher, there needs to be a logical connection between philosophy, theory, and method, or what we frequently refer to as *alignment* in this text. Miller and Burbank (1994) described that

> Philosophy, theory, and method are interrelated and fit together like pieces in a puzzle. . . . One's philosophical perspectives affect the way one chooses a theoretical perspective. Similarly, the underlying assumptions on which theories are based may limit the research methods. (p. 704)

Our focus in this chapter is the congruence between philosophy, theory, and method. We also emphasize why this congruence is important and provide you with strategies to achieve that congruence.

We begin with a review of two concepts in philosophy, ontology and epistemology. Next, we review the major philosophical perspectives, referred to in Chapter 1 as *paradigms*, found in social and health science research. Paradigms, which we commonly refer to as *philosophical orientations* in this chapter, refer to the constellation of assumptions about and orientations to reality that guide you to adopt particular approaches to inquiry and your methods. We will review the philosophical perspectives including logical positivism, sometimes just called positivism, postpositivism, and constructivism. Next, we present the essentials of theory development to orient you to the role of theory in research. We conclude the chapter with key questions that should guide your thinking regarding underlying assumptions and the role of theory in research.

Ontology

Ponterotto (2005) provided a definition of ontology as "the nature of reality and being. More specifically, ontology addresses the following question: What is the form and nature of reality, and what can be known about that reality?" (p. 126). Ontology addresses the underlying question of whether there is an objective, verifiable reality outside of the researcher, which is a position associated with *realism*, or whether reality is the result of individual interpretation or social construction, a position associated with *relativism*. Although there are many varieties of realism, for this discussion, realists believe that we perceive objects outside of ourselves, objects whose existence and nature are independent of our perception of them. Realists believe that there is one truth, even though we may never be able to know it. Truth, in this sense, corresponds directly to facts that are observable and knowable (Kukla, 2000).

Antirealism is the opposite of realism and includes many different philosophical positions such as pragmatism, idealism, rationalism, and relativism. Relativism will be considered here because it is an ontological position closely aligned with several nontraditional, interpretive methods such as hermeneutic phenomenology. Relativism was described by Clark (1992) as knowledge that is always a representation of reality from a person's particular perspective. Thus, there can be no objective point from which to evaluate the truth of our view outside of our own perspective of the world. Knowledge is determined by contextual circumstances, including historical, subjective, cultural, or institutional (Kim, 1999).

Epistemology and Ways of Knowing the World

Knowledge. Before beginning an examination of the philosophical and theoretical issues related to a philosophy of science, it is good to pause and reflect on how we come to generate knowledge about and understand the world. Knowledge is defined as belief that has been justified based on facts assumed to be true by an observer. There are multiple ways knowledge can be generated. The first is through perception. We know things because we have experienced them with our senses. Experiential knowledge associated with sensory understanding is called *empirical knowledge*.

Another way of knowing is through reasoning. Knowledge is produced when logical and reasoned analysis, either inductive or deductive, is used. For example, professors and employers alike value critical thinking skills, but one does not need to be a scientist to invoke critical reasoning to solve problems. In classes, you regularly are asked to support your ideas with credible evidence. You might review a number of sources and look for convergence of evidence that lends support to your opinion. You are using critical thinking skills but not necessarily invoking the scientific method.

A third type of knowledge is memory. Many of us use memory of situations to arrive at conclusions (knowledge) about what is happening at any given point in time. Introspection, or reflection on the experience of our own mind, is yet another way of gaining knowledge that may or may not involve inductive or deductive reasoning and logical thought. Other ways of knowing include intuition, telepathy, or clairvoyance. For example, people sometimes use hunches to come to understand their world, thereby reaching conclusions about the world that may be based partly or not at all in fact.

Johnson and Duberley (2000) wrote, "Everyone adheres to some theory about what constitutes warranted knowledge—a set of epistemological commitments which provide us with criteria for distinguishing between reliable and unreliable knowledge" (p. 5). Later, they wrote that epistemology provides answers to questions that include "(a) What are the origins, nature, and limits of scientific knowledge? (b) What constitutes scientific practice? and (c) What are the processes through which scientific knowledge advances?" (p. 5). Houghton, Hunter, and Meskell (2012) defined epistemology as "the relationship between researcher and what can be known" (p. 34), which highlights the symbiotic role of researcher and knowledge.

Epistemology concerns knowledge. It is the study of knowledge and guides us to ask questions about what we can know and how we can know it; it also includes questions that explore the limits of knowledge. From the perspective of science, epistemology concerns what can be known using a scientific approach to understanding the natural world. For example, what can science and the scientific method reveal? Is what is revealed the *truth*? And is there one *Truth*, or are there multiple and relative *truths*? Epistemological discussions also include whether the researcher can be a truly objective observer of reality or whether knowledge is really generated through dynamic interpretation of phenomena.

Mittwede (2012) also described epistemology in a somewhat different way as the connection between the "investigator and the investigated" (p. 26). Some have believed that this relationship could be viewed as completely objective, with the researcher viewing the object of research or participant objectively in an unbiased, value-free way. Thus, epistemology is concerned with the nature of scientific knowledge and the limits to which science will add new knowledge. Also critical is the examination of our relationship to that knowledge and whether an objective reality or truth actually can exist.

We agree with Hanson (1958) who argued that observation is theory laden and unbiased observation is not possible because we are human beings with biases that cannot be completely overcome. This position is often called *modified objectivity*. Other researchers who use nontraditional or interpretive science approaches believe that the best and sometimes only way to generate knowledge about human beings is through a subjective relationship between the researcher and the participants who together cocreate knowledge. In this perspective, objectivity is not possible, and the lack of objectivity is fundamental to the process of knowledge generation.

Facts. The scientific method is based on empirical evidence. Scientists discover and confirm facts, which are based on what can be observed and measured by multiple people. That the sun shines is a fact because it is directly observable and can be verified by multiple bystanders. However, facts are not necessarily stationary; tomorrow, the fact may be that the sun is not shining because you cannot see the sun and it is raining (although we all know the sun still exists behind the clouds!). Facts also result from immutable laws of nature. For example, sodium and chloride combine in predetermined proportions to make salt, which we accept as fact because an experiment repeated multiple times and in multiple contexts achieves the same result.

Facts can change based on new knowledge and new understanding (e.g., it was accepted as fact for centuries that the sun revolved around a stationary earth until scientific evidence showed otherwise). As another example, for many years, scientists believed that gastric ulcers were caused by stress and excess stomach acid. Although Warren and Marshall's (1983) work identified a bacterium as the cause of gastric ulcers, the medical community did not accept these findings for many years. Continued scientific research and discovery ultimately provided undisputable facts, and the accumulation of facts ultimately led to shared understanding of a bacterial cause of ulcers. The two scientists were awarded the Nobel Prize in 2005 for their important work. Warren and Marshall's work provides an example that scientific fact is not always immediately established and requires debate and verification.

Therefore, epistemology in the philosophy of science refers to scientific knowledge and what constitutes scientific knowledge, how knowledge is generated from the practice of the scientific method, and the nature of the relationship between researcher and what is being researched. Epistemology is about what we accept as fact. When we are using the scientific method, we are stating first that we are empiricists and that we are basing what we know (or what we have come to know) on facts that are observable and verifiable by others. We systematically observe a phenomenon, assemble the facts, and use critical reasoning to arrive at understanding (knowledge) of the phenomenon. Ontology concerns the nature of reality and the perception of what is *truth*. Thus, as researchers, we are interested in ontological and epistemological considerations because they are important considerations for understanding the nature of truth and reality as well as how knowledge is generated. How you understand ontology and epistemology orients you toward certain approaches to research; this will become clearer in the discussion of paradigms/philosophical orientations.

Universal and Individual Perspectives

Another axis on which the researcher orientation can be described is sometimes called the etic-emic distinction (Berry, 1989; Ponterotto, 2005). In cross-cultural psychology, the etic-emic distinction has been used to describe the focus of research as being the search for phenomena that are universal and transcend culture (etic) or for phenomena that are unique and culturally specific (emic). However, the distinction has more broad application in terms of the tension between the search for universal laws and the focus on individual uniqueness. This distinction drives the underlying objectives of quantitative and qualitative research. Whereas the etic perspective searches for generalizations, the emic perspective focuses on understanding individual units such as persons, organizations, and cultures with no intent to generalize.

Philosophical Orientations/Paradigms. It is important to reference a philosophical orientation when you create a research proposal because it can "help create a bridge between the aims of a study and the methods by which to achieve those aims" (Denzin, Guba, and Lincoln, 1994). Your specific philosophical orientation also plays a role in predisposing you to particular forms of research questions that involve particular kinds of research methods. A paradigm is the collection of facts, assumptions, and practices that guide a particular orientation to knowledge generation. The ones we review here include positivist, postpositivist, and constructivist perspectives. Guba and Lincoln (1994) described these positions as "a set of basic beliefs (or metaphysics) that deal with ultimates or first principles. It represents a worldview that defines, for its holder, the nature of the 'world,' the individual's place in it, and the range of possible relationships to that world and its parts" (p. 107). Lincoln, Lynham, and Guba (2011) also included critical theory as one of the philosophical orientations parallel to the other three. Although important, our view is that it is usually not considered a philosophy of science orientation but rather a theoretical perspective from the relativist position. We will come back to critical perspectives briefly later in the chapter.

Positivism/Logical Positivism

Positivism emerged primarily from the successes of knowledge generation in the natural sciences. Positivism has its roots in the scientific revolution that occurred during the 15th through 17th centuries and the Enlightenment period that extended from the late 17th into the early 19th centuries. During the period of the scientific revolution, the natural sciences were flourishing. Nicolaus Copernicus (1473–1543) had used systematic observation to show that the earth revolved around the sun and not the other way around, although this discovery was not accepted as truth at the time. Isaac Newton (1642–1726) invented the calculus and discovered laws governing the physical world. The advances in knowledge, principally due to the success in the natural sciences, influenced philosophers during the enlightenment. During this time, philosophers viewed knowledge as originating from thinking (rationalism) and observation (empiricism) (Johnson & Duberley, 2000, p. 13). Many consider John Locke (1632–1704) to be the father of modern empiricism. He described that data collected via the senses are internalized, reflected upon, and form the basis of ideas that grow in increasing complexity with the addition of new sense data. In essence, he created a logical argument for the role of experience in the production of knowledge. Scientists examine data, as perceived by the senses, and inductively generate laws based on those data.

Auguste Comte (1798–1857) was the first philosopher to use the term *positivism*. He used the term *positive* to distinguish scientific knowledge from fictitious knowledge, associated with religious teachings that existed during his era, and abstract knowledge associated with metaphysics (Johnson & Duberley, 2000). For Comte, positive knowledge was generated from facts that were derived from sense perceptions; anything that was not directly attributable to the senses was relegated to the metaphysical realm and thus not valid (Mittwede, 2012). Comte believed that just as laws govern the natural world, there are universal laws that would be discovered that govern social behavior. What emerged from Enlightenment thinking is a view of a science that seeks universal laws through objective observation, description, explanation, prediction, and control of natural phenomena.

This view of science was extended to the emerging social sciences in the 19th century. For example, the science of psychology, born in the German university system, facilitated the split between psychology and religion. Early German psychologists focused on directly measurable and observable aspects of behavior. The scientific method, proving to be highly successful in the natural sciences, was replicated to explain human behavior and cognition. John Stuart Mill (1806–1873) formalized a set of procedures that would form the basis of the scientific method in which the scientific experiment was promoted as the optimal means for controlling variables and identifying causal mechanisms. Positivism, based in the methods of the natural sciences, thus became the foundation for knowledge generation in the emerging social sciences.

Logical positivism emerged from a group of philosophers who created a positivist philosophy of science by formalizing the language of theories. Logical positivism has also been referred to as the *received view* to recognize the extended influence of positivism on scientific research. The language of theories included theoretical vocabulary and strict rules of logic. Johnson and Duberley (2000) described the four epistemological commitments of logical positivism. The four commitments include the following:

- Observations of the world, through our senses, provide the sole foundation for knowledge. Following Comte, those observations could be made in a very neutral, value-free manner.
- What is not observable or is unconscious cannot be included in the realm of scientific knowledge. Anything that is tested empirically must be able to be verified.
- Methods used in the natural sciences provide the gold standard for scientific knowledge generation.
- The goal of science is prediction and control. (pp. 23–27)

Each of these commitments has clear roots in positivism; what distinguishes positivism from logical positivism is that in the latter, all data must be observable and verifiable. This distinction provided the clear demarcation between what can be considered scientific and what belongs in the metaphysical realm. Logical positivism is also known for its belief that explanations for all phenomena (natural, social, and behavioral) could be made using the same language of logic. The approach of the logical positivists is very reductionist, much in the same way that physical laws reduce explanations to the mechanisms that occur at the atomic level.

Postpositivism

As knowledge in the disciplines of physics and the other physical and mathematical sciences rapidly progressed into the early 20th century and greatly enhanced our understanding of the world, philosophers of science began to see that logical positivism was not a useful or accurate way of describing scientific theories. For example, it became clear among philosophers of science that there was no way to prove empirical claims to be universally true. Karl Popper (1902–1994) and Carl Hempel (1905–1997) are widely considered as the principal originators of postpositivism. Popper was one of the key opponents of the view that verification of theory is the defining feature of science. Rather, in his critical rationalist philosophy, he believed that researchers are inherently human and thus fallible. For Popper, *falsifiability*, the ability of a theory to be shown to be false, was a much more important defining feature of science. Popper also realized that metaphysical knowledge, defined as knowledge that cannot be seen, sometimes spurs scientific discovery. Such knowledge, for example, can come through hunches that lead to hypotheses that can be tested. Thus, he found the complete rejection of metaphysical knowledge by the logical positivists to be overstated. Two other postpositivist philosophers of science, Polanyi (1958) and Hanson (1958), effectively disputed the idea of value-free observations, arguing that all observations are biased in some way by the observer and his or her values and past experience.

Constructivism

Constructivism is a particular relativist perspective that has its roots in the philosophy of Immanuel Kant (1724–1804), who posited that reality, or the external world, is shaped by our experiences. We cannot experience reality directly; instead, we experience phenomena that are then interpreted by our senses. A relativist perspective holds that there is not an external, verifiable reality outside of the observer and, because of this, there can be no value-free objective observations on the part of the researcher. There can also be no shared truth among researcher and participant. Constructivism departs significantly from positivism in that, as a relativist perspective, it does not assume the existence of any single *true* reality. Meaning and knowledge are constructed through the interactions of individuals, and it is through these interactions that shared meanings and *truths* are created. What is truth in one context may not be truth in a different context, and all realities are equally valid. Meaning is typically hidden and requires reflection through shared experiences to be discovered. From a research standpoint,

> a distinguishing characteristic of constructivism is the centrality of the interaction between the investigator and the object of investigation. Only through this interaction can deeper meaning be uncovered. The researcher and her or his participants jointly create [coconstruct] findings from their interactive dialogue and interpretation. (Ponterotto, 2005, p. 129)

Guba (1990, p. 25) identified the four points where constructivism is at odds with positivist and postpositivist views. The first is that from a constructivist perspective, facts only make sense in the context of a given theory or value proposition, which means that there can be no single reality. The second is that theories can never be fully tested; therefore, there will always be competing theories, and no one theory will ever be found that explains the facts completely. Third, inquiry can never be value free. Finally, objectivity is not possible because the observer's interpretation of what is observed will always be shaped by the value and theoretical systems of the observer. A constructivist approach reflects a position of relativism, described by Guba in this way: "The key to openness and the continuing search for ever more informed and sophisticated constructions. Realities are multiple, and they exist in people's minds" (p. 26). Gershenson (2013) summarized one consequence of constructivism in the following way: "Ontological statements, i.e. about things as they 'really' are, independent of an observer, are fictitious" (p. 782). Thus, relativist approaches question the plausibility of a single, objective reality or truth. The richness of relativist perspectives lies in the multiple interpretations of the experiences in which humans engage.

In an even more extreme type of constructivism, physical reality is seen as *caused* by consciousness (Harman, 1991). There are some researchers who challenge a *matter over mind* position; these researchers tend to represent those interested in Eastern spiritual traditions. One example is the power of meditative states and how meditation is seen as a causal mechanism for changes in physical states (self-healing). Some popular press authors who espouse the practice of intentionality are applying this philosophical perspective of constructivism. They propose that one may control events in the physical world with thoughts, affirmations, and/or intentions (Byrne, 2010; Day, 2010; Hay, 1984; Hicks & Hicks, 2006; McTaggart, 2003).

Each of these three paradigms—logical positivism, postpositivism, and constructivism—is important and can be described in terms of their essential ontological and epistemological positions. The positions not only represent an evolution of scientific understanding, but also reflect the various positions scientists take when thinking about the nature of reality. Understanding paradigmatic, ontological, and epistemological positions provides a path for clearly communicating your assumptions concerning your views of the fundamental nature of reality and the nature of knowledge generation. There are many fine tables that provide clear presentation of these positions (e.g., see Guba, 1990; Lincoln et al., 2011, p. 98). Figure 2.1 is an adaptation and simplification of information provided by Lincoln et al. (2011, p. 98). We show here only two paradigms representing extreme positions; note that your particular philosophical orientation may lie somewhere between the two.

Each of the three paradigms is also closely related to the research approaches to inquiry (qualitative, quantitative, and mixed methods). For example, most quantitative research tends to be more aligned with positivism and postpositivism, whereas qualitative research tends to be aligned more with constructivism. These are not hard and fast distinctions, however. For example, a qualitative researcher may code the data with the intent to reduce the data to bits that can be analyzed quantitatively; this approach would be more aligned with a positivist epistemology and ontology.

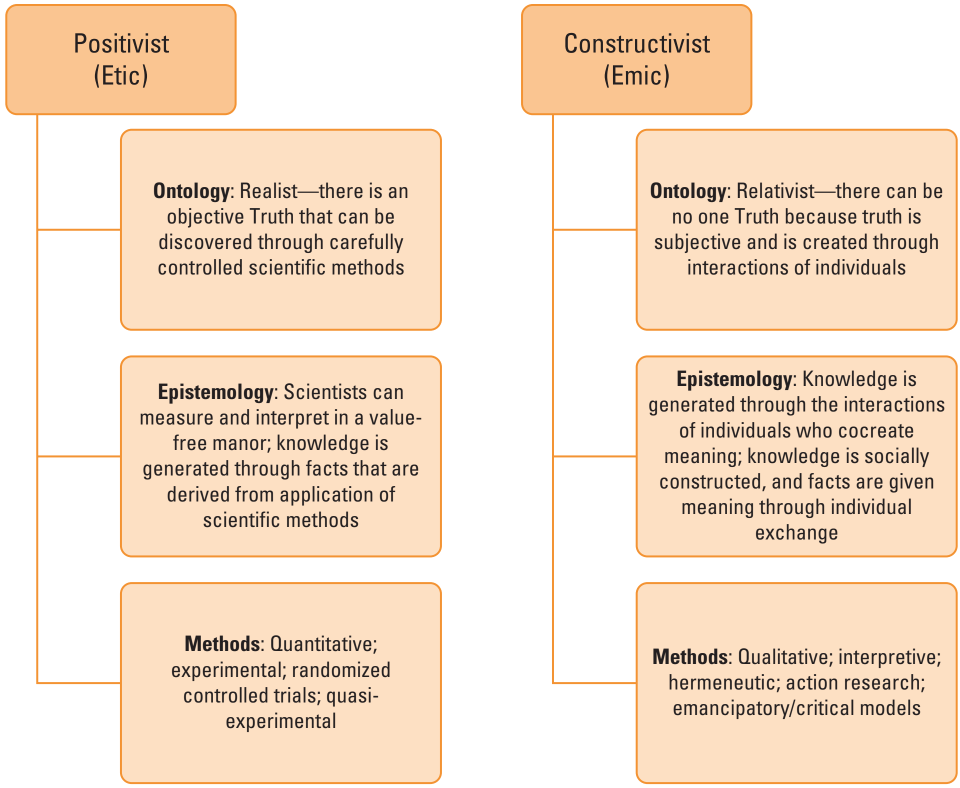


FIGURE 2.1 **Positivist–constructivist ontology, epistemology, and methods.**

Thus, there are three things to consider related to ontology and epistemology that require you to answer some important questions. These questions reflect the way you see the world and influence your natural approach to research. The first consideration is your own ontological orientation. For example, do you believe that there is a reality outside of yourself that is independent of your perception of it, or do you believe that there are multiple truths that depend upon individual interpretations of phenomenon? The second consideration is your own epistemological orientation. For example, does the generation of knowledge happen through a neutral, objective relationship between researcher and the object of research, is objectivity a goal but not truly achievable, or is knowledge discovered or created through the relationship between researcher and the participants you are researching? The third consideration is related to your goals for the research. Are you interested in a distant relationship between yourself and the object of research (the participant), or is your goal knowledge generation through interaction with participants in the research? Each of us has a natural orientation to one or more of these positions, and it is important that you read and reflect upon and discover your orientation.

Once these questions have been addressed, the next step is to explore the literature in more detail on these different perspectives. Which perspectives have been presented in the research in your area of interest? Is there room for other perspectives that challenge current thinking and potentially add new knowledge in the context of that challenge? What theoretical perspectives influence ontology and epistemology of your discipline? The pursuit and exploration of answers to these questions serve to enrich the quality of your research.

Critical Perspectives

Although we opted to not include critical perspectives as one of the paradigms, this perspective is a very important one that deserves some attention. Critical perspectives derive from Marxist and neo-Marxist theories that examine the ways in which "relationships of domination and exploitation are embedded in the dominant ideas of society" (Burbank & Martins, 2010, p. 30). At the heart of critical theory is acknowledgment of the nature of oppression and how subordinate acceptance of their social status continues to reinforce oppression (Kincheloe, McLaren, & Steinberg, 2011).

In their essence, critical theory and critical perspectives seek to understand the nature of oppression and seek the emancipation of oppressed peoples. Emancipation occurs through the process of research with coresearchers (typical of a participatory action framework; see Hacker, 2013), the critical evaluation of the contexts in which the research is situated, and using the results of the research to make change. Thus, critical theory is constructivist in that there can be no one truth and the relationship between the researcher and what is researched is constantly in flux. Various methods can be used to answer research questions; thus, critical perspectives can be viewed also as being very pragmatic. Critical theorists have established lines of scholarship in several important areas, including feminist theory (Gannon & Davies, 2007; Lather, 1994); queer theory (Plummer, 2011); research on indigenous peoples (Smith, 1999); and the intersections of race, gender, and sexuality (Bowleg, 2012), among many others.

Figure 2.2 provides a guide to help you understand some of the key relationships among the goal, philosophy, and approach to inquiry of research.

Research goal, philosophy, and approach		
Predictive	**Interpretive**	**Emancipatory**
Goal: Prediction **Philosophy**: Positivist **Approach**: Primarily quantitative	**Goal**: Understanding **Philosophy**: Constructivist and interpretive **Approach**: Primarily qualitative	**Goal**: Emancipation **Philosophy**: Constructivist **Approach**: Qualitative or quantitative, focus on critical perspective

FIGURE 2.2 **Relationship between research goal, philosophy, and approach.**

The goals of research as predictive, interpretive, and emancipatory are consistent with the writings of Habermas (1968/1971).

Next, we turn our attention to the role of theory in research. We begin the discussion with a broad understanding of what theories are and how they are constructed. Following the discussion of theories, there are definitions of key components of theories: concepts, constructs, and hypotheses. Next, we include a brief overview of the role of causality in social science research and its connection to theory. Finally, we close with a presentation of the concept of a continuum of inquiry that brings together the discussion of philosophical perspectives/paradigms and theory in a practical way.

The Role of Theory in Research

What Is Theory?

Reynolds (2007) described four roles for scientific knowledge. These are "(1) a method of organizing and categorizing 'things,' a typology; (2) predictions of future events; (3) explanations of past events; and (4) a sense of understanding about what causes events" (p. 2). Reynolds conveyed that each of these four roles must contribute significantly to how theory serves as an organizing principle of knowledge and the expected relationships between categories of objects.

The term *theory* is, unfortunately, used differently in different contexts. Part of the reason for this misuse is that, from a philosophy-of-science standpoint, the definition of theory and its implications has evolved over time, and even philosophers of science do not necessarily agree on the definition (refer to the previous discussion regarding positivism and postpositivism). To complicate matters, the word *theory* is used carelessly by new scientists as well as by the general public; the terms *theory* and *hypothesis* are frequently used synonymously. We sometimes hear a new student espousing a *theory* when describing his or her research interest. For example, a student might say he has a theory that "people in the workplace would respond better if they were offered more vacation time." But this is not a theory; rather, it is a speculation that needs to be tested empirically via the scientific approach. The *Oxford English Dictionary* defines *theory* (2006) as follows:

> A scheme or system of ideas or statements held as an explanation or account of a group of facts or phenomena; a hypothesis that has been confirmed or established by observation or experiment, and is propounded or accepted as accounting for the known facts; a statement of what are held to be the general laws, principles, or causes of something known or observed. (para. 1)

Wilkins (2006) provided the following definition: "Currently, a theory is regarded in the philosophy of science as a conceptual entity that plays an explanatory role in a science" (p. 2185). Both definitions are important in that they convey that a theory is an entity comprising multiple statements that describe concepts associated with a particular phenomenon and relationships among them. Thus, a theory is more than a single statement to be supported or proven false. We will adopt the definition of theory as *a set of concepts and relational statements that organize scientific knowledge in a* *focused way*. In addition, we will focus on what Reynolds (2007, p. 98) called the causal process form of theories. The causal process form focuses on the potential causal nature of relational statements comprising theory. In the following sections, we describe each of these components of the definition in more detail.

Theories typically have four purposes that include describing, explaining, predicting, and controlling or changing phenomena. These four purposes are hierarchical, with each subsequent purpose including the previous one. At the descriptive level, theories simply describe concepts regarding phenomena and the relationships between those concepts. Thus, this description serves to increase our understanding about a particular phenomenon. Note, however, that description also can serve as explanation in that the way we describe a phenomenon orients the reader to particular kinds of explanations (Geertz, 1973). Theories that *explain* answer the question regarding why a phenomenon occurs. Theories that explain typically refer to occurrences in the past, whereas theories focused on *prediction* build on explanation by focusing on future events. If a theory works well to explain something that happened in the past, it should be able to predict what will happen if the same situation occurs in the future. Lastly, theories can be applied to *change or control situations* such as a disease state or other undesirable outcome. Some theoreticians, however, do not like the word *control* and have chosen other terms to reflect prediction and control. For example, Kim (1993) uses *prescribe* in place of prediction and control.

Theories may be used in research in two ways. First, theories are used as a guide for the research process, including selecting research questions and interpreting findings. Second, the research may be a test of the theory itself, measuring the concepts and testing the relationships between the concepts to determine the adequacy of the theory.

Concepts. Concepts are representations of things that exist in reality and can be specific or vague, abstract or concrete. For example, *dog* would be a very specific, concrete concept; it is a representation that is widely agreed upon. Concepts that are more abstract may be agreed upon but lack specificity. For example, weight is an abstract concept that is independent of time and location (Reynolds, 2007). If one were to change this to read, *weight of the chair*, this would become a concrete concept. A *construct* is a type of concept that is theoretical in nature, agreed upon in terms of what it generally conveys, but cannot be directly observed. The social sciences contain many examples of constructs, including IQ, attitudes, and values. Using operational definitions, scientists measure these constructs even though they cannot see them.

Theoretical and Operational Definitions. Concepts have theoretical and operational definitions. Theoretical definitions are descriptions in words of the meaning of the concept. These may be different for the same concept used in different theories. One example is intelligence. Traditional measures of intelligence, developed by researchers such as David Wechsler, focus on quantitative and verbal reasoning that tend to be highly related to academic achievement. Other researchers have proposed other theoretical definitions of intelligence. For example, David Gardner has advanced the theory of multiple intelligences, and Peter Salovey has developed a theory around emotional intelligence. Therefore, theoretical definitions must be linked to operational definitions. To measure theoretical concepts (and constructs), scientists must determine how they will be defined and measured. Although all scientists may not agree on how a concept is theoretically or operationally defined, you must be very clear in stating how the concept is defined in your immediate study.

To be able to measure a concept, you must create an operational definition. Reynolds (2007) defined operational definitions as "a set of procedures that describes the activities an observer should perform in order to receive sensory impressions (sounds, visual or tactile impressions, etc.) that indicate the existence, or degree of existence, of a theoretical concept" (p. 52). For concepts such as temperature, the operational definition is very clear: It includes use of a thermometer that has been appropriately calibrated and tested. There is little choice or deliberation in the matter over how you might define and measure temperature.

For other kinds of concepts and constructs used in the social sciences, operational definitions typically require some kind of assessment that can be either self-reported or administered by a tester. In our example of intelligence, the operational definition is the particular scale that is used to assess it. For example, the Stanford-Binet Intelligence Scales is one instrument used to assess traditional intelligence, and the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) is used to assess emotional intelligence. Different theorists have different theoretical definitions of intelligence and thus operationalize it differently using different scales. This applies as well in qualitative research. A qualitative education researcher may be interested in measuring motivation; the raise of a hand in class may be one operational definition of motivation. What you should take away from this discussion is that every concept that is the focus of a research study must have a theoretical definition that is accompanied by its appropriate operational definitions. Theoretical and operational definitions communicate clearly to other researchers the choices that guide your study.

Relational Statements. Relational statements describe the relationship between two or more concepts. Relational statements form the basis for theories because theories consist of a number of relational statements that indicate how concepts are (or are hypothesized to be) related to one another. Reynolds (2007) identified two kinds of relational statements. Associational statements describe the concepts or constructs that occur (or do not occur) together. A simple example of an associational statement is the correlation statement, with which you may be familiar. For example, *lower fat diets are associated with lower risk for cardiovascular disease* is an associative statement that represents a positive correlation between dietary fat and risk for disease. The second kind of relational statement is the causal statement. The correlational statement can be converted to a causal statement: *If men and women in the United States eat lower levels of fat in their diet, this will cause lower risk for cardiovascular disease*. In a causal statement, there is a clear independent variable (dietary fat) that is manipulated to cause a change in the dependent variable (risk for cardiovascular disease). It is important to not confuse correlation and causation; we will return to this point later in the discussion.

Hypothesis. A hypothesis is a special kind of relational statement that provides a conjecture of the relationship between two or more variables that can be directly tested empirically. Hypotheses are best guesses about the relationships among the concepts of a theory. A review of the literature supports the potential relationship between the concepts, but there is no empirical evidence for the relationship yet. The researcher's goal is to test the hypothesis and provide evidence that supports or refutes it. In the ideal case, hypotheses that are derived from existing theory serve as further tests of that theory. It is very important to note that the results of a test of a hypothesis do not *prove* or *disprove* the larger theory. Rather, they provide evidence that supports and continues to strengthen the theory or that does not support the theory (see Kuhn, 2012, for an excellent discussion on the role of confirmation and disconfirmation of theory).

Reynolds (2007) provided a way to think about the relationship among relational statements and empirical support. "Those with no support are considered hypotheses, those with some support are considered empirical generalizations, and those with 'overwhelming' support are considered laws" (p. 88). In the social sciences, there are probably very few *laws* because the variability of human behavior requires extensive testing over multiple contexts and time periods.

Model. Sometimes, researchers use the terms *theory*, *theoretical model*, and *theoretical framework* synonymously. Theoretical models are visual representations that demonstrate how a subset of concepts comprising the theory is hypothesized to relate. A model can (but may not) include all concepts of the theory. An example is the health belief model. In public health research, the health belief model has been used extensively to explain people's health-seeking behavior. The model links concepts together, such as external environmental cues and perceived health risk. This model is a representation of concepts that have their roots in field theory originally conceptualized by Kurt Lewin (1890–1947).

Evaluating Theories

Theory evaluation is the process of systematically examining a theory using a set of criteria. These criteria vary by author. McEwen (2011) reviewed the nursing literature and identified the criteria most often used by researchers; although the review was focused on nursing literature, the criteria have broad generalizability across disciplines. McEwen identified the most widely used criteria internal to the theory as well as the most used external criteria, which involve examination of how well a theory works in the world. There are four key internal criteria: clarity, consistency, logical adequacy, and simplicity. Theories should be *clear* in terms definitions of their major concepts with the relationship among concepts clearly specified. Theories should be internally *consistent* in terms of their major suppositions, logical connections among concepts, and philosophical underpinnings. *Logical adequacy* refers to the propositions in the theory being plausible and creating a coherent structure for the theory. Many argue for *simplicity*, or parsimony, in that the simplest theories that maximally explain the phenomena should be chosen over more complex theories. However, Dudley-Brown (1997) also suggested that theories might be necessarily complex, depending on the phenomenon of interest.

Major external criteria include scope or generality, testability and empirical adequacy, fruitfulness, and usefulness or significance. Theories should have *scope or generality* appropriate to their level of development. Whereas broad scope would be expected of more advanced middle and higher range theories, which are characterized by more fully developed relationships among concepts that have been empirically tested, narrow scope may characterize newer theories. A theory that is *testable* has qualities that enable it to be tested empirically to determine if it can be supported or not. A theory is *empirically adequate* if the results of the studies conducted to test the theory support it. *Fruitful* theories are those that provide a rich set of concepts that generate multiple relational statements as well as hypotheses for testing. Fruitful theories can thus promote generation of new theories that yield different hypotheses to test. *Useful* theories, or those that have significance within or across disciplines, are those that can be used to address phenomena of interest or problems within the discipline. Useful theories enjoy increasingly broad consensus and acceptability.

Reynolds (2007) takes a slightly different position and advocates that theories should not necessarily be compared. He offers four reasons for this (p. 136). First, because science is not focused on searching for a unitary truth, multiple theories are useful to capture more accurate descriptions of phenomena. Second, theories will never be able to be totally rejected by a single study, and in reality, multiple studies are required over time and across contexts to eventually support or discount a theory. Third, theories may describe processes that impact variables in particular contexts, and these processes may be described differently in different theories. It makes more sense, in his view, to test the extent of influence each process has rather than trying to make a decision about which one is best or correct. Competing theories in this case make sense. Fourth, theories generally cannot be directly compared. Competing theories likely arise because the ones that are in existence fail to account for certain relational statements; thus, comparing them directly would be challenging.

Regardless of the stance you choose, you should understand the theories that guide a specific phenomenon through careful examination in terms of their concepts and relational statements. It may be that your study is able to test relational statements as hypotheses that would be consistent with two different theories. In this case, the goal would not be to compare and contrast, but rather to understand the assumptions and limitations of each of the competing theories.

Choosing a Theory to Guide Research

Much research appears to be atheoretical in that studies are often published without any reference to a theory guiding the research. The results seem just as valuable and important whether or not there is a theory, so an important question arises in terms of why theory matters. Even though a research study may not explicitly state a theory, the author of the study always has a particular view of the world. Often, the researcher's worldview can be inferred from the type of research study. The study author also has a particular set of beliefs about the concepts and/or variables and relationships between them that comprise theory. It is very important to make these theoretical connections explicit in your own research. The connections, when discussed in relation to the existing research in the discipline, provide greater depth of understanding to bridge the relationships between the variables. Acknowledging the theoretical and philosophical connections helps other scientists interpret the results of the study as well as the limitations of those results.

When you begin to explore an area and conduct a literature review, theories that others have used typically become apparent. Selecting a theory can seem like a daunting task. As you read more about the theories relevant to your topic area in the discipline, you will be drawn more to some theories than to others. The theories to which you are drawn are most likely theories that fit best with your worldview. Before selecting one particular theory to guide your research, examine the underlying assumptions of the theory to ensure that they do fit with your philosophical perspectives. One question that is useful in examining assumptions is to ask yourself, "What do I have to believe about the world and about human beings in order for me to accept or use this theory?" For example, if you are choosing a theory of human behavior and you hold a basic belief that human beings are unique and active creators of their own actions, then your assumptions are not congruent with those of the theory of behaviorism. Behaviorism posits behavior as a response to external stimuli. Thus, behaviorism would not be an obvious choice for the underlying theoretical and philosophical perspective for your research. However, if you take the position that human behavior is the result of responses to environmental stimuli, as posited by B. F. Skinner, for example, the choice of behaviorism as a theoretical perspective would yield congruence of theory and fundamental assumptions.

Cause and Effect

One final topic needs to be addressed: that of cause and effect. This is another principle that is frequently misused. Correlation, as described earlier, does not necessarily imply causation. In correlational relationships, one can have a positive relationship, negative relationship, or no relationship. In the case of a correlation relationship that is positive, the best you can conclude, in the absence of any other evidence, is that when the value of one variable is high (or low), so is the other. You cannot determine which variable is causing the relationship. It may be that neither variable is the cause, but the cause is due to another variable that was not tested. One example of what is commonly described as the *third variable problem* is the correlation between bars and churches. In general, when the number of bars in a community is high, so is the number of churches (it would seem rather silly that opening bars would somehow cause more churches to open). In this example, one indeed does not cause the other; rather, the increase in each is directly influenced by population growth.

Bullock, Harlow, and Mulaik (1994) provided several criteria required to establish causation. The first is that the two variables must be associated (correlated). The second is that the variables must be isolated so that other extraneous variables are removed as possible causes for the covariation (correlations). The third is that the variable hypothesized to be the causal variable must precede the other in time.

In the laboratory, these factors related to cause and effect can be imposed by the research design. For example, we know that when we heat a gas, the heat causes it to expand in a very predictable way. We also know that the expanding gas does not cause the temperature to rise. We know with certainty that increasing temperature causes gases to expand. Determining causal relationships in social behavior is much more challenging. The influences on behavior are so great and varied that controlling for all possible effects is virtually impossible, even in a laboratory setting. Even if we could control all of the variables, measuring the factors hypothesized to cause behavior is a daunting task. Measuring behavior is at best imprecise. There are also limits to control of variables based on ethical principles—we cannot impose the constraints we would need to in order to truly understand the impact of some concepts on behavior. For example, a researcher could not create an experimental group of nonsmokers and ask them to begin smoking in order to examine causal effects of smoking on depression. People also tend to behave differently when they know they are being observed, so it is not typically possible to know if the precise, hypothesized cause of the behavior is being measured.

Each of these challenges limits the degree to which we can assert causation in almost every situation involving human behavior. However, the programs of research in which we engage can help scientists make progress in the search for causal relationships. First, longitudinal studies, those that assess people's behavior at multiple time points, can allow researchers the opportunity to determine sequencing in time, a critical requirement for causality. Second, replicating studies across multiple contexts and times can help solidify evidence that a particular relationship is durable and likely to be causal. One example of how evidence over time leads to causal assumptions is the extensive evidence of the correlational relationship between smoking and lung cancer. It is not ethical to put nonsmokers into two groups and ask one group to start smoking, and thus, it is difficult to provide the experimental controls needed to ascertain causality. But the cumulative evidence of the association between tobacco use and lung cancer, controlling for the potential effects of other variables, has been overwhelming. Thus, correlation studies test for the extent to which variables change in a coordinated fashion and typically in naturalistic settings without determination of why the variables behave as they do. Cause-and-effect studies focus on determining which variable is causing the change.

The Research Continuum

In this chapter, we have described some of the epistemological and ontological underpinnings of the key research paradigms. The discussion of these underpinnings was followed with an explanation of the role of theory in knowledge production and research. In this section, we bring the two together into the research continuum. The research continuum (Figure 2.3; L. Milanesi, personal communication, April 26, 2015) provides a visual way to understand how paradigm, approach to inquiry, theory, and method come together to influence the overall research proposal.

The far left of the continuum represents an area in which the state of research is undeveloped. This may be a new area of inquiry, or it may be an area of inquiry that has been established but has not been extensively tested across

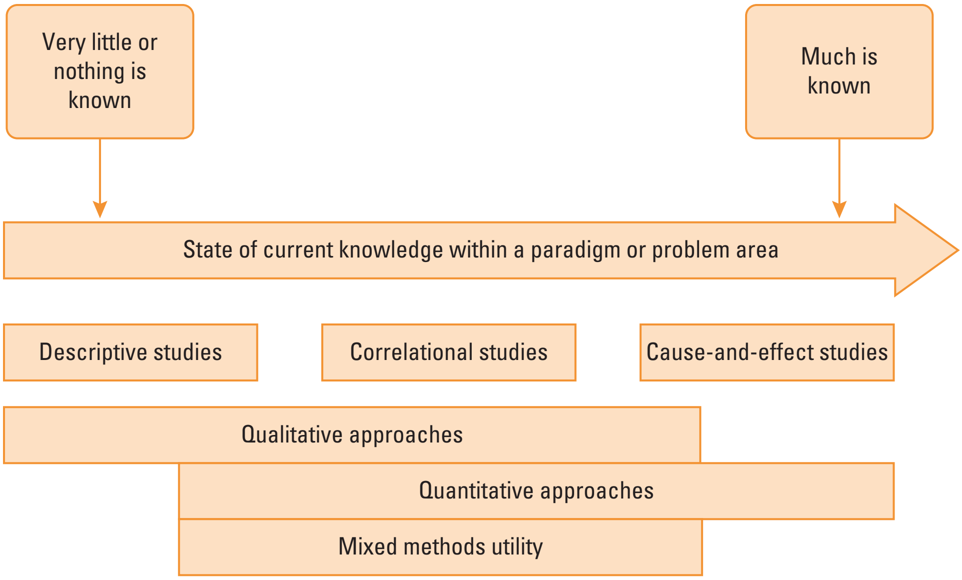


FIGURE 2.3 **The continuum of research. (Used with permission of L. Milanesi, personal communication, April 26, 2015.)**

contexts. It is in this stage that theory generation is important. There is a need to conduct deep analysis to clearly understand the concepts of interest related to the phenomenon and begin to establish relational statements. Such analysis would likely require interviews to understand the experiences of people involved with the phenomenon. A positivist paradigm may guide the worldview of the researcher, but critical and constructivist epistemologies that emphasize the transactional nature of the relationship between the researcher and participant would likely be preferred. The qualitative approaches to inquiry are appropriate—for example, phenomenology, narrative analysis, and grounded theory—and help generate the data regarding personal experiences that may form the essential components for new theory or important revision to existing theory.

The far right side of the continuum involves quantitative approaches to inquiry. Here, theories are fairly well developed in terms of their essential concepts and relational statements. Carefully controlled quantitative studies, such as randomized controlled trials, are used to isolate causal mechanisms. Extensive longitudinal studies may be important here as well because longitudinal studies can help establish the cross-time requirement for establishing causality. Paradigmatic orientations here may be more positivist in orientation.

The space in between these two ends of the continuum represents varying levels of maturity of theory. There is no complete theory and no theory that is 100% correct in all cases. In fact, much of theory in the social sciences requires cross-cultural testing, and replication studies to examine stability of concepts and relational statements during different times are important. Testing in new populations, for example, may require extensive interviews to understand the impact of cultural differences on measurement tools. This qualitative inquiry can then be followed up with quantitative surveys to assess the reliability and validity of new instruments that result from qualitative studies. This is an example of one application of mixed methods research.

Concluding Remarks

The research process can be very exciting, particularly for those who are new scientists learning for the first time how to be independent researchers. We have provided some context about research philosophy and theory that can provide you with a rich set of ideas on the types of studies available to you. It is always good to spend time reflecting on your orientation to the world. For example, are you more positivist in terms of ontology (you believe that there is one true reality) or constructivist (you believe that there are multiple local realities based on unique circumstances of time and place)? Understanding the answer to this question can help orient you to the approaches to inquiry that make the most sense and are consistent with your own views of the world.

We also want to ensure we leave you with the idea that the ways that worldview and paradigm, theory, approaches to research, and methods interact are complex. Figure 2.4 provides a visual to demonstrate that the research process is not necessarily as linear as one might think. Your choices of worldview influence the theories on which you focus and the research approaches with which you are most comfortable; however, it is also true that the process of research can change your view of the world. In addition, researchers should not be constrained by any one method. Multiple methods can be brought together to answer your research question. Understanding

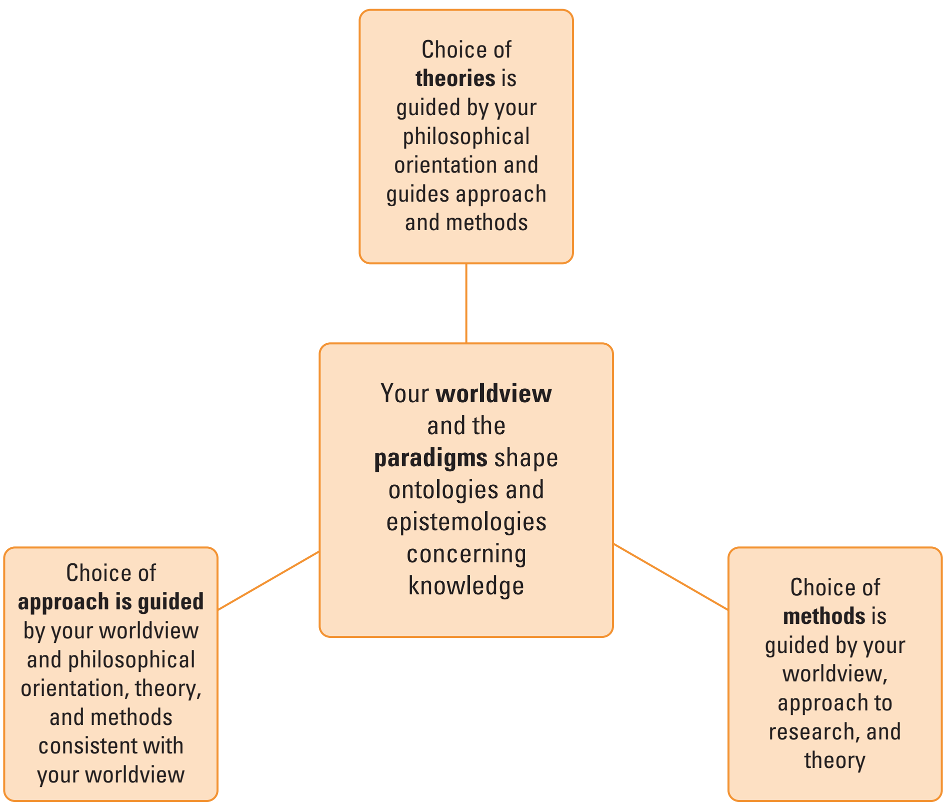


FIGURE 2.4 The interplay between worldview and paradigm, theory, approach to inquiry, and methods.

the critical roles that paradigm, theory, approach to inquiry, and methods play in the research process can guide you to create a research project that is as traditional or as critical/emancipatory as your goals dictate.

There are several implications that arise from our discussion for new researchers in particular. These are summarized below and serve as questions for you to consider.

- *How well do I know the philosophical and theoretical orientations that guide my discipline? What are the assumptions underlying those orientations, and in what ways are they congruent with my own worldview?*
- *What is the ultimate goal of my research? Am I seeking more traditional outcomes, such as those that can be achieved through experimental studies, or is my goal more emancipatory and participatory in that the outcomes are influenced by my coresearchers and that coresearchers influence the direction of the research project?*
- *What is the state of research in my discipline? What does an exhaustive review of the literature tell me about where we are on the research continuum?* This will be critical to guide the type of research that adds new knowledge. It will also give you some ideas about the kinds of research studies and range of research questions that are consistent with the needs of the discipline and your own philosophical orientation.
- *How well have the theories been tested across contexts?* Inadequate testing creates opportunities to expand knowledge and the scope of theory in the discipline. Thus, there are opportunities to do replication studies, which are important for establishing validity as well as confidence in the measurement tools.
- *What are the kinds of approaches to inquiry that will help to generate new data that inform concept development (or expansion of description of existing concepts) and hence theory development?*
- *What are the most appropriate quantitative designs to advance knowledge in my particular discipline?* Sophisticated statistical modeling procedures can now be run quickly using computer software, and the mining of large data sets ("big data") of population-based data is providing new opportunities to control for large numbers of variables in analysis.

Chapter 3 focuses on conceptual and theoretical frameworks. Dr. Linda Crawford describes and defines both kinds of frameworks and distinguishes between the two. Her chapter is a natural follow-up to the present discussion on theory and the paradigms that guide choices for theory and approaches to inquiry.

KEY SOURCES

Green, K., & LeBihan, J. (1996). *Critical theory and practice: A coursebook*. London, United Kingdom: Routledge.

Hayden, J. A. (2013). *Introduction to health behavior theory*. Burlington, MA: Jones & Bartlett Learning.

Johnson, P., & Duberley, J. (2000). *Understanding management research: An introduction to epistemology*. Thousand Oaks, CA: Sage.

Kuhn, T. (2012). *The structure of scientific revolutions: 50th Anniversary edition*. Urbana, IL: University of Chicago Press.

Reynolds, P. D. (2007). *A primer in theory construction*. Boston, MA: Pearson.

REFERENCES

Berry, J. W. (1989). Imposed etics-emics-derived etics: The operationalization of a compelling idea. *International Journal of Psychology*, *24*, 721–735.

Bowleg, L. (2012). The problem with the phrase "women and minorities": Intersectionality, an important theoretical framework for public health. *American Journal of Public Health*, *102*(7), 1267–1273. doi:10.2105/AJPH.2012.300750

Buckingham, W., & Burnham, D. (2015). *The philosophy book: Big ideas simply explained*. London, United Kingdom: Penguin.

Bullock, H. E., Harlow, L. L., & Mulaik, S. (1994). Causation issues in structural modeling research. *Structural Equation Modeling Journal*, *1*, 253–267.

Burbank, P. M., & Martins, D. C. (2010). Symbolic interactionism and critical perspective: Divergent or synergistic? *Nursing Philosophy*, *11*, 25–41.

Byrne, R. (2010). *The secret*. New York, NY: Atria Books.

Clark, T. W. (1992). Relativism and the limits of rationality. *The Humanist*, *52*, 25–42.

Day, L. (2010). *How to rule the world from your couch*. New York, NY: Atria Books.

Denzin, N. K., Guba, E. G., & Lincoln, Y. S. (1994). *Handbook of qualitative research*. Thousand Oaks, CA: Sage.

Dudley-Brown, S. L. (1997). The evaluation of nursing theory: A method for our madness. *International Journal of Nursing Studies*, *34*, 76–83.

Gannon, S., & Davies, B. (2007). Postmodern, poststructural, and critical theories. In S. N. Hesse-Biber (Ed.), *Handbook of feminist research: Theory and praxis* (pp. 71–106). Thousand Oaks, CA: Sage.

Geertz, C. (1973). *The interpretation of cultures: Selected essays*. New York, NY: Basic Books.

Gershenson, C. (2013). The implications of interactions for science and philosophy. *Foundations of Science*, *18*, 81–790.

Guba, E. (1990). *The paradigm dialog*. Newbury Park, CA: Sage.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105–118). Thousand Oaks, CA: Sage.

Habermas, J. (1971). *Knowledge and human interests* (J. J. Shapiro, Trans.). Boston, MA: Beacon Press. (Original work published 1968)

Hacker, K. (2013). *Community-based participatory research*. Los Angeles, CA: Sage.

Hanson, N. R. (1958). *Patterns of discovery*. Cambridge, United Kingdom: Cambridge University Press.

Harman, W. (1991). *A re-examination of the metaphysical foundations of modern science*. Petaluma, CA: Institute of Noetic Sciences.

Hay, L. (1984). *You can heal your life*. Carlsbad, CA: Hay House.

Hicks, E., & Hicks, J. (2006). *The law of attraction: The basics of the teachings of Abraham*. Carlsbad, CA: Hay House.

Houghton, C., Hunter, A., & Meskell, P. (2012). Linking aims, paradigm and method in nursing research. *Nurse Researcher*, *20*(2), 34–39.

Johnson, P., & Duberley, J. (2000). *Understanding management research: An introduction to epistemology*. Thousand Oaks, CA: Sage.

Kim, H. S. (1993). Putting theories into practice: Problems and prospects. *Journal of Advanced Nursing*, *18*, 1632–1639.

Kim, H. S. (1999). Critical reflective inquiry for knowledge development in nursing practice. *Journal of Advanced Nursing*, *29*(5), 1205–1212.

Kincheloe, J. L., McLaren, P., & Steinberg, S. R. (2011). Critical pedagogy and qualitative research: Moving to the bricolage. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 162–177). Thousand Oaks, CA: Sage.

Kuhn, T. (2012). *The structure of scientific revolutions: 50th Anniversary edition*. Urbana, IL: University of Chicago Press.

Kukla, A. (2000). *Social constructivism and the philosophy of science*. London, United Kingdom: Routledge.

Lather, P. (1994). Critical inquiry in qualitative research: Feminist and poststructural perspectives. In B. F. Crabtree, W. L. Miller, R. B. Addison, V. J. Gilchrist, & A. J. Kuzel (Eds.), *Exploring collaborative research in primary care* (pp. 103–114). Thousand Oaks, CA: Sage.

Lincoln, Y. S., Lynham, S. A., & Guba, E. G. (2011). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 96–127). Thousand Oaks, CA: Sage.

McEwen, M. (2011). Theory analysis and evaluation. In M. McEwen & E. M. Wills (Eds.), *Theoretical basis for nursing* (3rd ed., pp. 89–106). Philadelphia, PA: Lippincott Williams & Wilkins.

McTaggart, L. (2003). *The field: The quest for the secret force of the universe*. London, United Kingdom: Element.

Miller, J. R., & Burbank, P. M. (1994). Congruence among philosophy, theory and method in nursing research. In Workgroup of European Nurse Researchers, *The contribution of nursing research: Past-Present-Future*: *Vol. 2*. *Proceedings at the 7th Biennial Conference* (pp. 704–709). Oslo, Norway: Falch Hurtigtrykk.

Mittwede, S. K. (2012). Research paradigms and their use and importance in theological inquiry and education. *Journal of Education & Christian Belief*, *16*, 2–40.

Philosophy. (2006). In *Oxford English Dictionary*. Retrieved from http://www.oed.com.ezp.waldenulibrary.org/view/Entry/142505?rskey=U5jI01&result=1#eid

Plummer, K. (2011). Critical humanism and queer theory: Living with the tensions. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (pp. 195–212). Thousand Oaks, CA: Sage.

Polanyi, M. (1958). *Personal knowledge: Towards a post-critical philosophy*. Chicago, IL: University of Chicago Press.

Ponterotto, J. G. (2005). Qualitative research in counseling psychology: A primer on research paradigms and philosophy of science. *Journal of Counseling and Clinical Psychology*, *52*(2), 126–136.

Reynolds, P. D. (2007). *A primer in theory construction*. Boston, MA: Pearson.

Smith, L. T. (1999). *Decolonizing methodologies: Research and indigenous peoples*. London, United Kingdom: Zed Books.

Theory. (2006). In *Oxford English Dictionary*. Retrieved from http://www.oed.com.ezp.waldenulibrary.org/view/Entry/200431?rskey=DS8Spn&result=1#eid

Warren, J. R., & Marshall, B. (1983). Unidentified curved bacilli on gastric epithelium in active chronic gastritis. *Lancet*, *1*, 1273–1275.

Wilkins, J. S. (2006). Theory. In J. H. Birx (Ed.), *Encyclopedia of anthropology*. Retrieved from http://knowledge.sagepub.com.ezp.waldenulibrary.org/view/anthropology/n866.xml?rskey=sageBg&row=5

3

Conceptual and Theoretical Frameworks in Research

Linda M. Crawford, PhD

Introduction

At the outset of planning your research, you as researcher set the study into a framework that justifies the study and explains the study structure or design. This framework is like a foundation for a house. It provides the essential support for the study components and also clarifies the context of the study for the reader, much like a house blueprint. By constructing this framework, you not only justify and explain the study to others but also check your own understanding of the need for the study, how the study is conceived, what knowledge it will add regarding the topic, and how the elements of the study design align with the problem identified for the study.

This chapter addresses the framework for a research study. One of the difficulties for new researchers in developing a framework for a study is that conceptual and theoretical frameworks are defined and described differently by different authors, and the definition of what is considered a study framework may vary by institution. This chapter will explore those various definitions to provide a spectrum of understanding of conceptual and theoretical frameworks. This chapter also discusses the purpose of the conceptual framework, sources from which these frameworks are derived, and how conceptual frameworks are presented. Given the discrepancy in definitions of conceptual and theoretical frameworks, the term *conceptual framework* will serve as the overriding term for the chapter and will be differentiated in the chapter from the term *theoretical framework*. Finally, some guidance in how to approach the conceptual and theoretical frameworks for your study is provided.

Literature-Based Definitions of Conceptual Frameworks

Before exploring the various understandings of conceptual frameworks in depth, it is helpful to compare multiple definitions of the term. Some authors view conceptual and theoretical frameworks as being synonymous. Interestingly, some research design authors do not provide description or definition of either conceptual or theoretical frameworks, even if they discuss theory; for example, see this omission in Leedy and Ormrod (2013). Please note, however, that this omission from texts does not justify excluding a conceptual framework from a study. A conceptual framework provides the orientation to the study and assists both the researcher and the reader in seeing how the study contributes to the body of knowledge on the topic, how elements of the study align, and how the study design and methodology meet rigorous research standards. In summary, a conceptual framework is incredibly important.

Table 3.1 displays various authors' definitions of conceptual framework. A conceptual framework may be defined broadly as theory or literature review, or it may be defined more narrowly as the factors and variables addressed in a study (Maxwell, 2012; Miles, Huberman, & Saldaña, 2014). Essentially, all definitions of conceptual frameworks provide a context for the study, but the scope of that context varies among authors.

Ravitch and Riggan

Ravitch and Riggan (2012) presented the most comprehensive understanding of a conceptual framework. Indeed, they devoted an entire book to the topic. Their main point of view was that a conceptual framework is an argument for the study and that argument has two parts. First, the argument establishes the importance of and intended audience for the study. Second, the argument demonstrates alignment among research questions, data collection, and data analysis, as well as the use of rigorous procedures to conduct

TABLE 3.1
Definitions of Conceptual Frameworks

Author	Definition
Ravitch and Riggan (2012)	"An argument about why the topic one wishes to study matters, and why the means proposed to study it are appropriate and rigorous" (p. 7).
Miles et al. (2014)	"[An explanation], either graphically or in narrative form, [of] the main things to be studied—the key factors, variables, or constructs—and the presumed relationships among them" (p. 20).
Maxwell (2013)	"The actual ideas and beliefs that you hold about the phenomena studied, whether these are written down or not; this may also be called the 'theoretical framework' or 'idea context' for the study" (p. 39).
Marshall and Rossman (2016)	"The first major section of the proposal—the conceptual framework—demands a solid rationale. In examining a specific setting or set of individuals, the writer should show how she is studying instances of a larger phenomenon. By linking the specific research questions to larger theoretical constructs, to existing puzzles or contested positions in a field, or to important policy issues, the writer shows that the particulars of this study serve to illuminate larger issues and therefore hold potential significance for that field" (p. 6).

the study. They posited that the conceptual framework both informs and describes the development of research questions, design selection, data collection, data analysis, and presentation of findings.

Miles, Huberman, and Saldaña

A major contribution to the idea of conceptual framework presented by Miles et al. (2014) is the graphical representation of the conceptual framework, which will be explored later in this chapter. They promoted spending significant time in developing and representing the conceptual framework. That process encourages a closer assessment of how a study's variables are related, how study participants are characterized, and how data collection instruments are selected.

Maxwell

Maxwell (2013) discussed conceptual frameworks in relation to qualitative research design. For Maxwell, the conceptual and theoretical frameworks are synonymous. Maxwell presented the terms as synonymous because he viewed the conceptual framework as presenting a theory of the phenomenon under investigation (p. 39). A major point of Maxwell's contribution is that the researcher must build, or construct, the conceptual framework from personal experience, prior research, and published theory into a coherent representation of the study.

Marshall and Rossman

Marshall and Rossman (2016) described the conceptual framework as providing a rationale for the study. The idea of rationale is close to Ravitch and Riggan's (2012) view of the conceptual framework as an argument for the study. Marshall and Rossman also emphasized the importance of the foundation of a conceptual framework as grounded in the literature published on the topic under study.

All definitions demonstrate the importance of the relationship of the conceptual framework to the roots of the study purpose and the alignment of study parts. They also indicate ways that a conceptual framework makes the construction of a study clearer, cleaner, and more straightforward. However, another consideration is how researchers define the term *theoretical framework*, particularly in relation to the conceptual framework.

Literature-Based Definitions of Theoretical Frameworks

The definitions of conceptual framework are confounded by the fact that some authors do not differentiate between conceptual and theoretical frameworks. Maxwell (2013), Robson (2011), and Merriam (2009) consider the terms to be synonymous. Anfara and Mertz (2015) do not explicitly relate conceptual and theoretical frameworks, but they imply a synonymous relationship between them. Some authors (Marshall & Rossman, 2016; Miles et al., 2014) offer no discussion of the relationship between conceptual and theoretical frameworks.

Merriam (2009) defined theoretical framework as "the underlying structure, the scaffolding or frame of your study" (p. 66), which seems close to some of the definitions of conceptual framework provided earlier. Anfara and Mertz (2015) defined theoretical frameworks as "any empirical or quasi-empirical theory of social and/or psychological processes, at a variety of levels . . . that can be applied to the understandings of phenomena" (p. 15).

A very clear definition of theoretical frameworks and the relationship between theoretical and conceptual frameworks comes from Ravitch and Riggan (2012). They defined theoretical frameworks as follows:

> In the case of theoretical frameworks, the "parts" referred to in this definition are *theories*, and the thing that is being supported is the relationships embedded in the conceptual framework. More specifically, we argue that the parts are *formal* theories; [*sic*] those that emerge from and have been explored using empirical work. (p. 12)

Ravitch and Riggan required that the theoretical framework be based on published, identifiable theories. Private conceptualizations or theoretical constructions do not qualify. In addition, they held that the theoretical framework resides within the conceptual framework and is not synonymous with it. *In other words, the conceptual framework presents the overall structure of the study, and the theoretical framework within it explains the relationships that are explored within the study*.

Recommended Definitions of Conceptual and Theoretical Frameworks

I recommend that conceptual and theoretical frameworks not be considered synonymous, and I align the definitions used in this text with the guidance provided by Ravitch and Riggan (2012). I adopt Ravitch and Riggan's definition of conceptual framework as "an argument about why the topic one wishes to study matters, and why the means proposed to study it are appropriate and rigorous" (p. 7). For example, a conceptual framework for a study on learning styles would present the reason(s) why studying the particular aspect of learning styles is important, with that reason rooted in the literature; for whom studying the particular aspect of learning styles might make a difference; and how the planned conduct of the study is appropriate and rigorous.

Furthermore, I differentiate between conceptual and theoretical frameworks, conceiving theoretical framework as an explanation of how the study relates to the generation or testing of theory. Building on Ravitch and Riggan (2012), I define theoretical framework as an element of a conceptual framework that situates the relationships explored in the study into the context of developing or testing formal theories.

Consistent with Ravitch and Riggan (2012), the theoretical framework should

1. Identify the theory cluster. A theory cluster combines theories into categories, such as theories of learning style, organizational communication, and language acquisition.
2. Identify specific theories relevant to that cluster, including the originator or source and the major propositions and hypotheses of each theory.

3. Identify the theory selected for the study. This includes specifying the specific theory within the cluster that will be used, the propositions of the theory that relate to the specific study, and the review of prior studies using that theory as a focus.
4. State how the study will contribute to the body of knowledge related to the theory.

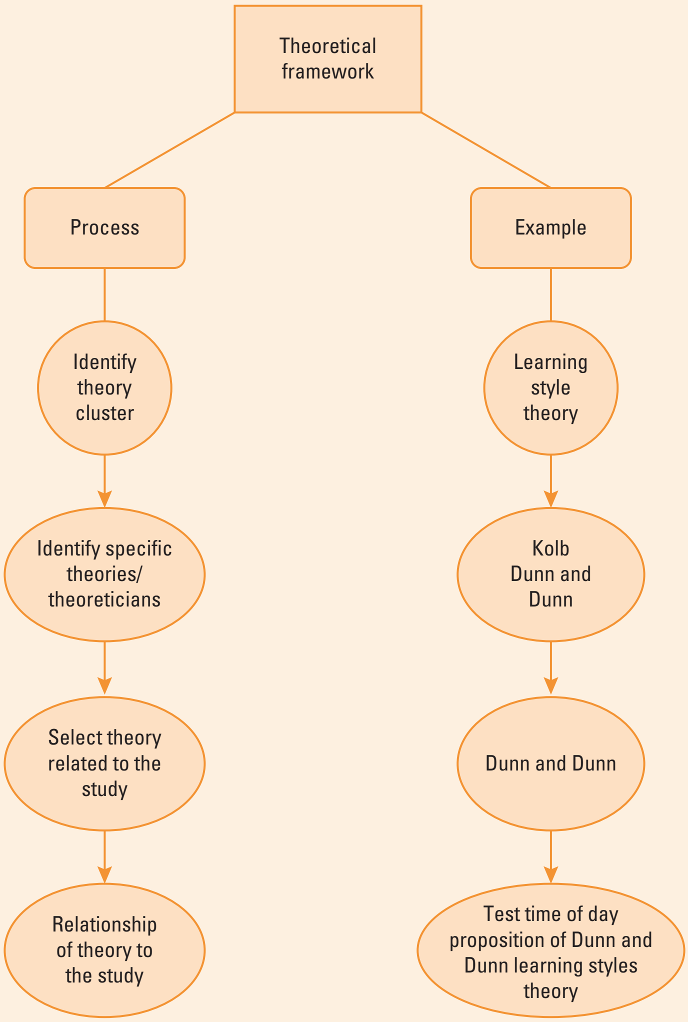
Following the earlier learning style study example, the theory cluster would be learning style theory. There are several different learning style theories, such as Kolb's experiential learning theory model (Kolb, 1984) and the Dunn and Dunn learning style model (Dunn, Dunn, & Price, 1984). Of course, there are more learning style theorists, but these two are presented for the purpose of this example. Within a theoretical framework, if you were doing this study, you would present the major theories that are relevant to the study.

Notice that a theory often has the originator's name associated with it, such as Einstein's theory of relativity, Gardner's theory of multiple intelligences, and Freud's psychosexual theory of human development. Notice, too, that the publication dates for theories are often old. Theories require significant testing over time in order to be verified. Theories supported by research survive the test of time; they may endure or may be replaced as understanding in a particular field deepens. Theories not supported by research tests lose usefulness and eventually fall away, are revised, or are replaced by new theories.

Having identified the theory cluster and the specific theories within the cluster that are related to the study problem, the theories must be explicated. In other words, their major propositions or hypotheses need to be presented. For example, Kolb's (1984) theory holds that individuals show a preference for one of four learning styles—accommodating, converging, diverging, and assimilating—and each style has a certain set of characteristics. Those styles and characteristics would need to be summarized along with any other major propositions or hypotheses of the theory. As another example, Dunn and Dunn's theory (Dunn et al., 1984) offers five stimulus areas—environmental, emotional, sociological, physiological, and psychological—and each of those five areas are associated with certain elements. Those areas and elements, along with any other major propositions or hypotheses, would need to be summarized. These two learning theories are very different from each other. The next task for the researcher is to select the theory most relevant to the study.

For the example of a learning style theory study, consider that the researcher is investigating whether student standardized test scores vary according to the time of day students are tested in relation to their preferred learning style. Since the researcher has explicated both Kolb's and Dunn and Dunn's theories, the researcher has shown that the Dunn and Dunn theory has explicit propositions in regard to time of day as a factor in learning, whereas the Kolb theory does not. Therefore, the researcher selects the Dunn and Dunn theory for inclusion in the theoretical framework, giving that rationale. A review of research on the physiological element of time of day in Dunn and Dunn's theory situates the proposed study within the professional conversation that is related to that theory. Finally, the researcher describes how the proposed study will contribute to using the theory for explanation and prediction.

Example of how a theoretical framework relates to a study.



Purpose of Conceptual Frameworks

To construct an informative conceptual framework, the researcher must understand the purpose of a conceptual framework. Different authors present the purpose of conceptual frameworks in different ways (Table 3.2). Some authors focus on the conceptual framework as argumentation for the study (Marshall & Rossman, 2016; Ravitch & Riggan, 2012). Other authors see the conceptual framework as explanatory (Anfara &

TABLE 3.2
Purpose of Conceptual Frameworks

Author	Purpose
Ravitch and Riggan (2012)	Argue for why the topic matters and why the proposed design and methodology are appropriate and rigorous (p. 7).
Miles et al. (2014)	Explain relationships among key factors/variables/constructs of the study (who and what will be studied; p. 20).
Maxwell (2013)	Clarify, explain, and justify methodology (pp. 39–40).
Robson (2011)	Specify variable relationships and research design (p. 67).
Marshall and Rossman (2016)	Argue for study in terms of meaning and contribution to improving the human condition (p. 67).
Merriam (2009)	Generate study problem, research questions, data collection, data analysis, and interpretation of findings (p. 67).
Anfara and Mertz (2015)	Explain variable relationships (p. 15).

Mertz, 2015; Miles et al., 2014). Merriam (2009) viewed the conceptual framework, which she terms as theoretical framework, as generating elements of the research design and methodology, whereas Robson (2011) emphasized variable relationships and research design. Maxwell (2013) combined purposes of the conceptual framework into clarification, explanation, and justification.

Figure 3.1 reduces the various purposes discussed earlier into three main purposes: argumentation, explanation, and generation. Argumentation focuses on the importance of studying the topic, the appropriateness of the design, and the rigor of the methodology. Explanation stresses the

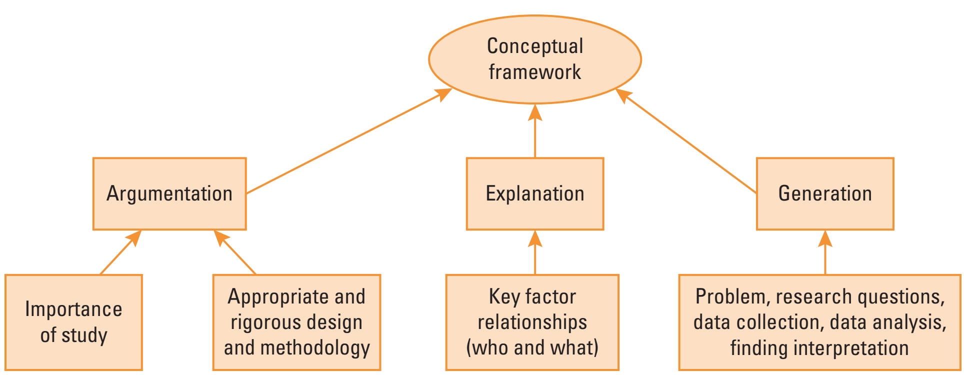


FIGURE 3.1 Purposes of conceptual frameworks.

relationships among who and what will be studied. Generation gives rise to the problem, research questions, and methodology of a study.

I recommend that you incorporate the three major purposes—argumentation, explanation, and generation—when constructing your conceptual framework. By doing so, you will build a comprehensive model that will aid in justifying your study, clarifying the relationships explored in the study, and aligning design elements. As you build the conceptual framework toward these purposes, you must root the framework in verifiable sources, which are explored next.

Sources of Conceptual Frameworks

In this section, the main sources of a conceptual framework are discussed. A source for a conceptual framework is the primary element forming the basis for the development of the framework (Ravitch & Riggan, 2012). You may think of it as the impetus for the conceptual framework. There are three sources for a conceptual framework: experience, literature, and theory.

Experience

Ravitch and Riggan (2012), Maxwell (2013), Robson (2011), and Marshall and Rossman (2016) all allowed for personal interests, experiences, intuitions, and hunches as stimuli for a conceptual framework, although none of them believed that personal experience alone is sufficient. For example, your personal experience observing leadership styles in an organization may stimulate in you a desire to conduct a study on a certain aspect of leadership. Booth, Colomb, and Williams (2008) also believed that personal insights are insufficient to support a study. Personal issues may point you in the direction of a study topic, but the study must have meaning for others in the field. In other words, there must be evidence that others in the field share your concern and that addressing the concern will advance knowledge. Such evidence rests in a literature and theoretical base in order to support a conceptual framework for a study.

Literature

An essential source for your conceptual framework is the published research literature related to your topic. Ravitch and Riggan (2012), Maxwell (2013), Robson (2011), Merriam (2009), and Marshall and Rossman (2016) advocated for rooting the conceptual framework in the literature associated with the topic of study. Of primary importance is that your study is based on a need documented from the literature; your study cannot be based solely in your own personal experience. For example, following the idea in the prior paragraph, your personal experience may point to a desire to study a certain aspect of leadership. Of importance, though, is that you find out from the literature the extent to which that aspect has already been studied, what is still not understood about it, and whether the discipline needs to remedy the lack of knowledge. It may be that the profession needs to understand a phenomenon better (Booth et al., 2008). In either case, the literature review provides the evidence for the argumentation contained in a conceptual framework.

Theory Framework

An additional source for your conceptual framework is theory (Anfara & Mertz, 2015; Marshall & Rossman, 2016; Maxwell, 2013; Ravitch & Riggan, 2012; Robson, 2011), and this source is expressed in the theoretical framework. The study may be focused on generating new theory or on testing theory that has already been constructed (Creswell, 2013). For example, your study may focus on describing how leaders distribute power in an organization. In other words, the focus is on developing an explanation, or theory, of how power distribution functions in a certain kind of organization. Or, your study may focus on testing some theory of power distribution that has already been developed to determine if it accurately explains how power is distributed within a certain group. Whether generating or testing theory, the conceptual framework contains the theoretical framework, or theoretical context, for the study.

Summary of Sources of Conceptual Frameworks

As shown in Figure 3.2, there are three primary sources, or stimuli, for creating a conceptual framework: experience, literature, and theory. Although personal experience may instigate a research idea, personal experience alone is not sufficient to support a conceptual framework for a research study. The conceptual framework must be rooted in the professional literature. The literature provides the rationale for the study by exposing what is not yet known or understood about a phenomenon. The third source for a conceptual framework is theory, integrated as the theoretical framework. Is there already a theory that needs to be tested? Is there no existing viable theory of the phenomenon, and thus, does one need to be developed? Thus, although personal experience may prompt a conceptual framework, the literature must provide the argumentation for pursuing the research idea, and the study must be situated in relation to generating or testing theory.

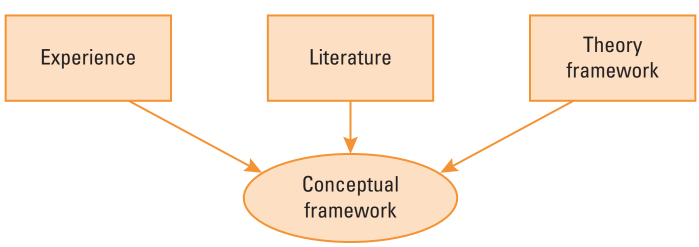


FIGURE 3.2 Primary sources of conceptual frameworks.

Presentation of Conceptual Frameworks

There are two ways to present a conceptual framework—graphically and narratively. If you are crafting a research study for a thesis or dissertation, your institution will probably expect that at a minimum, you describe the conceptual framework narratively, with optional figures to support clarity of presentation. In this section, means of exhibiting a conceptual framework will be examined.

Graphic Presentation

Some authors favor a diagrammatic portrayal of a conceptual framework using a concept map, with or without an accompanying narrative (Marshall & Rossman, 2016; Maxwell, 2013; Merriam, 2009; Miles et al., 2014; Robson, 2011). A concept map is a pictorial portrayal of relationships. It shows how one idea or concept connects to other ideas or concepts.

Miles et al. (2014) provided several fine examples of graphic presentations and concept maps describing conceptual frameworks. As an additional example, Figure 3.3 shows a graphical conceptual framework for a mixed methods study that examined the influence of specific dimensions of supervisor support (mentoring, coaching, social support, and task support) on transfer motivation and training transfer to determine whether transfer motivation mediates the relationships between dimensions of supervisor support and training transfer (Schindler, 2012).

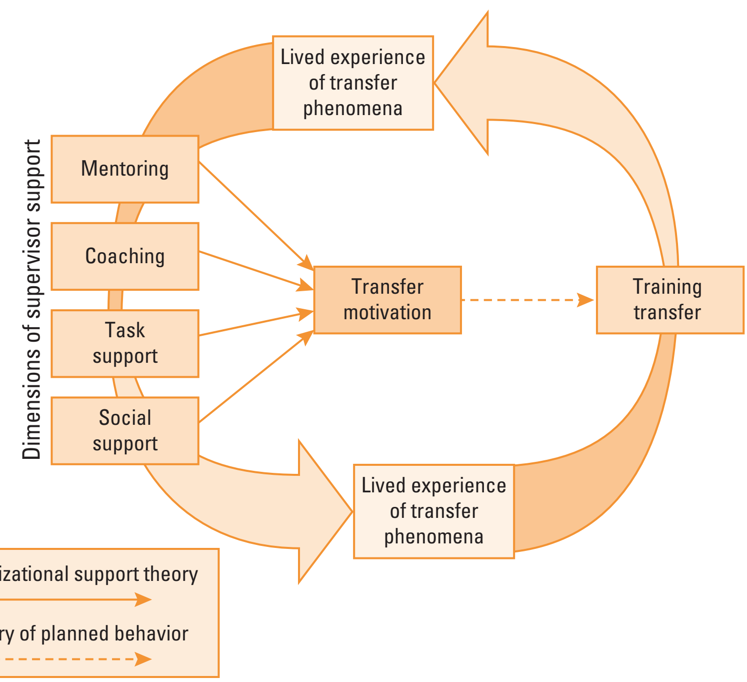


FIGURE 3.3 **Conceptual framework for the influence of supervisor support on training transfer (L. A. Schindler, personal communication, July 15, 2015).**

Figure 3.3 reflects the purpose of the mixed methods study, which was to understand the influence of specific dimensions of support (mentoring, coaching, task support, and social support) on transfer motivation and training transfer. In the quantitative portion of this study, the author examined the relationships between dimensions of supervisor support and training transfer and the degree to which transfer motivation mediates those relationships. In the qualitative portion of the study, the author explored participants' lived experiences of transfer phenomena (i.e., supervisor support, transfer motivation, and training transfer). Both organizational support theory and the theory of planned behavior provided support for this study (Schindler, 2012).

Miles et al. (2014) noted that forcing the graphic onto one page rather than multiple text pages allows you to see and adjust all the parts of the study as a unit as well as to see inconsistencies and contradictions. Going through this process lends cohesiveness to the study design. You should expect development of the graphic to be an iterative process with several versions until it finally accurately represents the study. During this iterative process, how you are writing about the study in text and how you are graphically representing the study become mutually informative. Miles et al. further suggested that you should challenge yourself to avoid overly global graphics with ubiquitous two-way arrows that do not clearly demonstrate the flow of the study.

Like Miles et al. (2014), Robson (2011) advocated presenting the conceptual framework in graphic format. Robson provided six specifications for developing that graphic:

1. Contain the graphic on one page.
2. Include multiple inputs, such as prior research, including pilot studies; relevant theories; hunches in regard to the phenomenon or variable relationships; and thoughts of other professionals in the field.
3. Attain internal consistency within the graphical map.
4. Expect to produce multiple iterations of the framework graphic.
5. Include an item, rather than exclude it, if unsure.
6. Simplify the graphic as you learn from experience.

If you attend to each of the six specifications listed, you will develop a solid graphical presentation of your conceptual framework.

Narrative Presentation

Ravitch and Riggan (2012) were less supportive of a graphical presentation. Although they saw that graphical and narrative presentations of the conceptual framework can work well, they preferred a text-based presentation of conceptual framework when there is a question about presentation. Ravitch and Riggan provided strong examples of narratively presented conceptual frameworks in relation to design, data collection, data analysis, and presentation of findings.

Recommendation

I advocate narrative presentation of the conceptual framework accompanied by an appropriate graphic. The effort to create a one-page graphical model of a study will assist you in coherently conceptualizing the study, determining appropriate alignment of research design elements, and communicating the essential elements to others. I also consider that another benefit of a graphical conceptual framework is that it lifts you from the burden of words, in which some researchers can become mired, and allows you to see the study and interrelationships as a picture. In that way, a framework provides you with an organizing tool. In a published document, the best figure will convey meaning to readers more simply than written text. I maintain, though, that you must also present the conceptual framework in clearly written text in the published document. Narrative presentation of the conceptual framework clarifies key aspects of the study foundation and conveys an understanding of the overall study in the context of knowledge in the discipline.

Summary

In this chapter, I explored definitions of conceptual and theoretical frameworks. I advocated that conceptual and theoretical frameworks should not be considered synonymous but should be understood as different concepts, congruent with the assertions of Ravitch and Riggan (2012). As shown in Figure 3.4, three main purposes were identified for conceptual

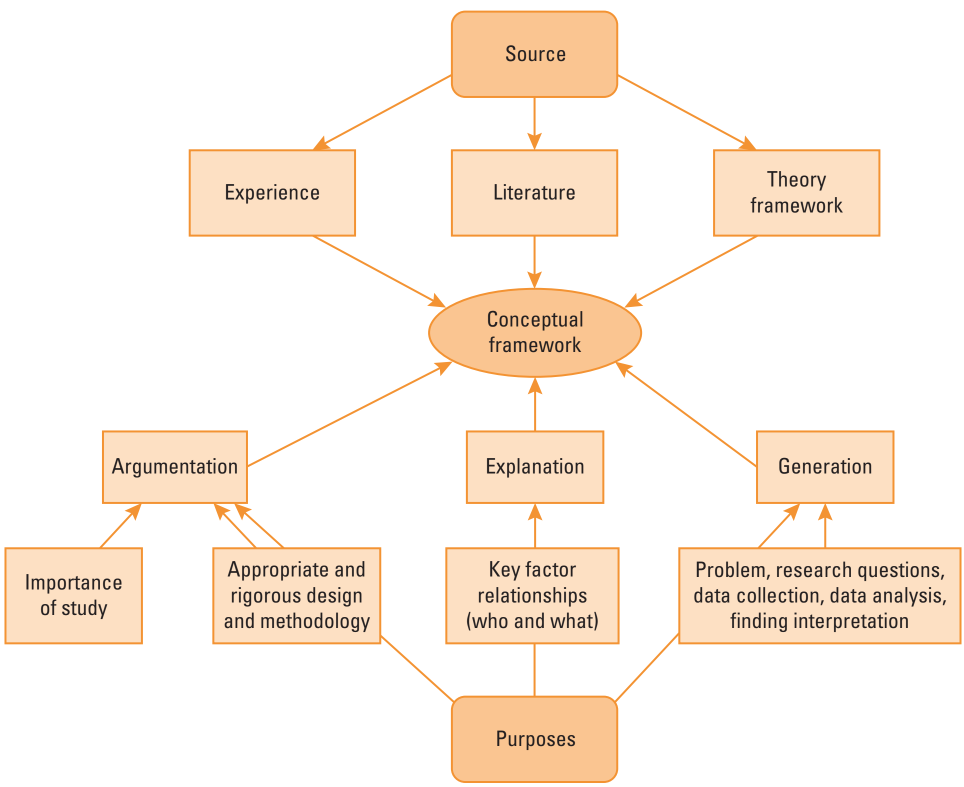


FIGURE 3.4 Purposes and sources of conceptual frameworks.

frameworks—argumentation, explanation, and generation—rooted in the three primary sources of experience, literature, and theory. As part of the conceptual framework, the theoretical framework shows how the study relates to generating or testing theory and explains the relationships that are explored within the study. Finally, I recommended graphical presentation of a conceptual framework accompanied by narrative explication.

The next three chapters—Chapters 4, 5, and 6—provide rich overviews of quantitative, qualitative, and mixed methods approaches to research. In Chapter 4, Dr. Kimberley Cox introduces basic principles of quantitative research designs. She begins with a brief discussion of frequently encountered terms and concepts in quantitative research, such as variables, operationalization, and levels of measurement. Dr. Cox follows this with an overview of common types of randomized experimental, quasi-experimental, and nonexperimental designs. Finally, she describes a number of sampling strategies associated with quantitative research.

KEY SOURCES

The sources listed below are seminal in understanding the concepts of conceptual and theoretical frameworks.

Anfara, V. A., & Mertz, N. T. (Eds.). (2015). *Theoretical frameworks in qualitative research* (2nd ed.). Thousand Oaks, CA: Sage.

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

Ravitch, S. M., & Riggan, M. (2012). *Reason & rigor: How conceptual frameworks guide research.* Thousand Oaks, CA: Sage.

REFERENCES

Anfara, V. A., & Mertz, N. T. (Eds.). (2015). *Theoretical frameworks in qualitative research* (2nd ed.). Thousand Oaks, CA: Sage.

Booth, W. C., Colomb, G. G., & Williams, J. M. (2008). *The craft of research* (3rd ed.). Chicago, IL: University of Chicago Press.

Creswell, J. W. (2013). *Qualitative inquiry & research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Dunn, R., Dunn, K., & Price, G. E. (1984). *Learning style inventory*. Lawrence, KS: Price Systems.

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.

Leedy, P. D., & Ormrod, J. E. (2013). *Practical research: Planning and design* (10th ed.). Boston, MA: Pearson.

Marshall, C., & Rossman, G. B. (2016). *Designing qualitative research* (6th ed.). Thousand Oaks, CA: Sage.

Maxwell, J. (2012). Foreword. In S. M. Ravitch & M. Riggan (Eds.), *Reason & rigor: How conceptual frameworks guide research* (p. xii). Thousand Oaks, CA: Sage.

Maxwell, J. (2013). *Qualitative research design: An interactive approach* (3rd ed.). Thousand Oaks, CA: Sage.

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation: Revised and expanded from qualitative research and case study applications in education*. San Francisco, CA: Jossey-Bass.

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

Ravitch, S. M., & Riggan, M. (2012). *Reason & rigor: How conceptual frameworks guide research*. Thousand Oaks, CA: Sage.

Robson, C. (2011). *Real world research* (3rd ed.). West Sussex, United Kingdom: Wiley.

Schindler, L. A. (2012). *A mixed methods examination of the influence of dimensions of support on training transfer* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (Order No. 3547454). Retrieved from http://search.proquest.com/docview/1268753273?accountid=14

4

Quantitative Research Designs

Kimberley A. Cox, PhD

Introduction

Research designs describe in detail your plans for answering your research questions. In quantitative research, this includes such elements as the specific structural features of your study, measurement of variables, and strategy for sampling. Together, these elements link your research questions to the data you collect. The goals of this chapter are to (a) introduce you to some of the specialized vocabulary that is used in quantitative research and (b) provide an overview of quantitative designs to guide you toward selecting an appropriate and credible design for your research projects.

I begin this chapter with a brief discussion of frequently encountered terms and concepts in quantitative research, such as variables, operationalization, and levels of measurement. I follow with an overview of common types of randomized experimental, quasi-experimental, and nonexperimental designs. Finally, I describe some of the common sampling strategies in quantitative research.

Variables

The design of a study often begins with curiosity about an abstract idea or phenomenon that cannot be directly measured, such as motivation or autonomy. That is, interest may begin as a *concept*, an "abstraction, a representation of an object, or one of that object's properties, or a behavioral phenomenon" (Frankfort-Nachmias & Nachmias, 2008, p. 24) that does not exist in a physical sense (Loseke, 2013). A *construct* typically refers to a type of concept that theoretically holds an agreed upon meaning, but cannot be directly observed. Concepts are common in many disciplines because of their usefulness as "heuristics devices" of communication (Frankfort-Nachmias & Nachmias, 2008, p. 24). Due to their subjective nature, concepts hold different meanings from person to person. It is necessary, therefore, to translate concepts into variables.

A *variable* is a measurable attribute or characteristic that varies in value. Whereas a concept may hold many different meanings from person to person, a variable is a concept translated into a measurable attribute that holds different values. For example, a concept such as autonomy could mean one's sense of control in defining and carrying out their work in the context of a job setting (Heidemeier & Wiese, 2014), or it could mean one's experience of psychological freedom or interpersonal distance from persons in positions of authority, such as parents, in the context of adolescent development (Van Petegem, Vansteenkiste, Soenens, Beyers, & Aelterman, 2015). The process of translating a concept, such as motivation or autonomy, into a variable that can be empirically tested is called *operationalization*; the result of this translation is an operational definition (Loseke, 2013).

Operationalization

Your task in the process of operationalization is to determine the procedures for measuring your variables of interest. The result is an *operational definition*, which is a definition that describes a variable based on the specific procedures for measuring or manipulating it, and in doing so, "bridge[s] the conceptual-theoretical and empirical-observational levels" (Frankfort-Nachmias & Nachmias, 2008, p. 28). For any particular variable, there are potentially countless possible operational definitions.

For example, returning to the concept of autonomy, Van Petegem et al. (2015) conducted a longitudinal study to examine whether adolescents' oppositional defiance to their parents leads to greater autonomy. The researchers measured autonomy on two dimensions: distance and volition. Specifically, Van Petegem and colleagues operationally defined autonomy along these two dimensions as the adolescents' self-reported scores on the Emotional Independence subscale of the Psychological Separation Inventory (Hoffman [as cited in Van Petegem et al., 2015]) and the Choicefulness subscale of the Self-Determination Scale (Sheldon, Ryan, & Reis [as cited in Van Petegem et al., 2015]), respectively. Heidemeier and Wiese (2014) also conducted a longitudinal study on autonomy, but they examined the concept through a different research question—whether women's achievement goals and autonomy would affect their readjustment to work upon returning to the workplace after maternity leave. Heidemeier and Wiese operationally defined autonomy as the women's self-reported scores on a Freedom scale that was part of a larger battery of assessments.

These examples illustrate that translating a concept into a variable can result in various definitions, but through the process of operationalization, researchers provide a definition for their readers that clearly describes the variable based on the specific procedures for measuring or manipulating it. As such, these examples are intended to give you some context about the decision-making process in operationalizing concepts. Just as these other researchers defined the same concept in different ways, you too must make a series of decisions to arrive at operational definitions that are aligned with your particular research question(s). A good starting point is the selection of a variable's level (or scale) of measurement.

Levels of Measurement

Measurement involves the assignment of values, such as labels or numbers, to represent the attributes of a variable. Values are assigned using four levels of measurement: nominal, ordinal, interval, or ratio. Some variables can have values on any one of these scales depending on the procedure used for its measurement. For example, educational attainment could be represented by category, such as high school diploma, bachelor's degree, master's degree, and doctoral degree. Educational attainment could also be represented by a number, such as the number of years of education completed. The choice of a variable's level of measurement typically coincides with writing an operational definition. This choice is important because a variable's level of measurement will inform the type of statistical analysis that you can perform to analyze the data, which, in turn, influences your interpretation of the findings.

Nominal Scale. The values of a nominal scale reflect a name, label, or category represented with words. For example, values for educational attainment might be high school diploma, bachelor's degree, and so on. This level of measurement enables you to classify things, individuals, or responses into groups that share a common attribute. The values reflect differences in quality (type), not quantity (amount), and therefore have no corresponding mathematical properties. A common example of a nominal variable is gender. Although researchers can assign numbers to represent the values of the categories they create, such as 1 5 female and 2 5 male, these numbers are arbitrary and have no intrinsic meaning.

Ordinal Scale. The values of an ordinal scale share the properties of the nominal scale in that they classify things, individuals, or responses into groups that share a common attribute, but the values imply a meaningful order or rank on some dimension. This level of measurement, therefore, communicates the position of values relative to each another, such as lower or higher or less than or more than. The distance between values, however, may not be equal. Therefore, it may not be possible to make meaningful mathematical comparisons between any two values on an ordinal scale. A common example of an ordinal variable is socioeconomic status (SES), which can be classified most generally as low, middle, and high status. The difference between low and middle SES may not be the same as the difference between middle and high SES.

Interval Scale. The values of an interval scale share the properties of nominal and ordinal scales, but the distance between any two values is equal. This level of measurement, therefore, communicates the rank of values with equal differences in the distance between each value. Temperature on the Fahrenheit or Celsius scales, for example, represents an interval scale. A difference in temperature of 72°F and 75°F is the same as a difference in temperature of 89°F and 92°F such that the 3-degree difference represents the same amount of energy regardless of where it is located on the scale. An interval scale, however, does not have a true or fixed zero point that reflects the absence of the attribute being measured. There is no absence of temperature, for example, at 0°F.

In some disciplines, there is debate about treating the scores obtained from rating scales such as a Likert scale, which presents response options that range along a continuum (e.g., *strongly disagree* to *strongly agree*) as interval data rather than ordinal data (Norman, 2010). This is because the differences between the categories of responses may be perceived differently depending on the perception of the responder. However, in practice, many researchers treat Likert scale data as an interval scale. There are also more sophisticated and accessible statistical procedures that can be used that treat data as ordinal. Further discussion on this point, however, is beyond the scope of this book.

Ratio Scale. The values of a ratio scale share the properties of nominal, ordinal, and interval scales, but this scale has a fixed zero point. This level of measurement, therefore, communicates the rank of values with equal differences represented by the distance between each value, and there is a true zero point that reflects the absence of the attribute being measured. This means that researchers measure values on a ratio scale starting at zero. Length and weight, for example, are data that are measured on a ratio scale because they have a fixed zero point representing absence of the attribute.

Types of Variables

There are three main types of variables in quantitative research designs: independent, dependent, and confound. A concept, as described earlier, can potentially be any one of these types of variables.

Independent Variable. The independent variable is the presumed factor that causes a change in the situation or phenomenon under study. In experimental designs, it is the variable that is varied or is manipulated across groups or conditions. For example, if you were interested in studying the impact of a recycling campaign on recycling behavior, the type of recycling campaign (with two levels, such as its presence or absence) would serve as the independent variable. By instituting the campaign (independent variable), you anticipate change in recycling behavior because of its presence.

Some researchers only use the term *independent variable* for experimental designs that involve the manipulation or control of the potential causal factor. The term *predictor variable* is commonly used instead when a research design is nonexperimental, such as in correlational research. Others use the term *independent variable* to represent any potential influential or causal factor. Researchers also tend to refer to variables that reflect intrinsic characteristics of the study population that cannot be manipulated or controlled, such as gender or age, as *subject variables* or *attribute variables* rather than independent variables.

Dependent Variable. The dependent variable is the presumed effect or outcome. In experimental designs, it is the variable that is measured to determine if the independent variable had an effect. Put another way, the outcome of the dependent variable depends on the independent variable. Returning to the recycling campaign example, the dependent variable in this study would be recycling behavior. In nonexperimental research designs, such as correlational research, the term *criterion variable* is often used rather than dependent variable.

In a study on residential curbside recycling behavior, Schultz (1999) conducted an experiment to investigate whether feedback interventions designed to activate personal and social norms—that is, beliefs about one's behavior and beliefs about other people's behavior, respectively—lead to an increase in recycling among 605 households. The independent variable was the feedback intervention with five levels: plea, plea plus information, plea plus neighborhood feedback, plea plus individual household feedback, and the control condition. To measure curbside recycling behavior, Schultz collected data over a 4-week period on the households' frequency of participation in the curbside recycling program, average amount of material recycled on a weekly basis, and proportion of nonrecyclable materials placed in recycling bins. Each of these three behaviors represented three dependent variables.

Confound. A confound generally refers to two categories of variables that can potentially have an undesired influence on the results of a study. One category of variables called *extraneous variables* refers to factors other than the independent variable that were not measured or were deemed not of interest that could confound or affect the dependent variable. Researchers can use design techniques to minimize confounds, such as random assignment, which refers to assigning participants to groups or conditions such that each participant has the same chance of being assigned to any given group or condition. In random assignment, values of the unmeasured variables should be evenly divided among the groups, and their effect therefore is cancelled. Researchers can also use statistical techniques, such as analysis of covariance, which allows for statistically controlling for the effects of other variables.

Intervening variables refer to factors that occur during the study that could also potentially affect the dependent variable, such as boredom or motivation on the part of participants. For example, during the course of a study involving a lengthy recycling campaign, participants might lose motivation over time, or their awareness of being observed might lead to any increases found in recycling behavior rather than the campaign's components.

Types of Quantitative Research Designs

The research design describes the plans for answering a research question(s), including such elements as the specific structural features of the study, the measurement of variables, and the strategy for sampling. All research designs have strengths and weaknesses. The selection of a research design rests largely on weighing these strengths and weaknesses given the study's research question, research problem, and purpose.

Although only a selection of common examples from three main types of designs—randomized experimental, quasi-experimental, and nonexperimental—are presented here, their basic components can serve as a guide for designing more sophisticated studies to meet the needs of a variety of research contexts. The notation used in this section is based on Campbell and Stanley's (1966) seminal work. This notation reflects a shared language that researchers commonly use to communicate about quantitative research designs by way of a diagram, where the letter X refers to the independent variable or treatment, and the letter O refers to the dependent variable or observation. When a design calls for more than one independent variable (e.g., multiple interventions) and/or more than one dependent variable (e.g., multiple measures), a subscript numeral is often used, such as X_1 and X_2 or O_1 and O_2, respectively. The letter *R* is used to denote random assignment. Each group is displayed on a separate line in the diagram. Although the examples presented below depict two groups, the designs can include more. Last, the placement of the notations conveys a temporal sequence that should be read from left to right, reflecting the steps in the study procedures.

Randomized Experimental Designs

The goal in designing an experiment is to determine whether there is a causal link between the independent variable and dependent variable. The design of an experiment identifies the operational definitions of these variables and the randomization procedure, which refers to the method used to randomly assign participants to groups or conditions. An experiment that is designed well and that is aligned with the research question(s) and hypotheses provides a strong foundation for making inferences about causation.

Many experimental designs call for a comparison between one or more groups or conditions. The experimental group or condition is exposed to a treatment of the independent variable, where a treatment could mean such things as a task, event, or intervention, and the control group or condition is not. The important difference, then, between the groups or conditions is the independent variable; both groups are identical and encounter the same conditions with the exception of exposure to the independent variable.

A significant feature of randomized experimental designs is *control*. Control refers to the procedures that researchers use, such as sampling and random assignment, to minimize threats to the internal validity of a study. Controlling for differences between groups before introducing the independent variable minimizes differences in participants within each of the groups or conditions and minimizes any potential confounds that could affect the observed outcome. The strength of randomized experimental designs is this feature of control that maximizes internal validity, that is, the extent to which the treatment (independent variable) is responsible for the outcome (dependent variable).

One method for addressing control is to design the experiment such that any factors or procedures are held as constant as possible across groups with the exception of manipulation of the independent variable. For example, within this design, all participants might complete a test or activity under the same conditions, such as location, time of day, and so on. Although such steps minimize differences, participants are inherently not all the same. Therefore, a second method used to address control is called *random assignment*.

Random assignment is a distinguishing feature of randomized experimental designs. It refers to a procedure for assigning participants to groups or conditions such that each participant has the same chance of being assigned to any given group or condition. Shadish, Cook, and Campbell (2002) present seven methods of randomization, including simple random assignment, batch randomization, and random assignment from matches or strata. Randomization affords the assurance of equivalent groups except for the treatment (independent variable). This equivalence enables causal inferences with more confidence: If the groups differ on the dependent measure after the treatment, it can be said with confidence that the difference was caused by the independent variable, not by the characteristics of the participants. Randomization is not without its controversies in scientific circles, however. One such controversy is the debate around the potential ethical implications of random assignment. The main arguments in favor and opposition of random assignment are covered in more detail in Shadish et al.

To broadly summarize, experiments feature (a) control methods that are executed through careful planning of experimental procedures and random assignment of participants and (b) a comparison between groups or conditions based on manipulation of an independent variable. The following section describes several popular randomized designs.

Posttest-Only Control Group Design. In a posttest-only control group design, participants are randomly assigned to the experimental and control groups before presentation of the independent variable (X). Participants in the experimental group are exposed to a treatment, and the control group is not. The two groups are measured on the dependent variable (O) after (or during) exposure to the independent variable. This design can be illustrated as follows:

$$\begin{array}{ccc} R & X & O \\ R & & O \end{array}$$

The lack of a pretest eliminates any potential carryover or practice effects that are possible when multiple measures of the dependent variable are used. There are also some situations where a pretest may not be possible or practical. Therefore, this type of design is most appropriate when the availability or use of a pretest is prohibitive or known to interact with the independent variable. Although random assignment allows the assumption that the groups are equivalent, there is no complete assurance that this is the case. Without a pretest, your ability is limited to determine whether any differences observed between the groups on the dependent variable are due to the independent variable (or potentially a confounding variable). In addition, the potential for loss of participants due to attrition is a possible concern without a pretest because there is no way to compare participants who dropped out of the study with those who remained in the study (Shadish et al., 2002). Therefore, when possible, I recommend the addition of a pretest.

Pretest-Posttest Control Group Design. In a pretest-posttest control group design, participants are randomly assigned to the experimental and control groups. Each group is measured on the dependent variable (O_1) followed by presentation of the independent variable (X). Participants in the experimental group are exposed to a treatment, and the control group is not. Both groups are then measured again on the dependent variable (O_2). This design can be illustrated as follows:

$$\begin{array}{cccc} R & O_1 & X & O_2 \\ R & O_1 & & O_2 \end{array}$$

With the addition of the pretest, this design can speak to whether any differences exist between the groups before introduction of the independent variable. If the results reveal that the groups differ on the posttest measure (O_2) but not the pretest measure (O_1), this design provides you with the confidence to report that the independent variable is responsible for the observed outcome. This design is well suited to manage attrition, but potential threats to internal validity still exist, such as testing and instrumentation. Testing refers to the potential for a pretest to affect participants' performance on the posttest. Instrumentation refers to the possibility of procedural changes in the use of an instrument between administration of the pretest and posttest.

There are numerous variations to the pretest-posttest control group design that can be tailored to the needs of a research context. For example, this design can be extended to compare more than one treatment (e.g., two different types of instructional delivery) with two experimental groups each receiving a level of the independent variable and a control group. This design can be illustrated as follows:

$$\begin{array}{cccc} R & O_1 & X_A & O_2 \\ R & O_1 & X_B & O_2 \\ R & O_1 & & O_2 \end{array}$$

It is also possible to include a second control group if your goal is to determine if a pretest measure affects participants' scores on the posttest measure. This design, called the *Solomon four-group design*, can be illustrated as follows:

$$\begin{array}{cccc} R & O_1 & X & O_2 \\ R & O_1 & & O_2 \\ R & & X & O_2 \\ R & & & O_2 \end{array}$$

Factorial Designs. The designs presented so far allow you to vary one independent variable (X) at a time. Factorial designs, by comparison, permit you to vary several independent variables within a single experiment. For example, in the simplest of factorial designs, called a 2 × 2 factorial design, there are two independent variables, referred to as *factors*, each with at least two levels. Participants are randomly assigned to one of four possible combinations of experimental treatment (X_{A1B1} and so on). The four groups are measured on the dependent variable (O) after (or during) exposure to the independent variable. This design can be illustrated as follows:

$$\begin{array}{ccc} R & X_{A1B1} & O \\ R & X_{A1B2} & O \\ R & X_{A2B1} & O \\ R & X_{A2B2} & O \end{array}$$

The main advantages of factorial designs are that they allow you to test several hypotheses or predictions of the relationship between variables or the study's outcome at the same time because they accommodate combinations of treatments and they permit the testing of interactions among factors. This type of design would allow for an examination of a wider scope to determine, for example, if a combination of treatments is more effective than a single treatment. Factorial designs are typically more suitable for settings that afford a great deal of control over the various combinations, such as in laboratory research (Shadish et al., 2002).

Quasi-Experimental Designs

Not all research questions can be answered with a randomized experimental design, whether due to the nature of the problem under study, ethical issues, and/or practical constraints. Research in the social, behavioral, management, and educational disciplines often involves the study of naturally occurring groups or groups for which membership is self-selected or intact prior to an investigation. In these situations, it may not be possible or ethical to change an individual's group assignment or randomly assign group members into experimental and control groups. Quasi-experimental designs attempt to approximate randomized experiments. They share some commonalities with experimental designs, such as the comparison of groups on a dependent variable. The key difference, however, is the lack of random assignment of participants to groups.

The lack of random assignment in quasi-experimental designs means that the groups may not initially be equal or similar. This presents the challenge of ruling out other alternative explanations that could be responsible for any observed outcome. Groups for which individuals self-select their membership, for example, may differ for any number of reasons other than the treatment or manipulation of the independent variable. A quasi-experimental design, then, would not afford you with the same confidence in making causal inferences that you would have from a randomized experimental design.

Two main types of quasi-experimental designs identified by Cook and Campbell (1979) are presented in this section to illustrate their general elements: nonequivalent groups design and time series design. These designs involve collecting data from at least two nonequivalent groups and collecting multiple observations over time, respectively. There are, however, numerous types of quasi-experimental designs that can accommodate various research contexts. I recommend Campbell and Stanley (1966) and Cook and Campbell (1979) for their extensive coverage of quasi-experimental designs.

Nonequivalent Groups Designs. In nonequivalent groups designs, participants are not assigned to groups at random. The term *nonequivalent* is used because without randomization, there may be characteristics of the participants that are responsible for any observed difference between the groups other than the independent variable. When only a posttest is used, the *posttest-only design with nonequivalent groups* (Cook & Campbell, 1979) involves two or more groups, including an experimental group that is exposed to the treatment (X), where a treatment could mean such things as a task, event, or intervention, and a control group that is not. The two groups are measured on the dependent variable (O) after (or during) exposure to the independent variable. This design can be illustrated as follows:

$$\begin{array}{cc} X & O \\ & O \end{array}$$

This design has limited use without a pretest and randomization, and it is considered to be among the weakest designs. With the addition of a pretest, the *pretest-posttest design with nonequivalent groups* (Cook & Campbell, 1979) holds greater credibility. Here, both groups are measured on the dependent variable (O_1) followed by presentation of the independent variable (X). Participants in the experimental group are exposed to a treatment, and the control group is not. Both groups are then measured again on the dependent variable (O_2). This design can be illustrated as follows:

$$\begin{array}{ccc} O_1 & X & O_2 \\ O_1 & & O_2 \end{array}$$

Time Series Designs. In time series designs, participants in the experimental and control groups are measured on multiple observations of the same dependent variable (O_1, O_2, and so on for any number of observations) before and after exposure to some treatment (X). Time series designs enable you to identify trends in change of the dependent variable that may be attributed to the impact of the treatment. There are several variations of the simplest of time series designs, called the *simple interrupted time series design with a nonequivalent control group* (Cook & Campbell, 1979), which can be illustrated as follows in this example involving eight observations:

$$\begin{array}{ccccccccc} O_1 & O_2 & O_3 & O_4 & X & O_5 & O_6 & O_7 & O_8 \\ O_1 & O_2 & O_3 & O_4 & & O_5 & O_6 & O_7 & O_8 \end{array}$$

Nonexperimental Designs

Correlational Design. When your goal is to understand the nature of the relationship or association between naturally occurring variables that cannot be manipulated, a correlational design may be an appropriate choice. The key difference here from experimental designs is that the values of variables in a correlational design are measured, but they are not manipulated. This design, therefore, would not provide information about the direction of the causal relationship between variables, which is called the *bidirectionality problem* (Does A cause B, or does B cause A?). This design also cannot provide information on whether a confound is the plausible explanation of the relationship between A and B; this is known as the *third variable problem*, because it is not clear whether a third variable C explains the relationship between A and B. The lack of manipulation of variables, the bidirectionality problem, and the third variable problem result in a design that limits your ability to make casual inferences (Hatfield, Faunce, & Job, 2006).

A correlational design is not to be confused with the statistical test called correlational analysis or the correlational coefficient statistic. A researcher's lack of confidence in inferring causation from data obtained through a correlational design is due to the nature of this type of design, not the statistical test used to analyze the data (Hatfield et al., 2006). Researchers can perform correlational analysis on data collected from a study that used a randomized experimental design, such as the pretest-posttest control group design, and make causal inferences based on the results of this analysis because the data were collected with the rigor characteristic of experimental methods that includes experimental control and random assignment.

Despite the potential problems of bidirectionality and third variables, correlation research does enable researchers to make predictions. That is, if there is a correlation between two variables, then knowing the scores of one of the variables can be used to predict the scores of the other variable. Researchers compute these predictions using statistical tests. When there is one predictor variable, simple linear regression is the statistical test used to investigate the linear relation between the two variables. When there is more than one predictor variable, multiple regression is the statistical test used. In the social, behavioral, management, and educational disciplines, the findings from correlational research often lead to experimental research studies that are used to investigate potential causal relations under the tighter controls that experimental designs afford.

Sampling

The process of selecting participants for a research study is called *sampling*. The goal in sampling is to maximize generalizability and minimize *sampling error*, which refers to the variability or difference between characteristics of the population and sample. A *population* refers to an entire set or collection, and a *sample* refers to a subset of the population. The goal is to select a sample from the population that is most representative of that population. Strategies to select a sample can be broadly categorized as *probability sampling* or *nonprobability sampling*. Your choice of sampling strategy will determine the extent to which you can make generalizations from the findings about the sample to the population from which it was drawn.

Probability Sampling

In probability sampling (also called *random sampling*), the selection of participants is determined by chance, with each member of a population having an equal chance of being selected. As a result, you can state the probability for which each member of the population will be included in the sample. Probability sampling is advantageous because the sample is more likely to be representative of the population, which would allow you to make inferences about the population from the sample. There are three common types of probability sampling: simple random sampling, systematic random sampling, and stratified random sampling.

Simple Random Sampling. Simple random sampling involves drawing a sample so that every member of a population has an equal chance of being selected. This type of sampling implies that you can obtain a numbered list of the population to draw the sample. You could flip a coin or use a table of random numbers to select members when the sample size is small. For obtaining larger sample sizes, random number generators, such as those available in some software programs such as Statistical Package for Social Sciences (SPSS; IBM Corp., Armonk, NY), are more efficient.

Systematic Random Sampling. Systematic sampling is similar to simple random sampling in that it begins with randomly selecting a member from the population, but then sampling proceeds with selecting every *n*th member. To begin, the population size is divided by the desired sample size to obtain *n*, and then participants are selected from every *n*th member.

Stratified Random Sampling. Stratified random sampling requires prior knowledge of a characteristic from the population that may be related to the dependent variable. This characteristic is used to form sampling strata (subgroups). Members of the population are first categorized based on their value on this characteristic, and then a percentage of members from these strata is randomly selected based on the ratio known in the population. Nationwide surveys often use stratified random sampling where the population is stratified, for example, by geographic region. The goal of this sampling method is to obtain a sample that represents the population on the stratified characteristic. This method is appropriate to use when it is important that the ratio of subgroups in the sample reflect the population from which they are drawn.

Nonprobability Sampling

In nonprobability sampling (also called *nonrandom sampling*), the probability of selection is not known. That is, each member of a population does not have an equal chance of being selected. Researchers in the social, behavioral, management, and educational disciplines frequently use nonprobability sampling for a variety of reasons. Practical considerations, such as time or difficulty, for example, may preclude the feasibility of probability sampling. In these circumstances, nonprobability sampling strategies are more feasible. There are three common types of nonprobability sampling: convenience sampling, purposive sampling, and quota sampling.

Convenience Sampling. Convenience sampling refers to selecting a sample based on availability. For example, much psychological research is conducted with convenience samples of undergraduate college students because of their availability. Researchers are limited in their ability to speak to the representativeness of a sample of a population when it is selected through convenience sampling.

Purposive Sampling. Purposive sampling refers to selecting a sample based on a particular purpose that meets the needs of the research study. This sampling method is often used when there is interest in a specific characteristic that only certain people possess, such as a specific life experience or skill. For example, Chueh and Chang (2013) used purposive sampling to select a sample of male veterans residing in a nursing home for a study that sought to investigate the effects of a therapeutic intervention on veterans' depressive symptoms. As with convenience sampling, a weakness of purposive sampling lies in the sample's potential lack of representativeness of a population.

Quota Sampling. Quota sampling is similar to stratified sampling except the probability of selection is unknown. In this sampling method, a quota or proportion for each stratum (subgroup) is determined that will proportionally represent a particular characteristic of interest in the population from which to select a sample.

Conclusion

In this chapter, I provided an overview of quantitative research designs to help guide you toward selecting an appropriate and credible design for your research projects. In doing so, I presented some of the specialized vocabulary that is encountered in designing and conducting quantitative research. I defined the main types of variables in quantitative research designs and described the process of operationalization that entails translating concepts into variables to arrive at an operational definition. I then covered levels of measurement followed by examples of three common types of randomized experimental, quasi-experimental, and nonexperimental designs. Last, I discussed some common sampling strategies for selecting probability and nonprobability samples in quantitative research.

In Chapter 5, Dr. Linda M. Crawford provides an overview of qualitative research. She presents definitions and purposes of qualitative research as well as descriptions of predominant qualitative research designs. Dr. Crawford provides a discussion of key concepts for qualitative research, including sampling logic, trustworthiness, data analysis, and the researcher's role. She also offers guidance on how to select an appropriate qualitative design.

KEY SOURCES

Campbell, D. T. (1957). Factors relevant to the validity of experiments in field settings. *Psychological Bulletin, 54*(4), 297–312. doi:10.1037/h0040950

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

Carifio, J., & Perla, R. J. (2007). Ten common misunderstandings, misconceptions, persistent myths, and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Science*, *3*, 106–116. doi:10.3844/jssp.2007.106.116

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Skokie, IL: Rand McNally.

Crano, W. D., Brewer, M. D., & Lac, A. (2014). *Principles and methods of social research* (3rd ed.). New York, NY: Routledge.

Gardner, P. L. (1975). Scales and statistics. *Review of Educational Research*, *45*(1), 43–57. doi:10.3102/00346543045001043

Rubin, D. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688–701. doi:10.1037/h0037350

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

REFERENCES

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

Chueh, K., & Chang, T. (2013). Effectiveness of group reminiscence therapy for depressive symptoms in male veterans: 6-Month follow-up. *International Journal of Geriatric Psychiatry*, *29*, 377–383. doi:10.1002/gps.4013

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Skokie, IL: Rand McNally.

Frankfort-Nachmias, C., & Nachmias, D. (2008). *Research methods in the social sciences* (7th ed.). New York, NY: Worth.

Hatfield, J., Faunce, G. J., & Job, R. F. S. (2006). Avoiding confusion surrounding the phrase "correlation does not imply causation." *Teaching of Psychology*, *33*(1), 49–51. doi:10.1207/s15328023top3301_9

Heidemeier, H., & Wiese, B. S. (2014). Achievement goals and autonomy: How person–context interactions predict effective functioning and well-being during a career transition. *Journal of Occupational Health Psychology*, *19*(1), 18–31. doi:10.1037/a0034929

Loseke, D. R. (2013). *Methodological thinking: Basic principles of social research design*. Thousand Oaks, CA: Sage.

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Science Education*, *15*(5), 625–632. doi:10.1007/s10459-010-9222-y

Schultz, P. W. (1999). Changing behavior with normative feedback interventions: A field experiment on curbside recycling. *Basic and Applied Social Psychology*, *21*(1), 25–36. doi:10.1207/15324839951036533

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Van Petegem, S., Vansteenkiste, M., Soenens, B., Beyers, W., & Aelterman, N. (2015). Examining the longitudinal association between oppositional defiance and autonomy in adolescence. *Developmental Psychology*, *51*(1), 67–74. doi:10.1037/a0038374

5

Qualitative Research Designs

Linda M. Crawford, PhD

Introduction

This chapter is an introduction to qualitative research. Qualitative research focuses on generating theory, and it is more commonly used at the initial stages of understanding a phenomenon. As an overview of qualitative research, this chapter presents definitions and purposes of qualitative research; descriptions of predominant designs for qualitative research along with how to select a qualitative design; and discussion of key concepts for qualitative research, including sampling logic, trustworthiness, data analysis, and the researcher's role. The information in this chapter will provide you with baseline understanding of qualitative research. Our hope is that you use this information as a foundation for reading, studying, and practicing qualitative research.

Definition and Purpose of Qualitative Research

Experts provide varying definitions of qualitative research. Several of the most prominent definitions follow. Denzin and Lincoln (2011) defined qualitative research in terms of a *situated activity*.

> Qualitative research is a situated activity that locates the observer in the world. Qualitative research consists of a set of interpretive, material practices that make the world visible. These practices transform the world. They turn the world into a series of representations, including field notes, interviews, conversations, photographs, recordings, and memos to the self. At this level, qualitative research involves an interpretive, naturalistic approach to the world. This means that qualitative researchers study things in their natural settings, attempting to make sense of, or interpret, phenomena in terms of the meanings people bring to them. (p. 3)

Creswell (2013) extended Denzin and Lincoln's (2011) definition by placing more emphasis on the process of qualitative research.

> Qualitative research begins with assumptions and the use of interpretive/theoretical frameworks that inform the study of research problems addressing the meaning individuals or groups ascribe to a social or human problem. To study this problem, qualitative researchers use an emerging qualitative approach to

> inquiry, the collection of data in a natural setting sensitive to the people and places under study, and data analysis that is both inductive and deductive and establishes patterns or themes. The final written report or presentation includes the voices of participants, the reflexivity of the researcher, a complex description and interpretation of the problem, and its contribution to the literature or a call for change. (p. 44)

Merriam (2009) offered a simplified definition of qualitative research: "Basically, qualitative researchers are interested in *understanding meaning people have constructed,* that is, how people make sense of their world and the experiences they have in the world" (p. 13).

In contrast to Merriam (2009), Schwandt (2015) offered a complex definition of *qualitative inquiry*.

> *Qualitative* is a not-so-descriptive adjective attached to the varieties of social inquiry that have their intellectual roots in ***hermeneutics***, ***phenomenological sociology***, and the ***Verstehen*** tradition. Many scholars use the phrase *qualitative inquiry* as a blanket designation for all forms of social inquiry that rely primarily on qualitative data (i.e., data in the form of words). . . . To call a research activity qualitative inquiry may broadly mean that it aims at understanding the ***meaning*** of ***human action***. . . . Because the adjective is used in so many different ways, it does not clearly signal a particular meaning or denote a specific set of characteristics for qualitative research. . . . Broadly speaking, qualitative methods are procedures including unstructured, open-ended interviews and participant observation that generate qualitative data. . . . One, however, could easily generate qualitative data via an open-ended interview, transform those data into numbers, and analyze them by means of nonparametric statistics. Hence, what precisely comprises a so-called qualitative method is not all that clear. (pp. 256–257)

In essence, Schwandt (2015) spotlighted vagueness and breadth in the definition of qualitative research, Merriam (2009) centered on understanding the world, Creswell (2013) focused on the process of conducting qualitative research, and Denzin and Lincoln (2011) emphasized the possibility that qualitative research can transform the world. What sense, then, can the student make of qualitative research, and why do it?

There are some commonalities that can be derived from the various definitions of qualitative research.

1. Qualitative research occurs in natural rather than controlled settings.
2. Qualitative data are collected as words, pictures, or other kinds of aural, visual, or textual artifacts that can be analyzed. Common qualitative data collection techniques include individual and focus group interviews, videotapes, observations, and documents.
3. Qualitative research reports incorporate the participants' voice (the *emic* perspective) into the presentation of findings.
4. Qualitative research describes some phenomenon as experienced by individuals or groups.

The last point above leads us to the purpose of qualitative research. The primary purpose of qualitative research is to describe phenomena that

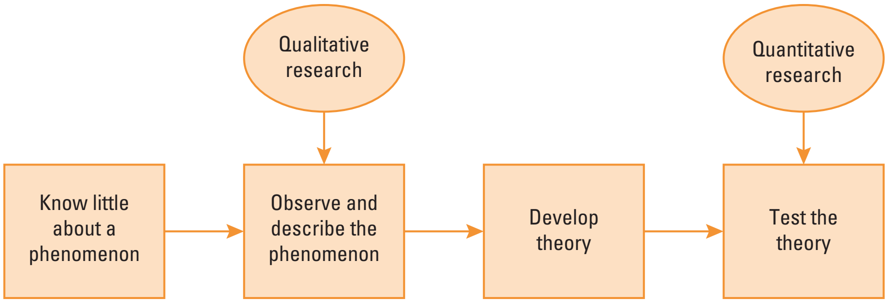


FIGURE 5.1 General process for development of knowledge through research.

occur in the world. By describing what is occurring, we develop complex understandings that can be used to build theories, or explanations, for how phenomena work in the world. The theories that are developed using qualitative research could then potentially be tested using quantitative research. Figure 5.1 displays this general process.

Qualitative research is exploratory, with the function of gaining understanding of complex phenomena through observation and description. Synthesizing the characteristics of qualitative research and its purpose, I define qualitative research as follows:

> *Qualitative research is an exploratory investigation of a complex social phenomenon conducted in a natural setting through observation, description, and thematic analysis of participants' behaviors and perspectives for the purpose of explaining and/or understanding the phenomenon.*

Qualitative research, similar to quantitative research, is structured according to a selected research design. I now explore qualitative research designs and how to select an appropriate design for a qualitative study.

Qualitative Research Design and Design Selection

A research study has an overall structure. That structure or design of a study carries certain conventions that have been accepted over time by researchers to establish rigor. For quantitative studies, there are three main types of research design: nonexperimental, quasi-experimental, and experimental (Campbell & Stanley, 1966). Qualitative researchers, on the other hand, use various terms to designate types of qualitative research, such as *genre* (Marshall & Rossman, 2016), *approach* (Creswell, 2014), *strategy* (Robson, 2011), and *tradition* (Creswell, 1998). Paralleling the quantitative categorization, in this chapter, the term *design* will be used to indicate the overall structure of a qualitative research study.

Qualitative Research Designs and Purposes

All qualitative research is descriptive, but different qualitative designs accomplish description in different ways (Glesne, 2016). There are many variations of qualitative designs, which can be distilled into five main categories commonly presented in qualitative research texts: case study, ethnography, phenomenology, grounded theory, and narrative (Creswell, 2013; Marshall & Rossman, 2016; Robson, 2011). These five designs do not encompass all possible qualitative designs. For example, there are critical, sociolinguistic, and feminist qualitative designs (Marshall & Rossman, 2016). The five main categories of design encompass the kinds of research that scholar-practitioners are likely to perform. Therefore, the discussion in this chapter will focus on those five designs in terms of the purpose, unit of analysis, and predominant data collection tools and strategies. You already are familiar with the terms *purpose* and *data collection*. A new term, *unit of analysis*, is introduced and means the entity around which analysis is organized (Trochim, 2006b). Understanding the unit of analysis for each qualitative research design is one key to selecting the appropriate design for your study.

Case Study. The purpose of a qualitative case study is to describe the interactions of a bounded unit in relation to some phenomenon (Merriam & Tisdell, 2016). There are different schools of thought on what should be emphasized in a case study. Some argue that the process is the most distinct feature of a case study (Yin, 2014); others argue that the bounded unit is the identifying feature of a case study (Merriam & Tisdell, 2016). I hold, with Merriam, that to determine if a study qualifies as a case study, the bounded unit must be identified.

A bounded unit is "a single entity, a unit around which there are boundaries. I can 'fence in' what I am going to study. The case then, could be a single person . . . , a group, an institution, a community or a specific policy" (Merriam, 2009, p. 40). For example, if a study intends to describe how information gathered from patients during hospital admission was processed and used by the hospital for patient care, the study would have a bounded unit—hospital. There could be one case (i.e., one hospital) or several cases (i.e., several hospitals). If, on the other hand, a study intends to describe how patients experience the admission process, regardless of hospital or regardless of how the hospital uses the admission information, then there is not a bounded unit; the focus is on individual experience. In the first instance, the unit of analysis is the hospital, allowing for a case study; in the second instance, the unit of analysis is the individual patient.

Case studies are also differentiated from some of the other qualitative designs by the data collected. Case studies seek to paint a comprehensive picture of a bounded unit around some phenomenon, such as a picture of how one higher education institution responds to allegations of sexual harassment. To do that, multiple data sources are explored, including sources such as interviews, observations, meeting agendas and minutes, policies, reports, and artifacts. A case study, then, is much like a jigsaw puzzle. It requires putting together various pieces to form a composite picture of what is occurring in the bounded unit. Case study is quite close to ethnography, but there are also distinctions between case study and ethnography (Merriam, 2009).

Ethnography. Although both case study and ethnography investigate a bounded unit, ethnography differs from case study in that it requires long-term immersion in a cultural group in order to collect data. Ethnographic research is rooted in anthropology and sociology and has the purpose of interpreting the functioning of a cultural group in relation to some phenomenon. Although case study considers any bounded unit as eligible for study, ethnography insists on a culture-sharing group that has been intact long enough to have discernible patterns of ideas, beliefs, values, behaviors, and rituals (Creswell, 2013). What constitutes a cultural group is not entirely clear, however. Schwandt (2015) reported that over 300 definitions have been published for the term *culture*. Schwandt suggested that culture, for ethnographic research, may be defined "in terms of meaning, symbolism, language, and the discourse drawing respectively on phenomenology, cultural anthropology, structuralism, semiotics, ***cultural studies***, and critical theory" (p. 55). Merriam (2009) was more specific in defining culture as "the beliefs, values, and attitudes that structure the behavior patterns of a specific group of people" (p. 27).

Responding to Merriam (2009), then, a cultural group, the unit of analysis for an ethnographic study, would have a fairly common and discernible set of beliefs, values, and attitudes. For example, an ethnographer might study the communication patterns in a rural community, the power structure within teenage gangs, or gender roles within a certain religious denomination. Because cultural groups are bounded units, similar to case study, you may be wondering how case study and ethnography differ. The distinction lies predominantly in the manner of data collection.

A defining characteristic of ethnography, different from case study, is that the researcher lives within and participates in the culture for a period of time, sometimes for several years, rather than just observing it from collected information, as in a case study (Suryani, 2008). The researcher becomes immersed in the culture. Data sources are not so different from case study data sources (interviews, observations, meeting agendas and minutes, policies, reports, and other artifacts), but they have the added component of intense fieldwork diaries that serve as records of observations made on site. Ethnography requires fieldwork. For example, an ethnographic study of communication patterns in a rural community would require involvement in the community for a period of time, notating, for example, who goes to the coffee shop and who does not go to the coffee shop and what are the topics of conversation in the coffee shop. The researcher would likely also participate in the conversations. Ethnographic and case studies, however, do not explore individual experiences regarding a phenomenon. That is accomplished using phenomenological studies.

Phenomenology. Phenomenological studies seek to understand the lived experiences of a set of individuals who share a common experience. The purpose of a phenomenological study is to describe the lived experiences of individuals in relation to an identified phenomenon (Creswell, 2013).

The unit of analysis for phenomenology is some set of individuals that shares a common characteristic related to the phenomenon of study. For example, for a study of principals' experience of the applicability of their training programs to the real-life work of educational leadership, the researcher would interview a number of principals from multiple locations, rather than in a bounded unit (case study) or cultural group (ethnography). Based on the interviews, the researcher would derive themes that incorporate the essence of shared experiences of the group. In other words, the researcher would collect individual experiences and then discern across those individual experiences a shared, lived experience (Marshall & Rossman, 2016).

The primary data collection tool for phenomenology is in-depth individual interviews, although strategies such as observations and document analyses have also been used (Creswell, 2013; Moustakas, 1994). The individual interviews are analyzed and brought together into a description of shared experiences. If the desired focus is on individuality of individual stories, narrative analysis is the optimal choice.

Narrative. The purpose of narrative research is to tell stories, "first person accounts of experience told in story form having a beginning, middle, and end. Other terms for these stories of experience are biography, life history, oral history, autoethnography, and autobiography" (Merriam, 2009, p. 32). Narrative research does not seek to derive a shared meaning from individual experiences, as does phenomenology, but seeks to understand the meaning of individual experiences in relation to a phenomenon. Marshall and Rossman (2016) noted that narrative research has been criticized as overly focused on the individual rather than the collective experience of a phenomenon, but narrative research has indeed resulted in significant contribution to knowledge.

A fine example of the contributory nature of narrative research is that done by Paige (2007). Paige collected life stories from 13 formative technology users in order to discern the places of self-directed informal learning and formal career development processes in career-directed learning. Using the individual life stories as the unit of analysis, Paige likened self-directed learning to entrepreneurial learning and found that such learning benefits from strategic guidance and self-reflection. He concluded that there is a necessity for instructors, or trainers, to model their own learning process as a potential learning strategy for their students. This modeling, he found, would lead to students' recognizing their own learning strategies and creating their own learning stories. With that developed skill, students would gain more control over their own learning and harness their informal learning experiences for new learning. Based on the results of the narrative study, Paige recommended that corporate trainers incorporate self-directed learning, particularly story-based self-reflection, into formal career development programs. The findings from his award-winning dissertation have been disseminated in a book chapter (Paige, 2008), two articles (Paige, 2009a, 2010a), and three invitational conference presentations (Paige, 2009b, 2010b, 2013). In addition, his work has been cited by others in the field at least eight times.

The unit of analysis for narrative research is the individual. The researcher collects stories from individuals, primarily from interviews, which tell the story of the individual in relation to the phenomenon under study. Both phenomenological and narrative researchers collect information from individuals. One difference between phenomenological and narrative research is that phenomenological research requires the participation of multiple individuals, and even though narrative researchers usually collect stories from multiple individuals, narrative research could involve the story of only one person. A second difference between phenomenological and narrative research is that phenomenological researchers looks for themes among the individual stories, whereas narrative research analyzes each story as a distinct unit. In narrative research, other artifacts can add information to the telling of the story, such as documents, including journals, emails, and letters; photographs; and video and audio records (Marshall & Rossman, 2016). The additional artifacts serve to enhance and support the story told in interviews.

Grounded Theory. As noted earlier, the primary purpose of all qualitative research is to describe phenomena that occur in the world, leading to complex understandings that can build toward theories. Although all qualitative research designs contribute to that purpose, the grounded theory design is specifically focused on theory development. The purpose of the grounded theory design is to "build a substantive theory about the phenomenon of interest" (Merriam, 2009, p. 23).

Although there is consensus that the purpose of grounded theory studies is to develop theory (Corbin & Strauss, 2015; Merriam, 2009; Saveyne & Robinson, 1997), there is less agreement about the unit of analysis within this design. For example, some authors have argued that a set of participants with a shared experience is the unit of analysis (Creswell, 2013), whereas others have proposed that the social situation of interest is the more appropriate unit of analysis (Clarke, 2005). Robson (2011) stated that, for grounded theory studies, "[n]o particular type of 'field' is called for. Such studies have been carried out in a very wide variety of settings" (p. 148). The unit of analysis, then, is not a defining characteristic for grounded theory studies. The purpose of generating theory can be accomplished with a variety of participant sets.

Whatever the participant set, the major data collection strategy for grounded theory studies is interviews, although other forms of data collection, such as observations and documents, may be used (Creswell, 2013; Robson, 2011). Interviews in grounded theory research often are iterative, meaning that the questions can change as interviews progress and new areas of exploration are encountered. An initial interview might query participants on their experience related to the process or activity under study, with emphasis on the sequence of events related to the phenomenon. Follow-up interviews might query participants on specifics of the phenomenon itself, causes for the phenomenon, how the participant interacted with the phenomenon, and what happened as a result (Creswell, 2013).

Summary of Qualitative Research Designs and Design Selection

If a qualitative study is deemed appropriate for a research question study, the next task in crafting a proposal is to select a particular qualitative design. Table 5.1 provides a summary of five qualitative designs and their major purpose, unit of analysis, and primary data collection strategy.

By considering the study problem, purpose, and research question(s) in the context of Table 5.1, you can explore, and ultimately select, the qualitative design that best matches your study. Typically, you will need to consider several designs before deciding on the one that will provide the best overall structure for your intended study. Once a design is chosen, you must consider how to choose study participants, which is a matter of sampling logic.

TABLE 5.1
Summary of Qualitative Research Design Specifications

Qualitative Design	Major Purpose	Unit of Analysis	Primary Data Collection
Case study	Describe behavior of a bounded unit in relation to a phenomenon	Bounded unit	Multiple sources
Ethnography	Describe behavior of a cultural group in relation to a phenomenon	Cultural group	Immersion in culture for an extended period; multiple sources
Phenomenology	Describe themes and patterns of lived experiences across individuals in relation to a phenomenon	Individuals sharing a common experience	Interviews
Narrative	Describe individual stories in relation to a phenomenon	Individuals	Interviews
Grounded theory	Develop theory	Unspecified	Interviews

Sampling Logic

After selecting the research design, your next step would be to identify the sources that will provide information (data) to answer your research question(s). In quantitative research, this step requires identifying the population of interest and drawing a sample from that population sufficient in size and statistical power to represent that population. Having a large enough sample to represent the population is important for statistical analysis done in quantitative research. Quantitative studies use *probability sampling* to represent a population. Qualitative studies, though, are not concerned with representing a population but instead are focused on relevance to the research question(s). Qualitative studies use nonprobability sampling, also called *purposive sampling*, to identify those who can provide data for the study (Schwandt, 2015).

Participant Selection

A distinguishing factor between quantitative and qualitative sampling is the matter of sample size. For statistical analysis in quantitative studies, the size of the sample is important in relation to the size of the population. For qualitative studies, sample size (and hence generalization) is not so important because depth is more important than breadth. Instead, as Schwandt (2015) indicated, there are two critical issues in qualitative sampling logic, or purposive sampling. These are (a) establishing criteria for choosing study participants and (b) describing strategy for determining that selected participants meet the established criteria.

If you were conducting a qualitative study, you would first need to provide the criteria by which participants are eligible to participate. You would need to look closely at the research question and then identify the characteristics needed of study participants to answer that research question. For example, if your research question was "How do female executives in Fortune 500 companies experience work–family balance?" your participant criteria would be

1. Female gender
2. Executive position, and *executive position* would need definition
3. Employed in a Fortune 500 company, and *Fortune 500 company* would need definition
4. Family association, and *family association* would need definition

Defining terms is very important. For example, does *family* include only women who have a spouse and/or children, or does it include single women with no children but obligations to parents or other relatives? In this example, only women who meet the definitions of executive in a Fortune 500 company and who have a family would be eligible for the study. Notice that the research question does not ask about work–life balance but about work–family balance, so the research question specifies the participant pool. Clarity in specification of participant criteria supports study rigor.

With participant criteria clearly delineated, you would then need to describe how you find appropriate participants. Miles and Huberman (1994) summarized purposive sampling strategies, such as snowballing (asking the current participant for a referral to a next participant), convenience (using those who are readily available), and opportunistic (capitalizing on unexpected leads). Which sampling strategy to use depends largely on feasibility. Feasibility requires that there is a sufficient number of people meeting your participant criteria available to you and that you truly can access them. If, for example, you know of only a few people who meet participant criteria, but those few may know others, then snowballing may be appropriate. If, for example, the study requires access to financial records of a company, and those records are not publicly available, you must determine if it is feasible to obtain those records with permission of the company.

The mention of *records* in the last paragraph is important. In qualitative research, documents may serve as data sources as well as people. For example, case study procedures might include examining meeting agendas and minutes, policy statements, calendars, and other records. If you decide to collect such data sources, you will need to explain your rationale for doing so and your criteria for including those sources. The same requirements apply for including these data sources as they do for including human participants in your research.

At this introductory stage, it is not necessary to understand all the purposive sampling strategies. It is important to understand that qualitative sampling is based on relevance, rather than representativeness, and that, as a researcher, you must describe the criteria for inclusion of data sources in the study and how those data sources are selected, with consideration for their availability and accessibility. That does leave the question, though, of how many participants, documents, or other sources should be included in a qualitative study. That question is addressed through understanding the concept of *saturation*.

Saturation

In quantitative studies, sample size is dictated by a formula in relation to population size; the goal is generalization to the larger population. Qualitative studies are not designed to generalize to a population but to *describe* some unit of analysis. Important questions in qualitative research include the amount of data you should collect, from how many people and from how many documents or observations. *Saturation* is the relevant concept here. Saturation happens when continued data collection does not add new themes or patterns but instead reinforces what has already been derived from prior data analysis. Two criteria have to be achieved in order to reach saturation: (a) continued analysis yields no new information and (b) there are no unexplained phenomena (Marshall & Rossman, 2016; Saumure & Given, 2008). It is easy to think that you have achieved saturation when no new information is being derived from the data but remember that the second condition has to be fulfilled as well. For the second condition to be fulfilled, the data gathered must be sufficient to answer the research question(s).

This vagueness makes it difficult to decide on the specific number of people from whom data will be collected. You cannot always predict the number of people to be interviewed or documents to be reviewed to reach saturation. You should, though, provide an estimate of the number of people, documents, observations, or hours of recorded transcripts that are expected, with the option of continuing on if saturation is not attained or stopping earlier if saturation is already attained (Merriam, 2009; Patton, 2015). The estimate can be informed by reading other studies that have used similar designs and/or addressed the same or similar topics. You should also provide a description of how saturation will be recognized, which is a matter of design trustworthiness.

Trustworthiness in Qualitative Studies

Quantitative researchers rely on the concepts of reliability, internal and external validity, and replicability to legitimate their studies. The counterparts for qualitative researchers are dependability, credibility, transferability, and confirmability (Guba & Lincoln, 1989; Lincoln & Guba, 1985). In this section, I first define the four terms presented in the previous sentence; after that, I describe strategies related to evidencing trustworthiness of a qualitative study.

Definitions

Dependability. Dependability in qualitative research is akin to reliability in quantitative research. Reliability means that the instruments used to collect data produce consistent results across data collection occurrences.

For example, a metal tape measure will likely be more reliable than a cloth tape measure because a metal tape measure will not stretch as will a cloth one. Reliability can be estimated statistically. In qualitative studies, the concept of statistical reliability does not apply. Instead, qualitative studies must meet a standard of dependability. Dependability means that there is evidence of consistency in data collection, analysis, and reporting. It also means that any adjustments or shifts in methodology, which can occur in qualitative studies, are documented and explained in a fashion that is publicly accessible. Inquiry audits and triangulation are the most common methods of establishing dependability (Guba & Lincoln, 1989; Merriam, 2009).

Credibility. Internal validity in quantitative research confirms that the data collected matches the research question. In other words, if the research question is about motivation, the data collected must be about motivation. Credibility is the parallel concept for qualitative studies. Credibility means that the findings of the study are believable given the data presented (Merriam, 2009). Credibility is established using strategies such as prolonged engagement, persistent observation, peer debriefing, negative case analysis, progressive subjectivity, member checking, triangulation, and reflexivity (Guba & Lincoln, 1989; Merriam, 2009). If you were conducting a qualitative study, you would not need to use all the strategies related to credibility; rather, you would select the strategy or strategies most appropriate for your study.

Transferability. Transferability relates to the quantitative concept of external validity. External validity provides a measure of the extent to which the findings of the study, based on a sample, are generalizable to the population of interest for the study. Even though the purpose of qualitative research is not to generalize from a sample to a population, a qualitative study must have some meaning beyond the immediate instance of the study. Determining the applicability of the findings of a qualitative study to other situations is predominantly a responsibility of the person in the other situation. The researcher's responsibility in regard to transferability is to provide sufficient description of the setting and the assumptions of the study so that a reader can make informed application of the findings of the study (Trochim, 2006a). Transferability is supported by using thick descriptions and maximum variation (Merriam, 2009).

Confirmability. Quantitative research aims at objectivity, extracting the researcher from the study as much as possible so that the findings of the study are disassociated from any researcher bias. Qualitative research admits researcher subjectivity, but its methods must be based on verifiable procedures, analyses, and conclusions. In other words, confirmability requires that other informed researchers would arrive at essentially the same conclusions when examining the same qualitative data (Guba & Lincoln, 1989). Guba and Lincoln proposed a confirmability audit as the primary means of establishing confirmability.

Strategies to Evidence Trustworthiness

As identified earlier with each definition, there are multiple strategies available to establish trustworthiness of a study. Table 5.2 summarizes

TABLE 5.2
Summary of Trustworthiness Criteria and Related Evidentiary Strategies

Trustworthiness Criterion	Strategies
Dependability	• Inquiry audit • Triangulation
Credibility	• Prolonged engagement • Persistent observation • Peer debriefing • Negative case analysis • Progressive subjectivity • Member checking • Triangulation • Reflexivity
Transferability	• Reflexivity • Thick description • Maximum variation
Confirmability	• Confirmability audit

the strategies associated with each of the four aspects of trustworthiness defined earlier.

In this section, I will briefly describe each of the strategies listed in Table 5.2.

Audit Trails. I consider inquiry and confirmability audits together under the idea of audit trails. “An audit trail in a qualitative study describes in detail how data were collected, how categories were derived, and how decisions were made throughout the inquiry” (Merriam, 2009, p. 223). Audit trails are derived from field notes as well as from memos or reflection journals about decisions made during the research process. “Essentially, it [the audit trail] is a detailed account of how the study was conducted and how the data were analyzed” (Merriam, 2009, p. 223). These reflective documents become available to other analysts to assess the dependability and confirmability of the study.

Triangulation. Triangulation is the use of more than one source to verify the basis of a claim. Although researchers sometimes limit the concept of triangulation to multiple data sources, triangulation can also be established using “multiple investigators, multiple theoretical perspectives, or multiple methods. The central point of the procedure is to examine a conclusion (assertion, claim, etc.) from more than one vantage point” (Schwandt, 2015, p. 307). Triangulation might imply *three*, but it should be interpreted as *multiple*, which is more than one but not necessarily three.

Prolonged Engagement. Ethnographic studies, in particular, rely on prolonged engagement. Prolonged engagement requires presence and involvement at the study site for an extended period of time, often several years. There are three purposes for prolonged engagement: (a) to build rapport and trust, (b) to derive enough information to mitigate misunderstandings, and (c) to understand more deeply the context and culture of the study environment (Guba & Lincoln, 1989).

Persistent Observation. Persistent observation accompanies prolonged engagement. It means that the observations recorded are sufficient to provide depth of understanding through collection of details (Guba & Lincoln, 1989). This definition is vague because what is sufficient is a matter of judgment, but consider persistent observation in alignment with the concept of saturation described earlier.

Peer Debriefing. Guba and Lincoln (1989) proposed peer debriefing as a means of establishing credibility of a qualitative study. Peer debriefing involves engagement with a qualified colleague, someone who is not involved in the study, in ongoing discussions of study progress, data analyses, and tentative findings. The role of the peer is to pose questions that assist you in clarifying conclusions and excising researcher bias. A key to peer debriefing is rapport between you and the peer as well the peer's own experience as a qualitative researcher, as both will inform the kinds of questions the peer poses.

Negative Case Analysis. In quantitative studies, most, but not all, pieces of data collected support the overall findings of the study. As a parallel, for qualitative studies, you would need to note data collection instances that are divergent from the majority of derived themes and patterns. You would notate the strength of such negative cases. If the presence of such divergent cases is weak, then the derived patterns and themes receive more confidence (Guba & Lincoln, 1989). You should, however, be vigilant that a divergent, or negative, case may point in a direction that requires further study.

Progressive Subjectivity. Although subjectivity is inevitable in qualitative research, it is not appropriate to let this subjectivity run unchecked. You must acknowledge your subjectivity by revealing, monitoring, and controlling its influence on data collection and interpretation, a process known as progressive subjectivity. This progressive subjectivity is recorded in three steps: (a) prior to data collection, you would record your conceptualization and expectations; (b) during the study, you would regularly note developing conceptualizations and expectations; and (c) a peer reviewer would examine your notations and challenge your interpretations, as needed (Guba & Lincoln, 1989).

Reflexivity. Progressive subjectivity is closely associated with reflexivity. Reflexivity obliges you to document in field notes, memos, or journals your self-critical analyses of biases, your role in and responses to the research process, and any adjustments made to the study based on ongoing analysis (Schwandt, 2015).

Member Checking. Qualitative researchers sometimes interpret member checking only as having study participants review and confirm their interview transcripts, but member checking is more than transcript review. Member checking, also called respondent validation, "is systematically soliciting feedback about your data and conclusions from the people you are studying" (Maxwell, 2013, p. 126). Notice that member checking involves not only having participants examine data, such as transcripts, but also getting feedback from participants on findings as they emerge. As Merriam (2009) wrote, "The process involved in member checks is to take your preliminary analysis back to some of the participants and ask whether your interpretation 'rings true'" (p. 217).

Thick Description. Thick description means that the research report contains "a description of the setting and participants of the study, as well as a detailed description of the findings with adequate evidence presented in the form of quotes from participant interviews, field, notes, and documents" (Merriam, 2009, p. 227). There are, then, three elements to thick description: description of the setting, description of the participants, and evidence to support the findings. You must use your own judgment to determine what is adequate, and not too little or too much, evidence to support findings. Peer debriefing can support that judgment.

Maximum Variation. Merriam (2009) discussed maximum variation as a means of strengthening transferability. Maximum variation is a sampling strategy that intentionally diversifies study participants in order to create greater applicability of the study to a variety of situations. This strategy may not be useful if the study is focused on understanding a particular set of participants.

Thus far in this chapter, I have defined qualitative research and its purposes, examined qualitative study designs, and considered how to establish the trustworthiness of a qualitative study. In the next section, I introduce qualitative data analysis.

Qualitative Data Analysis

Different from quantitative studies in which data are collected as numbers and analyzed using statistics, data from qualitative studies are primarily textual in the form of words, video, and other kinds of artifacts (recognizing that these data could be converted to numerical form and analyzed descriptively). Because words are the primary data for qualitative studies, in this discussion, I will focus on how verbal data are analyzed in qualitative studies.

The primary procedure for analyzing qualitative data is coding. Coding is the process of organizing qualitative data by recognizing patterns in the data, creating categories for the patterns, determining interconnectedness among the categories, and then synthesizing the interconnected categories into a cohesive understanding of the unit of analysis in relation to the phenomenon. There are three stages to this process: open coding, axial coding, and selective coding (Leedy & Ormrod, 2013). Benaquisto (2008) succinctly explained these three stages:

1. Open coding: deriving themes and labeling categories from raw data
2. Axial coding: reassembling categories
3. Selective coding: identifying a core category around which other categories may integrate

Open Coding

After collecting data, transcribing interviews, and otherwise organizing the qualitative data, you begin the data analysis using open coding. In open coding, your goal is to look for common themes that are emerging and what kinds of categories emerge from their organization (Corbin & Strauss, 2015). For example, if the study is about faculty satisfaction with online curriculum, and multiple participants state lack of flexibility as an issue, then that might be recognized as a theme. You might create a category called *flexibility* to fit the code. A theme is what you first recognize as a commonality across participants; a category is the label you give to the theme. You might begin with some ideas for categories, but if relevant themes do not emerge from the data, those initial categories would need to be discarded. Open coding results in a set of categories derived from themes substantiated by the collected data.

Axial Coding

Axial coding is the process of discerning the interconnections among the categories created in open coding. Once the initial set of categories is determined, you would proceed through an iterative process of discerning interconnections among the categories (Corbin & Strauss, 2015). These interconnections may be derived from the conditions that create the category, the context of the categories, strategies people use to manage the categories, or the effects of the strategies people use to manage the categories (Leedy & Ormrod, 2013). Axial coding refines the categories created in open coding by combining several categories into one, subdividing one category into several, or creating a new category. In relation to the earlier example, you might see an interconnection between flexibility and participation in curriculum development. Therefore, you might create a new category that relates to both flexibility and participation. Axial coding is not a straight-line process. As you become more immersed in the data and see the interconnections, you will need to refine and revise the initial categories.

Selective Coding

Selective coding is the stage at which you synthesize the understandings developed in open and axial coding into a cohesive statement of findings from the study (Corbin & Strauss, 2015). Selective coding tells the story of the interaction between the participants and the phenomenon. It can lead to theory development because it produces hypotheses about how the phenomenon and the participants interact (Leedy & Ormrod, 2013; Merriam, 2009). Following from the earlier example, you might identify participation as a core category for faculty satisfaction with online curriculum, and you would integrate other emerged categories within that core.

Variations on Coding in Qualitative Studies

Coding is the primary form for analyzing qualitative data, but the coding procedures can vary according to the purpose of the study. Four main types of qualitative data analysis are thematic, conversation, narrative, and semiotics

(Glesne, 2016). Thematic analysis focuses on deriving themes by identifying similarities and differences among the categories derived in the coding process. Conversation analysis attends not only to the words spoken but also to the inflections, spacing, and other subtle interaction indicators. For example, Plum (2008) used conversation analysis to explore power distribution in conversations within special education child study team meetings. Narrative analysis inserts the coded themes into historical or other contexts of importance to the participants (Gubrium & Holstein, 2009). Semiotic analysis incorporates symbols into the coding process, including visual signs, such as colors; aural signs, such as ambient sounds; and linguistic signs, such as terminology, to construct meaning (Glesne, 2016). Those interested in analysis strategies tailored to specific designs should refer to an appropriate reference (such as those provided in the Key Sources at the end of this chapter).

Analytical Tools

Although qualitative data may be analyzed manually, doing so is becoming a less common approach. Just as most quantitative researchers use analytical software, most qualitative researchers now use analytical software. There are many free and for-purchase qualitative data analysis tools available. These tools do not supplant the need for you to create categories in open coding, rearrange them in axial coding, and synthesize them in selective coding. However, they can be helpful in managing large amounts of textual data, and I recommend their use. You are conducting the analysis; the software is only an assistive analytical tool.

Role of the Researcher

If you are undertaking a qualitative study, then you as the researcher will serve as the primary instrument of qualitative data collection. You serve as this instrument through your direct observations, your participation in interviews, and your analysis of documents. This role as primary instrument of data collection brings you into an intimate relationship with your setting, your participants, and your data analysis, yielding a duality to your presence in a study as both a participant and an observer in varying degrees ranging from complete participant to complete observer (Creswell, 2013).

Complete Participant

If you are a complete participant in a qualitative research study, you will behave as a full member of the unit of analysis. For example, if you were a complete participant in a study of neighborhood watch groups, you would become a member of the watch group, live in the neighborhood, and engage in neighborhood activities as any other community member would. Benefits of the complete participant role include very deep understanding of how the participants experience the phenomenon of focus for the study and the development of trusting relationships with participants. Challenges with complete participation include potential loss of perspective and difficulty recording observations accurately while fully engaged in activities.

Participant Observer

Another variation of the researcher role is the participant observer role. In the participant observer role, you share actively with the study participants but also step back from the activities so as to be an observer. For example, in a study of how members of a volunteer community group negotiate task distribution, you might join the group and participate in some activities, but you would also step out to observe other activities and conduct interviews. Benefits and challenges of the participant observer role are similar to those for the complete participant role but may be mitigated somewhat by the balancing of both roles, with weight toward the participant role.

Observer Participant

With the observer participant role, the balance shifts to the observer side. As observer, you would not engage in activities with your participants. As participant, you would be present in the setting, and the study participants would be aware of your direct observation. For example, for a case study of how school board members manage conflict, you might attend board meetings, interview board members, and examine related documents, but you would not participate in board deliberations, make presentations to the board, or otherwise interact with the board. Benefits of the observer participant role include facility of recording data contemporaneously with observations and the opportunity to collect data within settings and from groups to which the researcher does not have access of full membership. Deficits of the observer participant role include loss of information that comes from full engagement and potential influence of observer presence on participant behavior.

Complete Observer

The researcher in the complete observer role is not present in the setting or with the participants. The study participants do not see or hear you, and you either collect data remotely, through technology, or from archival documents. For example, in a study of the visible and audible reactions of patients to a painful medical procedure, you might collect videotapes, with permission of course, and later analyze the body, facial, and verbal language exhibited. Benefits of the complete observer role include the opportunity to collect data when researcher presence would be intrusive or the setting is at a distance. Another benefit is the absence of influence of the researcher's presence on participant behavior. Deficits of the complete observer role include loss of information from engaging directly with study participants and inability to use some data collection strategies, such as interviews and focus groups.

Each of the four research roles requires particular skills in order to be used effectively, such as how to observe a situation without influencing it and how to be fully engaged in an activity and yet collect data regarding it. Whichever role you assume depends on the circumstances of the study. You must disclose your selected role with a clear and well-articulated rationale. If the role changes during the course of the study, you must document that change and the rationale for the change.

Summary

In this chapter, I provided an overview of qualitative research. After defining and stating purposes for qualitative research, I offered guidance in selection of a qualitative research design by defining five designs and their major purpose, unit of analysis, and primary data collection strategy. Finally, I discussed the key concepts of sampling logic, trustworthiness, data analysis, and researcher role and addressed further development and understanding of how to implement qualitative research. If you are pursuing a qualitative research design, I advise you read the sources in this chapter's Key Sources section relevant for that design and also read multiple studies that have used the design you are considering. Many students think that because qualitative research analyzes words rather than numbers it is easier to do, and they gravitate toward it due to that assumption. Qualitative research is not easy. A substantial amount of practice is needed to conduct interviews, record observations, and analyze qualitative data.

In the next chapter, Dr. Bonnie Nastasi describes mixed methods studies. She draws upon the information from Chapters 4 and 5 to focus on the important issues relevant to mixing of methods. Framing the discussion is the assumption that mixed methods research can facilitate the exploration of complex research questions such as those related to translating research to practice and policy, thereby extending the potential contributions of qualitative or quantitative methods alone. Dr. Nastasi presents the essential issues for those considering mixed methods research and, in the process, describes key features of mixed methods study designs.

Acknowledgment

The author wishes to acknowledge the contribution of Deanna Boddie, PhD, to this chapter's development of the definition of qualitative research.

KEY SOURCES

Case Study

Stake, R. E. (1995). *The art of case study research.* Thousand Oaks, CA: Sage.

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Thousand Oaks, CA: Sage.

Ethnography

Agar, M. H. (1996). *The professional stranger: An informal introduction to ethnography* (2nd ed.). San Diego, CA: Academic Press.

Geertz, C. (1973). *The interpretation of cultures*. New York, NY: Basic Books.

Madison, D. S. (2005). *Critical ethnography: Method, ethics, and performance.* Thousand Oaks, CA: Sage.

Wolcott, H. F. (1999). *Ethnography: A way of seeing.* Thousand Oaks, CA: Sage.

Phenomenology

Moustakas, C. E. (1994). *Phenomenological research methods.* Thousand Oaks, CA: Sage.

Van Manen, M. (1990). *Researching lived experience: Human science for an action sensitive pedagogy.* Albany, NY: State University of New York Press.

Narrative

Clandinin, D. J., & Connelly, F. M. (2000). *Narrative inquiry: Experience and story in qualitative research*. San Francisco, CA: Jossey-Bass.

Czarniawska, B. (2004). *Narratives in social science research*. Thousand Oaks, CA: Sage.

Grounded Theory

Corbin, J., & Strauss, A. (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (4th ed.). Thousand Oaks, CA: Sage.

Glaser, B. G., & Strauss, A. L. (1999). *The discovery of grounded theory: Strategies for qualitative research*. New Brunswick, NJ: Aldine Transaction.

REFERENCES

Benaquisto, L. (2008). Codes and coding. In L. M. Given (Ed.), *The Sage encyclopedia of qualitative research methods* (pp. 86–89). Thousand Oaks, CA: Sage.

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

Clarke, A. E. (2005). *Situational analysis: Grounded theory after the postmodern turn*. Thousand Oaks, CA: Sage.

Corbin, J., & Strauss, A. (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (4th ed.). Thousand Oaks, CA: Sage.

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage.

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage.

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

Denzin, N. K., & Lincoln, Y. S. (2011). Introduction: The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (4th ed., pp. 1–19). Thousand Oaks, CA: Sage.

Glesne, C. (2016). *Becoming qualitative researchers: An introduction* (5th ed.). Boston, MA: Pearson.

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.

Gubrium, J. F., & Holstein, J. A. (2009). *Analyzing narrative reality*. Thousand Oaks, CA: Sage.

Leedy, P. D., & Ormrod, J. E. (2013). *Practical research: Planning and design*. Boston, MA: Pearson.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.

Marshall, C., & Rossman, G. B. (2016). *Designing qualitative research* (6th ed.). Thousand Oaks, CA: Sage.

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). Thousand Oaks, CA: Sage.

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: Jossey-Bass.

Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). San Francisco, CA: Jossey-Bass.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.

Moustakas, C. E. (1994). *Phenomenological research methods*. Thousand Oaks, CA: Sage.

Paige, R. D. (2007). *The relationship between self-directed informal learning and the career development process of technology users* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (Order No. 3244821). Retrieved from http://search.proquest.com/docview/304762834?accountid=14872

Paige, R. D. (2008). The relationship between self-directed informal learning and the career development process. In H. Nilsen & E. Perzycka (Eds.), *The teacher for the knowledge society* (pp. 71–95). Nesna, Norway: Organ for Research Publications.

Paige, R. D. (2009a, March). *Self-directed learning and the mind-set of successful entrepreneurial learning*. Paper presented at ATEA's 46th National Conference on Technical Education, Chattanooga, TN.

Paige, R. D. (2009b, May). *Self-directed learning and the mind-set of successful entrepreneurial learning*. Paper presented at NISOD's International Conference on Teaching and Leadership Excellence, Austin, TX.

Paige, R. D. (2010a). Beyond student-centered instruction: A model for teaching learning-to-learn strategies. *International Journal of Interdisciplinary Social Sciences*, *5*(5), 299–308.

Paige, R. D. (2010b, August). *Beyond student-centered instruction: A model for teaching learning-to-learn strategies*. Paper presented at the Fifth International Conference on Interdisciplinary Social Sciences, Cambridge, United Kingdom.

Paige, R. D. (2013, April). *The entrepreneurial learner: How people learn-to-learn and learn to apply what they have learned*. Paper presented at the 18th Annual Instructional Technology Conference, Middle Tennessee State University, Murfreesboro, TN.

Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). Thousand Oaks, CA: Sage.

Plum, C. (2008). *Interaction within individualized education program meetings: Conversation analysis of a collective case study* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (Order No. 3342449). Retrieved from http://search.proquest.com.ezp.waldenulibrary.org/docview/276002378/73A49141A19F4CA4PQ/4?accountid=14872

Robson, C. (2011). *Real world research* (3rd ed.). West Sussex, United Kingdom: Wiley.

Saumure, K., & Given, L. M. (2008). Data saturation. In L. M. Given (Ed.), *The Sage encyclopedia of qualitative research methods* (Vol. 2). Thousand Oaks, CA: Sage.

Saveyne, W., & Robinson, R. (1997). Qualitative research issues and methods: An introduction for educational technologists. In D. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 1171–1195). New York, NY: Palgrave Macmillan.

Schwandt, T. A. (2015). *The Sage dictionary of qualitative inquiry* (4th ed.). Thousand Oaks, CA: Sage.

Suryani, A. (2008). Comparing case study and ethnography as qualitative research approaches. *Jurnal Ilmu Komunikasi*, *5*(1), 117–127.

Trochim, W. M. K. (2006a). Qualitative validity. In *Research methods knowledge base* (2nd ed.). Cincinnati, OH: Atomic Dog. Retrieved from http://www.socialresearchmethods.net/kb/unitanal.php

Trochim, W. M. K. (2006b). Unit of analysis. In *Research methods knowledge base* (2nd ed.). Cincinnati, OH: Atomic Dog. Retrieved from http://www.socialresearchmethods.net/kb/unitanal.php

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Thousand Oaks, CA: Sage.

6

Mixed Methods Designs and Approaches

Bonnie K. Nastasi, PhD

Introduction

A major premise of this chapter is that research methods can be portrayed on a continuum from quantitative to qualitative, with mixed methods somewhere in the center. Chapters 4 and 5 address quantitative and qualitative methodologies. This chapter draws on that background to focus on the important issues relevant to mixing of methods. Framing the discussion is the assumption that mixed methods research can facilitate the exploration of complex research questions such as those related to translating research to practice and policy, thereby extending the potential contributions of qualitative or quantitative methods studies alone.

An important first step in decision making about research methods is consideration of the purpose of research. Newman, Ridenour, Newman, and DeMarco (2003) provided a typology of nine research purposes with specific connections to methods. The purposes that call for quantitative methods include prediction, extending knowledge, measuring change, and testing new ideas. Those that call for qualitative methods include understanding complex phenomena; generating new ideas; examining the past; and having personal, social, or organizational impact. Newman proposed that a ninth purpose, informing constituencies, relies on both qualitative and quantitative methods. Furthermore, he proposed that several purposes could be extended through the use of mixed methods: extending knowledge base; understanding complex phenomena; generating or testing new ideas; and having personal, social, or organizational impact.

Given this differentiation of method by purpose, how do we distinguish quantitative, qualitative, and mixed methods? Put simply, quantitative methods are based in the hypothetical-deductive and postpositivist traditions of science and rely on statistics to represent the occurrence of phenomena and to test predictions (hypotheses) based on deductions from existing research and theory. Qualitative methods are based on the constructivist and interpretive traditions and rely on verbal and visual representations (e.g., through narrative, observations, artifacts). In contrast to the deductive approach in quantitative research, qualitative research is an inductive form of inquiry. Finally, quantitative research embodies an etic perspective based on existing theory and research as interpreted by the researcher, and qualitative research is characterized by an emic perspective designed to capture the perspective of those being studied and/or identify theory grounded in real-life experience. Mixed methods research, characterized by the mixing of quantitative and qualitative, is intended to capitalize on the strengths of each of these traditions in order to extend our understanding of phenomena and to inform science, practice, and policy through the integration of emic and etic perspectives.

Given the coverage of quantitative and qualitative methods in preceding chapters, in this chapter, I focus only on issues related to the mixing of methods. Thus, the purpose of this chapter is to provide an introduction to mixed methods research (MMR), address key philosophical and design issues, and answer questions about MMR and why researchers choose to mix qualitative and quantitative methods. The chapter begins with an explanation of mixed methods research and why researchers choose a mixed methods approach. Following this is a discussion of philosophical underpinnings of MMR. Next is a discussion of types of MMR designs, followed by a discussion of the process of integrating quantitative and qualitative findings.

What Is Mixed Methods Research?

MMR, stated simply, is the integration of quantitative and qualitative data collection and analysis to answer specific research questions; it can be viewed as the center of the quantitative–qualitative continuum. However, MMR is much more complex than merely the collection of both qualitative and quantitative data in a single study. In the last two decades, methodologists have engaged in discussion and debate about what constitutes MMR, the types of MMR, when and how to apply mixed methods, and the interplay between the researcher's philosophy and choice of methods, among other issues. The result of these discussions is reflected in publication of two editions of the *Handbook of Mixed Methods in Social and Behavioral Research* (Tashakkori & Teddlie, 2003, 2010), the *Journal of Mixed Methods Research*, and a plethora of books and articles. The writings in the field not only address philosophical and design issues but also provide examples of applications of MMR. For methodologists, the consideration of the questions surrounding MMR can be overwhelming and impede its use. At the same time, lack of understanding of the complexity and expected rigor can lead to poorly designed studies. Planned and conducted in a systematic way, MMR can enhance our understanding of phenomena in ways that extend beyond the use of qualitative or quantitative designs alone.

Why Do Researchers Use Mixed Methods?

The decision to use MMR depends on your purpose. Newman et al. (2003) and others (e.g., Tashakkori & Teddlie, 2003) have argued that purpose drives the choice of design and methods. Greene, Caracelli, and Graham (1989) identified purposes for using MMR in evaluation research: triangulation (seeking convergence), complementarity (seeking clarification or elaboration), development (informing later stages of research program), initiation (seeking new perspectives), and expansion (extending breadth and depth). More complex schemes for classifying purpose have been suggested as well, for example, Bryman's (2006) 16-type scheme for classifying research purpose. As a general rule, research purpose and questions determine the choice of design and methods. For those of you choosing MMR, your decision must stem from research questions that cannot be answered by quantitative or qualitative methods alone, as illustrated in Box 6.1.

Morgan (2007) extends the considerations around research purpose and draws distinctions among qualitative, quantitative, and MMR based on the connections between theory and data, researcher and research process, and data and inference, which all follow from the research purpose:

1. *Theory–data connection*. In qualitative research, data analysis can be described as an *inductive* process; that is, the meaning of data is derived from the viewpoint of the participants. For example, children might describe what they view as psychological well-being. In quantitative research, data analysis is *deductive*, such that the meaning of data depends on existing theory and research. For example, hypotheses are derived by existing empirically supported theory. In MMR, researchers adopt an *abductive* approach, in which they integrate meanings of constructs, such as psychological well-being, based on an integration of the inductive and deductive reasoning. Thus, MMR researchers modify existing definitions based on perspectives of research participants as reflected in data.
2. *Researcher–research relationship*. In qualitative research, the researcher assumes a *subjective* stance in regard to the research process. For example, the researcher's perspective influences research questions and data collection and analysis. In quantitative research, the researcher assumes an *objective* stance, thus attempting to engage in the research process without personal bias; instead, existing theory might guide the development of research questions, data collection, analysis, and so on. MMR researchers assume an *intersubjective* perspective by attempting to integrate an objective stance drawn from existing theory with a subjective stance drawn from personal worldview.
3. *Data–inference relationship*. In qualitative research, the interpretation of data is assumed to be specific to the *context* in which the research occurred, for example, a specific elementary school. In quantitative research, the data are interpreted with an assumption of *generality*; that is, the findings can be applied to elementary schools in general. In MMR, findings are assumed to be *transferable* to other similar settings, for example, to elementary schools with similar contextual features (demographics, location).

These descriptions of MMR characterize the method as a complex yet powerful approach to research. If your research interests are aligned with an MMR approach, you need to integrate the inductive approach with one that is also deductive. You need to seek an intersubjective connection between yourself as the researcher and your topic of interest. Last, you need to characterize your data inference connection as one that is transferable between context-specific settings and one that is general. All of these considerations must be addressed. For further detail, see Morgan (2007) and Nastasi, Hitchcock, and Brown (2010).

BOX 6.1

Illustration of the Value of Qualitative and Quantitative Methods in MMR

Hausman et al. (2013) describe the use of MMR in the development of contextually relevant measures of outcomes in violence prevention programs. Through a participatory process of involving stakeholders in the community, the researchers developed community-defined constructs relevant to established measures. The qualitative data, gathered through focus groups, led to the development of community-defined indicators of the effectiveness of violence prevention programs. These indicators were then matched to items from existing, previously validated measures to inform the development of new culturally relevant measures; this was followed by quantitative data collection (administration of new measures) and factor analysis to establish instrument validity. The process resulted in development of new culturally relevant constructs not reflected in preexisting measures. This research demonstrates the value of mixed methods for enhancing the cultural relevance of existing psychological measures. The qualitative research component helped to understand key constructs from a cultural perspective, understanding that may not have been evident otherwise. This then led to developing a quantitative instrument to gather data to represent the population of interest. Using the modified instrument, the researcher could have more confidence that the constructs being measured were appropriate to the beliefs, values, and normative behaviors of, and thus valid for, the target population. For additional examples of the use of MMR in instrument development, see Hitchcock and Nastasi (2011), Hitchcock et al. (2005, 2006), and Nastasi, Hitchcock, Burkholder, et al. (2007).

In addition to considerations about research purpose, your philosophical and theoretical perspectives are also critical. Indeed, MMR has been identified as a third paradigm to distinguish it from quantitative (first) and qualitative (second) research. I explore these issues in the next section.

Philosophical Underpinnings of MMR: What Is the Third Paradigm?

Underlying methodological decisions are "philosophical assumptions that researchers bring to their inquiries . . . [assumptions about] what knowledge warrants our attention, how knowledge is learned, the nature of reality and values, and also the historical and political perspectives that individuals bring to research" (Creswell & Tashakkori, 2007, p. 305). R. Johnson and Onwuegbuzie (2004) defined a *research paradigm* as "a set of beliefs, values, and assumptions that a community of researchers has in common regarding the nature and conduct of research . . . a research culture" (p. 24).[1] Applying

[1] R. Johnson and Onwuegbuzie (2004) and Denscombe (2008) describe the MMR as a paradigm that depicts a *research community*, consistent with the thinking of Kuhn (1962). As discussed in Denscombe, the four characteristics of paradigm (according to Kuhn) are (a) a specific set of problems that are important to advancing science, (b) shared practices or research techniques, (c) shared identity among researchers, and (d) a community of researchers.

the concept of research culture, they describe MMR as the *third paradigm* with philosophical foundations that distinguish it from the first (quantitative) and second (qualitative) paradigm. The philosophical foundations of quantitative research are positivist or, more recently, postpositivist, with a focus on objectivity, context-free generalizability, and causality. The foundations of qualitative research are constructivist or interpretive, with a focus on subjectivity, context specificity, and multiple constructions of reality. The most common depiction of the third (alternative) paradigm for MMR is *pragmatism* (see R. Johnson & Onwuegbuzie, 2004; Maxcy, 2003; Morgan, 2007; Nastasi et al., 2010; Tashakkori & Teddlie, 1998, 2003). Morgan (2007) described pragmatism as grounded in *practical* questions that include purpose (Is MMR the best way to answer the research questions?), context (What is appropriate given the setting?), and consequences or implications (Will the research lead to change at a social or policy level?).

Two other perspectives or alternative paradigms for MMR have been proposed. Greene (2007), drawing on the intersubjective nature of MMR, has offered a *dialectical* perspective. Critical to this conceptual foundation is the potential divergence or conflict of views between researcher and researched and the focus on integration of these. For example, in participatory research, one might focus on integrating the etic perspective of the researcher and the emic perspective of the participants (Nastasi, Hitchcock, Sarkar, et al., 2007). The dialectical perspective is also inherent in Denscombe's (2008) *communities of practice* model, involving partnership of researchers and practitioners (e.g., teachers). Critical to the dialectical perspective is the value of MMR that results from the potential conflict between subjective and objective views inherent in qualitative and quantitative data, respectively. Applied to Denscombe's communities of practice, the tension between researchers and practitioners can facilitate a process that leads to thinking informed by both stakeholder groups. The differences in perspectives can challenge the etic view of researchers, that is, their worldview informed by theory and research. These differences can also make explicit the practitioners' emic views, worldviews informed by their experiences and understanding of the context. The integration of etic and emic views can help to ensure the application of research to practice and alter the perspectives of both researchers and practitioners. Harkness et al. (2006) illustrated the contributions of the dialectical perspective through the integration of etic and emic perspectives in an international collaborative study (see Box 6.2).

Mertens (2003, 2007) proposed the *transformative–emancipatory* paradigm that incorporates the pragmatist and the dialectical perspectives. Mertens (2007) described the transformative approach as "a cyclical model of mixed methods research . . . as a way of involving the community in research decisions and the collection of data that can be used for social justice purposes" (p. 224). The purpose of research from the transformative–emancipatory perspective is to understand social, cultural, and political realities from perspectives of stakeholders and to create change through researcher–participant partnerships. The pragmatic view is reflected in the intended consequences, to empower participants to bring about social change or action through the participatory research process. The dialectical perspective is evident in efforts to integrate emic and etic perspectives

BOX 6.2

Illustration of the Contributions of Dialectical Perspective: Emic + Etic = Derived Etic

Harkness et al. (2006) used a mixed methods design—specifically, qualitative semistructured interviews and quantitative structured questionnaires—to develop a cross-cultural child temperament measure that was relevant to cultures in seven countries. Using data from parent descriptors of their children (emic) and existing child development theories and measures (etic), the researchers developed instruments that reflected convergence of emic and etic perspectives (derived etic) that could be applied to cross-cultural research. Important to their work was the identification of descriptors that reflected convergence from parents' interview data across the seven countries, thus achieving consensus on what might be described as universal descriptors of child temperament. This method contrasts with the application of an etic perspective (e.g., measures of temperament developed and validated in the United States) in cross-cultural research.

through the participatory process. Thus, the transformative–emancipatory framework focuses on the consequences of research (the pragmatic) and involves research participants as collaborators (the dialectical).

Mertens (2007) described the educational reform work of Thomas (2004) as an example of transformative evaluation research that integrates scientific methods with political activism, "an evolving entity that is developed through a co-constructive process involving the evaluators, school staff members, parents, and students" (p. 223). This perspective is inherent also in participatory approaches focused on social change and capacity building in domestic and global contexts, such as those of Eade (1997), Nastasi, Hitchcock, Sarkar, et al. (2007), and Nelson and Prilleltensky (2005).

The use of MMR to address research questions from pragmatic, dialectical, and transformative perspectives speaks to MMR's potential to address more complex questions through the mix of qualitative and quantitative research. An example of this potential can be seen in the work of researchers who study the implementation of evidence-based interventions (EBIs) in order to answer questions about how to translate EBIs to practical settings. You might recall from Chapter 4 (Dr. Cox's chapter on quantitative research designs) the value of experimental research—for example, randomized controlled trials (RCTs)—for testing the efficacy of an intervention. Such testing occurs under highly controlled conditions that may not match the conditions in real-world contexts. Even RCTs conducted in natural settings to test intervention effectiveness require some level of control to ensure validity. This level of control is not always possible as practitioners attempt to implement EBIs. To address these challenges, implementation scientists have begun to use MMR to better understand the necessary conditions for effective interventions. To do this, researchers use quantitative designs such as RCTs to test effectiveness of the intervention and use qualitative methods such

BOX 6.3

Illustration of MMR to Facilitate Translation of Research to Practice

Goldstein, Kemp, Leff, and Lochman (2013) described and illustrated the use of mixed methods in participatory action research to facilitate the adaptation of EBIs to different contexts and populations. The process involves selecting an EBI, collecting qualitative data through focus group interviews with the target population, making revisions and piloting the revised program using an experimental (quantitative) design, collecting additional qualitative data from program implementers, making further revisions, and testing the intervention using RCTs with the target population. As described by Goldstein and colleagues, the process is an iterative application of a sequential process—qualitative (Qual) → quantitative (Quan) → Qual → Quan—in which data from each step informs the next. You might also consider the use of both qualitative and quantitative during each step of the process, thus Qual + Quan → Qual + Quan, and so on.

as observation and interviews to document the conditions under which the intervention was conducted (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005). Implementation scientists also use MMR to study the process of implementation and adaptation of EBIs (Lee, Altschul, & Mowbray, 2008). For example, researchers might use qualitative research (interviews, observations) to understand the current context and population, make relevant modifications to an EBI to better fit the target culture and context and then test the effectiveness of the modified intervention using quantitative methods. Furthermore, mixed methods can be used to explore the translation of research and policy to practice (Beidas et al., 2013). Box 6.3 illustrates the use of MMR to facilitate translation of research to practice, drawing on the work of Goldstein et al. (2013). Such research holds promise for the application of MMR to address real-world challenges and to facilitate social and political change.

Types of MMR

Mixed methodologists have proposed multiple schemes for categorizing mixed methods designs, ranging from basic to complex (Creswell, 2009; Nastasi et al., 2010; Onwuegbuzie & Collins, 2007; Teddlie & Tashakkori, 2006; Teddlie & Yu, 2007). Nastasi et al. (2010) proposed five typologies to depict the array of the MMR designs. In this chapter, I focus primarily on basic designs, with brief attention to more complex designs.

For the purposes of our discussion, a *research strand* refers to the completion of a three-phase/stage cycle of research: conceptualization, experiential, and inferential (see Teddlie & Tashakkori, 2009). *Conceptualization* refers to establishing theoretical and empirical foundations, identifying research purpose, and formulating research questions. *Experiential* refers to data collection and analysis. *Inferential* refers to interpretation and application/dissemination of findings. In the next section, I expand these three phases to encompass a 10-step research process.

There are three basic design typologies: Type I, based on the number of research strands; Type II, the manner in which data are mixed; and Type III, the stage of the research process in which mixing occurs. *Type I* refers to whether research is monostrand or multistrand (see Tashakkori & Teddlie, 2003; Teddlie & Tashakkori, 2009). *Monostrand* refers to a study involving a single cycle (strand) from conceptualization to inference; this is typical of a single research study in which one generates research questions, collects and analyzes data (in this case, using mixed methods), and disseminates the findings. *Multistrand* refers to use of multiple cycles as one might use in a program of research that includes a series of studies. That is, the researcher engages in several conceptualization–experiential–inferential cycles, which are linked by long-term research goals, typically with the findings of one study influencing the purpose and questions of the next study.

Type II addresses the mixing of qualitative and quantitative methods, specifically *types* of data (qualitative and quantitative); *dominance*, the priority or relative importance of qualitative and quantitative (e.g., are qualitative data considered primary and quantitative secondary or vice versa, or are both data types equal in importance); and *timing*, whether qualitative and quantitative are collected at the same time (simultaneous or concurrent) or sequentially (qualitative before quantitative or vice versa). Morse (2003) developed a notation system to facilitate communication among mixed methods researchers, which included (a) types of methods, Qual (qualitative) or Quan (quantitative); (b) priority or dominance of methods, QUAL (dominant) or qual (supplemental); and (c) timing of methods, qual + quan (simultaneous) or qual → quan (sequential). Thus, an MMR study whose research questions call for dominant focus on qualitative methods, using sequential timing with quantitative data collection following qualitative, is depicted as QUAL → quan.

Decisions about types of data, priority given to different types, and timing are dependent on the study purpose. The complexity of decision making related to Type II typology is reflected in the diversity of schemes for classifying design types, ranging from a continuum of qualitative–mixed methods–quantitative (R. Johnson, Onwuegbuzie, & Turner, 2007) to a scheme of eight types of designs representing all the variations in data type, timing, and priority (Leech & Onweugbuzie, 2009). You might find the scheme proposed by Creswell and Plano Clark (2007) to be the most useful, which includes four design types linked to study purpose: (a) *triangulation*, with equal weight and timing of qualitative and quantitative data, QUAN + QUAL; (b) *embedded*, with priority given to one type of data but simultaneous data collection, QUAL (quan) or QUAN (qual); (c) *explanatory*, in which qualitative data are collected subsequently to explain dominant quantitative findings, QUAN → qual; and (d) *exploratory*, in which the primary data collected first are qualitative, followed by quantitative data collection for confirmation, QUAL → quan. (For information about more complex schemes, see Leech & Onwuegbuzie, 2009; Nastasi et al., 2010.)

Type III typologies refer to the categorization of MMR designs based on the phase of the research strand in which mixing occurs, that is, the specific phases related to conceptualization, experiential, and inferential. Most important to these distinctions is (a) whether qualitative and quantitative methods are part of all phases of the research process, that is, pertaining to research questions, sampling, data collection, data analysis, inference, and so on or (b) whether mixing is stage specific, for example, only applying to sampling, data collection, or inference. In planning a study, you will need to consider the following: Do the research questions require both qualitative and quantitative data collection? Do you need sampling procedures specific to both qualitative (purposive) and quantitative (random and representative sample) data collection? Do you need to include procedures for collection and analysis of both qualitative (such as interviews) and quantitative (such as structured survey instrument) data? Are your inferences based on both qualitative (what do interviews tell you) and quantitative (what are findings from surveys) data as well as the integration of both data sets (do data from interviews confirm or disconfirm survey data)? Researchers have proposed typologies that address mixing across all phases, including R. Johnson and Onwuegbuzie (2004), Ridenour and Newman (2008), and Teddlie and Tashakkori (2009). Other researchers have proposed typologies that are stage specific, for example, sampling (Kemper, Stringfield, & Teddlie, 2003; Onwuegbuzie & Collins, 2007; Teddlie & Yu, 2007), data collection (B. Johnson & Turner, 2003), data analysis (Onwuegbuzie & Teddlie, 2003), and data inference/representation (Erzeberger & Kelle, 2003; Sandelowski, 2003; see Nastasi et al., 2010, for full discussion). In the next section, I return to discussing the integration of qualitative and quantitative data.

Integration of Qualitative and Quantitative Methods in MMR

Mixing or integration of qualitative (Qual) and quantitative (Quan) methods can occur at multiple points in the research process. Nastasi et al. (2010) proposed an inclusive framework to characterize mixing, in which questions regarding the inclusion of Qual and Quan are considered at each stage of the 10-stage research process: (1) existing theory, research, and practice/policy, describing the context for research; (2) worldview (theoretical or conceptual foundations); (3) research purpose and questions; (4) sampling; (5) data collection; (6) data analysis; (7) data inference; (8) inference quality; (9) data representation; and (10) application. Each stage is characterized by specific questions that guide decisions about mixing of qualitative and quantitative methods. This includes which methods to use, which are dominant, and the timing of the methods (Table 6.1). A full discussion of this inclusive framework can be found in Nastasi et al. (2010). Box 6.4 illustrates the mixing of Qual and Quan throughout the research process, drawn from the published doctoral research of Woolley (2009).

Summary

In this chapter, I introduced MMR, addressed key philosophical and design issues, and answered questions about the nature of mixed methods and why researchers choose to mix methods rather than use qualitative or quantitative methods alone. MMR is the integration of qualitative and quantitative methods. However, the field of MMR has expanded into a complex system in which the mixing of qualitative and quantitative research methods influences every stage of the research process. Whereas your initial decision about choice of methods—qualitative, quantitative, or mixed—must always follow

TABLE 6.1
Stage-Specific Questions to Guide Decisions About Qual–Quan Mixing in Mixed Methods Research

Stage	Questions	Application
Context	What context is the basis for your research? Is it existing theory, research, practice, and/or policy?	What are the implications of the context (e.g., theory or existing practice) for your choice of methods?
Worldview	What theory or conceptual framework guides your thinking about research?	How does your worldview influence the extent to which you take an inductive, deductive, or abductive approach to inquiry?
Purpose/ question	What is the purpose of your research? What are your research questions?	What do your purpose and questions imply about selection of Qual, Quan, or MMR methods? Do different questions require different methods (e.g., some Qual, some Quan)? What are the implications of questions for timing, priority, and mixing of Qual and Quan?
Sampling	What types of sampling are appropriate to your research purpose and questions?	How are sampling decisions influenced by whether questions require Qual, Quan, or MMR? Are different sampling methods necessary for collecting Qual and Quan data?
Data collection	What data collection methods are appropriate for answering your research questions?	What qualitative and quantitative data collection methods (e.g., observation, interviews, surveys) are needed to answer your research questions? Do different questions require different methods (e.g., some Qual, some Quan)? What are the implications of questions for timing, priority, and mixing of Qual and Quan? For example, what types of data are dominant? In what order do you need to collect different types of data (e.g., sequential, concurrent)?

continued

TABLE 6.1 *(continued)*

Stage-Specific Questions to Guide Decisions About Qual–Quan Mixing in Mixed Methods Research

Stage	Questions	Application
Data analysis	What are the implications of your theory, purpose, and questions for how you approach data collection in terms of choice of methods, priority, timing, and mixing? What data analysis approach is most appropriate to answer the research questions? How will you approach Qual and Quan data and their mixing? How will you make decisions about the need for more data collection (e.g., to achieve saturation or to make sure research questions are answered)?	Do your theory, purpose, and questions imply an inductive, deductive, or abductive approach to data analysis? How will you analyze Qual versus Quan data sets? Based on your purpose, how will you integrate the findings from Qual and Quan data analysis (e.g., will Qual be used to explain Quan data)? Following integration of Qual–Quan, how will you decide if you need to collect more data?
Data inference	How will you interpret the different types of data as they relate to research purpose and questions? How does the priority of data type (e.g., QUAL, quan) influence your interpretation of the data? How does the timing of data collection influence your interpretation? How will you resolve discrepancies between Qual and Quan? How will you decide if you have answered the research questions based on each type of data and their mixing?	What is the process for interpreting data based on timing (concurrent vs. sequential), priority (dominance of data type), and mixing (complementary, explanatory, etc.)? For example, will the dominant data type have more weight in interpretation? If you engaged in qual → quan sequence of data collection, will you analyze the types of data in the same sequence, and will findings from qual influence how you interpret quan? How will you make decisions when integrating data, based on your conceptual framework and purpose? For example, a transformative paradigm would imply that data are mixed in order to facilitate social or political change. If the purpose of your study is to generate and test theory, would this require that you use QUAL for generating and QUAN for testing, or can you use both Qual and Quan to inform both generation and testing of theory?

continued

TABLE 6.1 *(continued)*

Stage-Specific Questions to Guide Decisions About Qual–Quan Mixing in Mixed Methods Research

Stage	Questions	Application
Inference quality[a]	What techniques are necessary to ensure inference quality of Qual, Quan, and their integration? How will findings related to inference quality influence decisions about additional data collection?	Based on your study purpose and design, what are the inference quality techniques for Quan data? For Qual data? For mixing Quan and Qual? How is that influenced by priority and timing of data types? What if one or both types of data have low inference quality? How can the integration of Qual and Quan enhance inference quality?
Data representation	What is the best way to represent the data types? How will you present Qual, Quan, and their integration for different audiences? How will presentation of data potentially inform subsequent research?	If you operate from a transformative paradigm, what is the best way to represent findings so that stakeholders can use the information to inform social or political change? How would you represent data differently for lay and professional audiences?
Application	What are the potential applications of research findings? What do the context, worldview, and purpose imply about the use of findings?	To what extent can research findings be used for the following purposes: (a) to extend existing knowledge; (b) to facilitate social change; (c) to generate new research questions; (d) to provide direction within a program of research; (e) to validate an assessment instrument or intervention program, which could then be used in practice; or (f) to generate a new conceptual framework? Are these applications justified based on the study design, findings, and inference quality?

Note. An expanded discussion of these questions can be found in Nastasi et al. (2010).
Qual = qualitative; Quan = quantitative; MMR = mixed methods research.
[a]Inference quality is addressed in Chapter 7.

BOX 6.4

Illustration of applying mixed methods throughout the research process

Woolley (2009) illustrated the integration of mixed methods in the research process in carrying out his doctoral dissertation and provided a rationale for decisions relating to MMR. The process of decision making is described here, with notations of relevant steps (see Table 6.1). He wanted to investigate the interaction of personal agency and sociocultural contexts (i.e., gender and institutional setting) and explore both the variations and importance in development of young adults, ages 18 to 25 years, living in Derby, United Kingdom. He generated related research objectives and questions (Steps 1 and 3). His work was informed by a sociological perspective that required adopting MMR to generate a conceptualization of the agency–context interplay through the integration of qualitative and quantitative methods (Step 2). The study's sequential design included initial data collection of quantitative survey with 300 participants (using quota sampling related to objectives), followed by focus group interviews with 47 participants (a subset of the 300; purposive sampling), and subsequently individual interviews with a subset of 8 participants (4 male, 4 female; purposive sampling) (Steps 4 and 5). Data analysis of the quantitative data (factor analysis, analysis of variance, descriptive statistics) informed focus group questions. Data from focus groups were used to elaborate on quantitative findings. Subsequent biographical interviews were used to further explicate the interplay of agency and context, initially established through surveys and focus groups (Step 6). The integration of data from surveys, focus groups, and individual interviews informed the inferences drawn from the study. In the process of integration, Woolley also addressed apparent discrepancies, in part drawing on knowledge of the context from which the sample was drawn and discussed how this process influences data representation for dissemination (Steps 7, 9, and 10). Woolley did not explicitly address inference quality, although you might conclude that the systematic interplay of qualitative and quantitative data as described enhanced quality (Step 8). Woolley described the process as "linking the quantitative and qualitative components effectively [as] the basis for producing integrated findings that are greater than the sum of their parts" (p. 23).

from your study's purpose, this choice will lead to questions at every step of the research process (as depicted in Table 6.1). The most common conceptual framework or worldview/paradigm behind MMR is pragmatism, the focus on what methods best answer the questions and what is appropriate for the context of research and the intended consequences of the research. In addition, mixed methods researchers are likely to operate from dialectical (value of conflicting viewpoints) and/or transformative–emancipatory (research for social change) worldviews.

The growth of the field of MMR has resulted in a range of typologies for categorizing mixed methods—both simple and complex. The basic typologies address questions related to the number of research strands (single vs. multiple), the manner of mixing (type of data, relative priority of qual–quan, and timing of quan–qual data collection), and the stage in the research process in which mixing occurs. Ideally, your choice of MMR would result in mixing of qualitative and quantitative throughout the research process. In practice, you may only choose to mix or integrate in select phases, for example, only in data inference. In such cases, qualitative and quantitative data would be treated separately during data collection and analysis but combined for the purpose of interpretation. Mixed methodologists have proposed more complex schemes for characterizing the array of decisions about mixing of methods, which can occur at every stage of the 10-step research process, as summarized in Nastasi et al. (2010) and portrayed in Table 6.1. The purpose of this chapter was to provide an introduction to what has become a complex field of methodology. What is most promising about the field of MMR is the potential to explore more complex research questions such as those posed in implementation science, an interdisciplinary field focused on enhancing the translation of research to practice and policy.

I concluded this chapter with coverage of the qualitative–MMR–quantitative research continuum. In the next chapter, questions related to the inference quality (e.g., validity, trustworthiness) of research along this continuum are explored. Ultimately, the value of your research depends on the extent to which consumers, such as other researchers, practitioners, and policymakers, can be confident in research findings.

This concludes the three chapters devoted to the approaches to inquiry. The next chapter provides a detailed analysis of reliability and validity in research. Drs. Mary Stewart and John Hitchcock provide robust definitions of reliability and validity appropriate to quantitative and qualitative research. In their analysis, they provide you with some considerations that will help you to design your research studies in the most rigorous way.

KEY SOURCES

Creswell, J. W., & Plano Clark, V. L. (2010). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, *33*(7), 14–26. doi:10.3102/0013189X033007014

Mertens, D. M. (2007).Transformative paradigm: Mixed methods and social justice. *Journal of Mixed Methods Research*, *1*, 212–225.

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

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks, CA: Sage.

REFERENCES

Beidas, R. S., Aarons, G., Barg, F., Evans, A., Hadley, T., Hoagwood, K., . . . Mandell, D. S. (2013). Policy to implementation: Evidence-based practice in community mental health—Study protocol. *Implementation Science*, *8*, 38. doi:10.1186/1748-5908-8-38

Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research*, *6*, 97–113.

Creswell, J. W. (2009). Editorial: Mapping the field of mixed methods research. *Journal of Mixed Methods Research*, *3*, 95–108.

Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.

Creswell, J. W., & Tashakkori, A. (2007). Editorial: Differing perspectives on mixed methods research. *Journal of Mixed Methods Research*, *1*, 303–308. doi:10.1177/1558689807306132

Denscombe, M. (2008). Communities of practice: A research paradigm for the mixed methods approach. *Journal of Mixed Methods Research*, *2*, 270–283.

Eade, D. (1997). *Capacity-building: An approach to people-centred development*. Oxford, United Kingdom: Oxfam.

Erzeberger, C., & Kelle, U. (2003). Marking inferences in mixed methods: The rules of integration. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 457–490). Thousand Oaks, CA: Sage.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication No. 231). Retrieved from http://nirn.fmhi.usf.edu

Goldstein, N. E. S., Kemp, K. A., Leff, S. S., & Lochman, J. E. (2013). Guidelines for adapting manualized interventions to new target populations: A step-wise approach using anger management as a model. *Clinical Psychology: Science and Practice*, *19*(4), 385–401.

Greene, J. C. (2007). *Mixed methods in social inquiry*. San Francisco, CA: Jossey-Bass.

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.

Harkness, S., Moscardino, U., Bermúdez, M. R., Zylicz, P. O., Welles-Nyström, B., Blom, M., . . . Super, C. M. (2006). Mixed methods in international collaborative research: The experiences of the International Study of Parents, Children, and Schools. *Cross-Cultural Research*, *40*(1), 65–82. doi:10.1177/1069397105283179

Hausman, A. J., Baker, C. N., Komaroff, E., Thomas, N., Guerra, T., Hohl, B. C., & Leff, S. S. (2013). Developing measures of community-relevant outcomes for violence prevention programs: A community-based participatory research approach to measurement. *American Journal of Community Psychology*, *52*, 249–262. doi:10.1007/s10464-013-9590-6

Hitchcock, J. H., & Nastasi, B. K. (2011). Mixed methods for construct validation. In P. Vogt & M. Williams (Eds.), *Handbook of methodological innovation* (pp. 249–268). Thousand Oaks, CA: Sage.

Hitchcock, J. H., Nastasi, B. K., Dai, D. C., Newman, J., Jayasena, A., Bernstein-Moore, R., . . . Varjas, K. (2005). Illustrating a mixed-method approach for identifying and validating culturally specific constructs. *Journal of School Psychology*, *43*(3), 259–278.

Hitchcock, J. H., Sarkar, S., Nastasi, B. K., Burkholder, G., Varjas, K., & Jayasena, A. (2006). Validating culture- and gender-specific constructs: A mixed-method approach to advance assessment procedures in cross-cultural settings. *Journal of Applied School Psychology*, *22*, 13–33.

Johnson, B., & Turner, L. A. (2003). Data collection strategies in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 297–320). Thousand Oaks, CA: Sage.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, *33*, 14–26. doi:10.3102/0013189X033007014

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, *1*, 112–133.

Kemper, E. A., Stringfield, S., & Teddlie, C. (2003). Mixed methods sampling strategies in social science research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 273–296). Thousand Oaks, CA: Sage.

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

Lee, S. J., Altschul, I., & Mowbray, C. T. (2008). Using planned adaptation to implement evidence-based programs with new populations. *American Journal of Community Psychology*, *41*, 290–303. doi:10.1007/s10464-008-9160-5

Leech, N. L., & Onwuegbuzie, A. J. (2009). A typology of mixed methods research designs. *Quality and Quantity*, *43*, 265–275.

Maxcy, S. J. (2003). Pragmatic threads in mixed methods research in the social sciences: The search for multiple modes of inquiry and the end of the philosophy of formalism. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavior research* (pp. 51–90). Thousand Oaks, CA: Sage.

Mertens, D. M. (2003). Mixed methods and the politics of human research: The transformative-emancipatory perspective. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 135–166). Thousand Oaks, CA: Sage.

Mertens, D. M. (2007). Transformative paradigm: Mixed methods and social justice. *Journal of Mixed Methods Research*, *1*, 212–225.

Morgan, D. L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research*, *1*, 48–76.

Morse, J. M. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 189–209). Thousand Oaks, CA: Sage.

Nastasi, B. K., Hitchcock, J. H., & Brown, L. M. (2010). An inclusive framework for conceptualizing mixed methods design typologies: Moving toward fully integrated synergistic research models. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (2nd ed., pp. 305–338). Thousand Oaks, CA: Sage.

Nastasi, B. K., Hitchcock, J. H., Burkholder, G., Varjas, K., Sarkar, S., & Jayasena, A. (2007). Assessing adolescents' understanding of and reactions to stress in different cultures: Results of a mixed-methods approach. *School Psychology International*, *28*(2), 163–178.

Nastasi, B. K., Hitchcock, J. H., Sarkar, S., Burkholder, G., Varjas, K., & Jayasena, A. (2007). Mixed methods in intervention research: Theory to adaptation. *Journal of Mixed Methods Research*, *1*, 164–182.

Nelson, G., & Prilleltensky, I. (2005). *Community psychology: In pursuit of liberation and well-being*. New York, NY: Palgrave Macmillan.

Newman, I., Ridenour, C. S., Newman, C., & DeMarco, G. M. P., Jr. (2003). A typology of research purposes and its relationship to mixed methods. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 167–188). Thousand Oaks, CA: Sage.

Onwuegbuzie, A. J., & Collins, K. M. T. (2007). A typology of mixed methods sampling designs in social science research. *The Qualitative Report*, *12*, 281–316. Retrieved from http://www.nova.edu/ssss/QR/QR12-2/onwuegbuzie2.pdf

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. 351–384). Thousand Oaks, CA: Sage.

Ridenour, C. S., & Newman, I. (2008). *Mixed methods research: Exploring the interactive continuum*. Carbondale, IL: Southern Illinois University Press.

Sandelowski, M. (2003). Tables or tableaux? The challenges of writing and reading mixed methods studies. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 321–350). Thousand Oaks, CA: Sage.

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining quantitative and qualitative approaches*. Thousand Oaks, CA: Sage.

Tashakkori, A., & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage.

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

Teddlie, C., & Tashakkori, A. (2006). A general typology of research designs featuring mixed methods. *Research in the Schools*, *13*, 12–28.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks, CA: Sage.

Teddlie, C., & Yu, F. (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research*, *1*, 77–100.

Thomas, V. G. (2004). Building a contextually responsive evaluation framework: Lessons from working with urban school interventions. In V. G. Thomas & F. I. Stevens (Eds.), *Co-constructing a contextually responsive evaluation framework: New directions in evaluation* (pp. 3–24). San Francisco, CA: Wiley.

Woolley, C. M. (2009). Meeting the mixed methods challenge of integration in a sociological study of structure and agency. *Journal of Mixed Methods Research*, *3*(1), 7–25. doi:10.1177/1558689808325774

7

Quality Considerations

Mary S. Stewart, PhD, and John H. Hitchcock, PhD

Introduction

If you want to have an impact in academic, policy, business, and program evaluation settings, you must be able to conduct high-quality research. You must also have the skills to assess the rigor of other published research. A way to think about quality in research is to consider certain indicators to demonstrate that research findings accurately represent the subject, phenomenon, or process being studied. Failure to meet standards of quality may result in research that is misleading or inaccurate. For example, suppose that a study examined the effectiveness of a reading intervention by analyzing the test results of first graders before and after the implementation of the intervention. However, the study was done in a way that did not allow researchers to be reasonably sure that intervention exposure was the best explanation for any observed improvement. The findings of such a study would have limited use for educators and administrators because they would not know if they should use the intervention in question. Hence, it is important to be able to develop and critique studies that yield findings that can clearly inform decision making. This chapter introduces and offers examples of commonly used quality indicators in the context of different approaches to inquiry.

The most common way for researchers and scholars to address quality is to consider whether findings (and the data and inferences that form the basis of findings) are *valid*; in other words, they must reflect the actual phenomenon under study rather than reflecting coincidental relationships, the biases of the researcher, or the limitations of the study design. One aspect of validity is *reliability,* which refers to the consistency of results from a research instrument, strategy, or approach. That is, a reliable research instrument would be one that yields the same findings when administered multiple times on the same subject. In this chapter, we provide a conceptual overview of these two specific quantitative quality considerations (validity and reliability) as well as of their qualitative counterparts (credibility, transferability, and dependability), because these are the primary yardsticks by which research quality is gauged.

There are, however, some challenges to presenting such a broad overview, because even though these are fundamental terms, they have different meanings for different types of researchers, depending on their area of expertise and methodological training. These concepts take on slightly different definitions and are represented by competing terminology across approaches to inquiry. These different definitions can yield some disagreements across various subfields, although such disagreements are not inherently problematic; indeed, scholarly arguments are a necessary ingredient for improving academic disciplines. Instead of focusing on inconsistencies or evolving definitions, the purpose of this chapter is to make you aware of the conceptual bases of *quality* in social science research and some of the broad debates that shape these concepts. In this chapter, we do not go into great detail on method-specific issues related to validity and reliability, or the philosophical orientations aligned with different epistemologies; rather, our hope is to describe how your increased attention to research design, execution, and analysis can yield higher quality findings.

We provide a broad introduction to validity, specifically using experimental design as an example of how validity can be undermined or enhanced as a function of design choice. We also review the topic of trustworthiness, which is often used in qualitative research. We chose to highlight these two approaches to research because they are reasonably concrete and span a wide set of studies. Although these approaches are quite different, we argue that quality considerations apply to many different types of social science research (e.g., case studies, survey work, test design, phenomenology, single-case designs, and developing models that predict given outcomes). We also introduce several basic method-specific terms and techniques for improving research quality. Finally, although validity and reliability are two of the most central quality indicators, they are certainly not the only important indicators. To that point, this chapter cannot be thought of as a one-stop source for what you need to know. Indeed, quality considerations pertaining to the research process are involved in every step of the research process (Guest & MacQueen, 2008), and you should always investigate issues of quality that are specific to your chosen research designs.

Validity and Reliability

Validity

The meaning of *validity* is related to the concept of truth; in research, *valid* findings accurately describe or reflect the phenomenon under study. The concept of truth is also reflected in the qualitative term *trustworthiness*, which some scholars approximate to the quantitative notion of validity. Cook and Campbell (1979), two of the best-known scholars on experimental design methodologies, wrote that validity is “the best available approximation to the truth . . . of propositions” (p. 37).

There are several considerations in the research process that are necessary to promote valid findings, and these all relate to designing a study that is appropriate to the research question. Such considerations include understanding whether (a) the method of data collection (quantitative or qualitative) enables you to answer the specific research question, (b) the type(s) of data collected (interviews, attitude surveys, standardized test results) enable(s) you to answer the question, (c) the sample of data collected enables you to address a target question (i.e., did you question or test the appropriate types of people or other subjects?), (d) you asked the participants questions that were appropriate to the research question, and (e) you included enough participants such that results can be applied beyond the study. These are just a few of the details that you must consider when thinking about the quality of a study.

Although the concept of validity broadly reflects the idea that research findings reflect the true phenomenon, causal mechanism, or attitudes under study, different types of studies and methods necessitate different approaches to ensure validity. Some methodologists, primarily from qualitative and mixed methods traditions, sometimes use different terminology for concepts related to validity, such as *credibility*, *trustworthiness* (e.g., Lincoln & Guba, 1985; see also Guest & MacQueen, 2008; Onwuegbuzie & Johnson, 2006), *legitimation* (e.g., Johnson & Onwuegbuzie, 2004), and *inference quality* (e.g., O'Cathain, 2010). Some qualitative methodologists even reject the concept of qualitative validity altogether (e.g., Wolcott, 1990). Some of these disagreements are rooted in real and honest differences in philosophy, or how one thinks about the world. For example, some researchers espouse a postmodern framework to question the assumption that there can be one reality to portray a finding, or that even primary data, such as informant interviews, test results, or survey responses, are able to fully describe that reality (cf. Lofland, Snow, Anderson, & Lofland, 2009; Onwuegbuzie & Johnson, 2006). However, for the purposes of this chapter, we avoid this disagreement and operate from the assumption that certain aspects of reality can be observed and/or measured by researchers; the slight variations in the concepts of *validity* and *trustworthiness* correspond to their relationship with various approaches to inquiry and methods.

After choosing research questions, you must consider what kind of study design and methods are appropriate to address your questions at hand. Fortunately, there is an existing framework[1] for just about every type of design. Simply put, if you are going to pursue any project that takes several steps and careful thinking, it helps to have a set of guidelines to follow. Methodological guidelines, or frameworks, are available when doing surveys, case studies, psychometric studies (i.e., developing tests and measurement instruments), experiments, ethnographies, phenomenological studies, mixed methods investigations, and so on (see Table 7.1).

A framework is an established structure for the design and execution of a given type of study, including data collection methods, data management, and analytic methods. In addition, all frameworks include components for checking for quality, whether in terms of validity (quantitative) or trustworthiness (qualitative). Frameworks also tend to address validity as both a

[1] By framework, we refer to a set of ideas that can help us think through research processes and findings. Shadish, Cook, and Campbell (2002), for example, describe four facets of experimental validity and ways in which validity can be undermined during the course of an experiment. Some have called this the *Campbellian validity framework*. Lincoln and Guba (1985) offer one of the earlier sets of guidelines for strengthening qualitative inquiry. As another example, O'Cathain (2010) describes a framework for assessing the quality of mixed methods studies.

TABLE 7.1
List of Different Frameworks and Suggested Readings

Type of Inquiry	Description	Suggested Readings
Case studies	Used when trying to learn about phenomena in the context of a particular case (e.g., a person, school, etc.).	Yin (2009)
Experimental design	To be used when a central question is causal in nature, such as when obtaining evidence that a new teaching technique might yield better learning outcomes compared to another technique.	Shadish et al. (2002)
Ethnography	This form of inquiry tends to be used when studying a particular cultural group.	LeCompte & Schensul (2010)
General qualitative inquiry	For studies where the aim is to explore and learn about phenomena in natural settings.	Brantlinger, Jimenez, Klingner, Pugach, and Richardson (2005) Denzin and Lincoln (2005) Lincoln and Guba (1985) Nastasi and Schensul (2005) Patton (2014)
General statistical guidance	There is a lot of guidance around application of the general linear model, which is used in most statistical analyses readers of this text are likely to run into. We offer one text because we find it to be accessible and amusing.	Field (2013)
Meta-analysis	For use when synthesizing the results of multiple, existing studies to learn about aggregated levels of evidence.	Hedges and Olkin (1985) Lipsey and Wilson (2001)
Mixed methods	These apply when combining both qualitative and quantitative design elements.	Creswell and Plano Clark (2010) O'Cathain (2010) Tashakkori and Teddlie (2010)
Psychometrics	To be used when the central purpose of a study is to develop or refine a test/measurement instrument.	American Education Research Association, American Psychological Association, and National Council on Measurement in Education (2014) Crocker and Algina (1986)

continued

TABLE 7.1 *(continued)*
List of Different Frameworks and Suggested Readings

Type of Inquiry	Description	Suggested Readings
Phenomenology	This is a particular variant of qualitative inquiry that focuses on understanding the experiences of research participants.	Moustakas (1994)
Single-case (single-subject) designs	For studies that aim to test intervention effects on small numbers of people (e.g., ABAB and multiple baseline designs).	Horner et al. (2005) Kratochwill et al. (2010) Kratochwill et al. (2013) Kratochwill and Levin (2014)
Surveys	For studies that use surveys, typically when working with a sample and the intent is to learn more about some population of interest.	Dillman, Smyth, and Christian (2009) Fowler (2009) Groves et al. (2009)

Note. This list is not meant to be complete because there is such wide variation in the types of studies and designs within the broad arena of the social sciences. We selected a few on the basis that we think they are commonly used. We also did not attempt to be comprehensive with the citation list. These can be considered as a beginning set of resources to learn more. Later in the chapter, we cover ideas from experimental and qualitative frameworks in more detail. Finally, in some cases, we impose the word *framework*. Some of the citations we offer use this term and others do not, but we are otherwise confident that the authors of the citations would agree that their intent was to offer guidance on how to carry out the particular form of inquiry.

process, where following certain steps should help yield defensible findings, and as an outcome, where one examines the degree to which a set of findings are defensible (e.g., Kane, 2013; Onwuegbuzie & Johnson, 2006). Validity should not be considered as a one-dimensional goal. It is both a process and an outcome and requires an iterative process that continually helps us get a better understanding of whatever is being studied. On that note, notice that we refer to validity as something that you strive for; it is thus best thought of as a kind of continuum. That is, evidence for validity ranges from being poor to really great, as opposed to something being either valid or not. As Cook and Campbell (1979) have stated, "We should always use the modifier 'approximately' when referring to validity, since one can never know what is true. At best, one can know what has not yet been ruled out as false" (p. 37).[2]

Quality of Data Sources and Methods

The quality of data sources and data collection methods has implications for validity. In both qualitative and quantitative studies, there may be inconsistencies among data sources. For example, a subject's actions may not match what the subject says he or she does, or the topic of a given survey

[2] You may have been exposed to this sort of wisdom before. Consider the saying: *Believe those who seek the truth; doubt those who say they have found it.*

question, such as drug use or other illegal activities, may incentivize subjects to answer items inaccurately (Cronbach, 1946; Groves et al., 2009; Lofland et al., 2009). For example, if you wanted to study the prevalence of cheating, how might you gather such information? Would you conduct interviews to ask participants to confess their tendency to cheat? The chances are fairly high that your participants would underreport their behavior. If you chose a different approach, such as allowing them to self-disclose their cheating behavior via an anonymous survey, you might get a more accurate—or valid—portrayal of their behavior. You must therefore examine your data sources and data collection methods for problems that may undermine validity. For this reason, triangulating data sources and using mixed methods are often done to bring to light inconsistencies among qualitative sources (Denzin, 1989; Guest & MacQueen, 2008; LeCompte & Goetz, 1982; Onwuegbuzie & Johnson, 2006), and there are statistical methods for designing and checking the validity of specific survey and assessment questions (e.g., Borgers, Hox, & Sikkel, 2004).

The validity of a research instrument depends in part on its intended purpose and whether it is used for that purpose. In other words, when thinking about data quality, you should consider the evidence for using a specific instrument in a particular situation, rather than thinking of the instrument as valid or not (see Kane, 2013). Consider, for example, standard college or graduate school entrance examinations designed to assess achievement, such as the Scholastic Aptitude Test (SAT), American College Testing (ACT), and the Graduate Record Examinations (GRE). There may be some evidence that these assessments have a valid application in terms of deciding which students are likely to perform well if admitted to given schools (cf. Brewer, Knoeppel, & Lindle, 2014; Hamilton, Stecher, & Klein, 2002; Heubert & Hauser, 1999; Messick, 1994, 1995), but the evidence to support their use for assessing intelligence is far weaker because achievement and intelligence are currently understood to be two different things. Therefore, it is critical to always think of a measurement instrument as a tool, and then consider whether the tool is being used for its intended purpose (Shadish, 1995). Following the tool analogy, you might have a poorly made hammer, such that even after light work it breaks easily. In this case, the tool itself is problematic. But another consideration is the purpose for which the tool is used.

The hammer may be one of the very best ones ever made, but it still would be a poor tool to use when needing a screwdriver. This analogy applies equally well when thinking about what instruments to use when measuring psychological or educational traits. If, for example, we hoped to assess whether some new teaching technique yielded improvements in reading scores, we would not logically choose a mathematics test to use for the outcome measure. But reading is a complex skill with many subcomponents—such as fluency versus comprehension—and we must be able to distinguish the specific skill(s) being tested and which assessments will measure those specific skill(s). Above all, it is important to make sure you are using the right tool for the job.

A related aspect of understanding validity entails thinking about *contextual variables* of the study, including local cultures, time period, and environment (Onwuegbuzie & Johnson, 2006). Every variable, even in quantitative work, represents an attribute that is situated within a specific time and place, and depending on the focus of a given study, certain aspects of variables may be relevant to the defensibility of findings. It is therefore important to always consider the context in which data were collected and interpreted. For example, if you were to do a study on political values, you would need to note the current political climate in which you are doing the study. Or, if you were to read a study on political values, you would need to note the publication date on that study and make sure to take into consideration the political climate of that time period. Consider the two largest political parties in the United States: Democrats and Republicans. In this example, *time* is the specific type of context, because there has been a shift in the overarching political beliefs of these two parties over time. Decades ago, Republicans would have been thought of as being the more liberal of the two political parties. This example demonstrates that researchers, and readers of research, must understand the context of studies.

A related issue is the concept of social construction of variables, or the fact that the common understanding of a variable may be defined by the society in which it is situated, as opposed to being defined by scientific differences. Race is a well-known example of a socially constructed variable, in which racial identity has very little to do with biological differences among races. However, the experiences of people from different races tend to be systematically different within certain societies, based on how those societies consider different races. One example of how race can be socially constructed in multiple ways is the construction of *White* in the United States. Several immigrant groups that are now considered White, including Irish and Eastern Europeans, were once considered racially different from immigrant groups from other Western and Northern European areas (Jacobson, 1998). Racial identities can change as demographic groups assimilate with or differentiate themselves from other groups, and there are often social, economic, and political benefits and/or drawbacks to these changes. As a researcher, you must be aware of the possibility that an attribute has been socially constructed and, if so, whether and how that social construction affects the meaning of the variable. The ways in which variables are constructed will have considerable influence on study validity (cf. Reynolds et al., 2014; Spillane et al., 2010; Wells, Williams, Treweek, Coyle, & Taylor, 2012). Clear, straightforward definitions of each study variable can help to increase validity.

In summary, we reiterate that validity is somewhat synonymous with truth. And just as your definition and understanding of *truth* can be individually subjective as well as based on cultural and social interpretations, so it goes with the concept of validity. When thinking in terms of research quality, consider how to design studies that can yield defensible evidence that can be used to make a reasonably accurate inference or proposition. Being detailed, specific, and thoughtful in all of your design elements and analysis will help you increase validity. These efforts both help you as a researcher to keep your focus within the scope of the study and help the readers of your study to understand your specific research questions, methods, and variables.

Reliability

In a broad sense, reliability refers to the extent to which findings and results are consistent across researchers using the same methods of data collection and analysis. The heart of the concept is synonymous with the notions of consistency and accuracy (Crocker & Algina, 1986), and it is related to validity. You must consider the importance of data and methodological consistency (reliability) because consistency increases the likelihood that your interpretation of data has validity. On the other hand, findings might also be reliably wrong, and this is a critical difference between reliability and validity. To illustrate reliability, consider a scale used to measure a person's weight. If the scale yields a close approximation to a person's actual weight, then one would say the scale's measure is accurate, a key aspect of validity. But now consider the idea of consistency. Suppose a person is weighed once a day for a week, and the scale indicates values of 130, 150, 170, 110, 190, 145, and 155 pounds. Because the weight of one person will not vary so much in a span of a week, the conclusion to be drawn is that the scale is broken. Because of the lack of consistency, or reliability, there is little reason to trust the validity of any single measurement. In this sense, *valid estimation requires consistent, or reliable, scores*.

It is also instructive to see that just because measurements are reliable, they are not necessarily valid. A person may weigh 150 pounds, but suppose that the measurements produced by the scale across 1 week are 191, 189, 190, 191, 190, 190, and 189 pounds. These scores consistently indicate that the person weighs about 190 pounds, but the estimates are consistently wrong. This idea reflects an often-repeated phrase in research methods: *Reliability is a necessary aspect of validity, but insufficient if used alone as a measure of quality*. When using tests and surveys to measure a phenomenon, it is thus critical to understand the properties or the measurement tool and consider whether the measurement tool has been well designed and suited for the job at hand. When engaging in qualitative tasks such as observations and interviews, think of strategies for assessing whether any conclusions to be drawn from these data collection approaches are likely to be consistent (reliable) and accurate (valid).

Reliability Issues in Data Collection and Analysis. Reliability checks can take place at two stages of research: during data collection and during analysis. To test reliability at the data collection stage, another researcher could collect data using the same sampling strategy as you to see if consistent data are being collected. An example of this strategy would be two researchers, with two different scales, each measuring the same person. If both scales indicate the same result, you have some assurance of reliability. It is also possible to have two groups analyze the same set of data with the same analytical methods to see if the two groups come to the same results.

The types of data collected have significant influence over the reliability and replicability of a study (Peräkylä, 1997). Quantitative data sets are often easily accessible and transferable among researchers. Furthermore, quantitative data, once collected and recorded, are not usually subject to any detailed interpretation beyond understanding what a number is supposed to represent. For example, consider a five-option response scale, where 5 = *strongly disagree*, 4 = *disagree*, 3 = *neutral*, 2 = *agree*, and 1 = *strongly agree*. The number 4 is understood to have one meaning (*disagree*), and researchers tend not to conjecture further without having special reason to do so. In contrast, qualitative data are often products of the researcher's filtering and interpreting of information during data collection via observation and interview notes. For example, a researcher creates notes about an observed lesson, and these notes become part of the data set. However, the researcher cannot observe, or record, every detail of the lesson, due to the limited capacity of human observation and as well as choices—conscious or unconscious—that the researcher makes about which details to notice and record. In contrast, data that are not filtered by the researcher at the time of collection include documents, tape-recorded interviews, and videos of observations. The researcher does not have to transfer heard or observed data into a tangible record, such as drawings or notes, because the data are already in a tangible format. Qualitative researchers generally agree that a combination of machine-recorded and interpretive data is ideal in order to achieve a full understanding of the phenomenon under study (see, e.g., Lofland et al., 2009). Research design conceptualization, whether qualitative, quantitative, or mixed method, should entail examining the trade-offs of different kinds of data collection and should incorporate plans for increasing the reliability of the research, such as by using multiple researchers, multiple data sources, detailed data audits, or other strategies.

LeCompte and Preissle (1993) discussed the challenges inherent in trying to replicate the data collection phase of a qualitative study, comparing it with quantitative methods: "Unique situations cannot be reconstructed precisely because even the most exact replication of research methods may fail to produce identical results. Qualitative research occurs in natural settings and often is undertaken to record processes of change, so replication is only approximated, never achieved." Furthermore, they point out that replication may not even be appropriate for qualitative research: "Researchers whose goals are generation, refinement, comparison, and validation of constructs and postulates may not need to replicate situations. Moreover, because human behavior is never static, no study is replicated exactly, regardless of the methods and designs used" (p. 332). One reason for these issues, suggest LeCompte and Preissle, is the longer history of discussions of reliability within quantitative research arenas. These discussions are still fairly new to qualitative methodologists, as demonstrated by the diversity of opinions regarding theory and standardized practices for achieving reliability in qualitative studies.

In summary, you will need to consider whether your data are collected in a consistent manner and whether the type(s) of data collected will help you develop inferences and propositions that approximate the reality of your studied phenomenon. These are not the only important considerations, however. Even when your data are of high quality and are appropriate for the research questions, there are additional considerations to keep in mind when making analytic inferences from these data. That is, you must choose the most appropriate methods to analyze and interpret data in order to reach valid conclusions about your object of study. We use the methodology of experimental design as one example to demonstrate how analytic methods can positively or negatively affect the validity of conclusions.

Validity Considerations in Experimental Design

Shadish and colleagues (2002) provided an overview of an experimental design validity framework that focuses on validity in the context of experimental designs. There are four components to the framework: internal, external, statistical-conclusion, and construct validity. Each type of validity is briefly reviewed in this section, and we focus on internal and external validity in this chapter because these offer relatively concrete examples of how the truth of a proposition can be defended or undermined. Overall, understanding these validity issues is necessary in order to create rigorous study designs as well as to enable critique when reading empirical work done by other researchers. Note that this particular validity framework was developed to help researchers specifically assess *causal mechanisms*; that is, it is used to determine whether a particular condition or treatment causes better outcomes compared to some alternative. This is one example of a framework, as discussed earlier, which provides researchers with standards by which to judge the validity of conclusions. The following discussion of internal validity prompts the use of the experimental design framework, due to the element of causation.[3]

Internal Validity of Experimental Findings

Consider the following statement: *I took some aspirin and my headache went away; therefore, aspirin reduced my pain*. This statement contains a causal inference: Taking aspirin caused the reduction in pain. The degree to which this inference is valid reflects the degree of *internal validity*. In experimental designs, researchers examine whether some variable (the *independent* variable), rather than others, produces some result or change (*dependent* variable; Shadish et al., 2002). Consideration of *internal validity* begs the question: How truthful is the proposition that a change in one variable, rather than changes in other variables, causes a change in outcome?

Causal inference, and thus internal validity, can be surprisingly tricky. For any given proposition about a causal inference, there are rival explanations; these explanations are referred to as *threats* to a statement's validity. For our example, we might assume that you usually take aspirin with water; if the headache had been caused by dehydration, then it is possible that the water—not the aspirin—was the actual cause of pain relief. Alternatively, it is possible that the headache eventually subsided on its own, and thus it was the natural recovery processes—not the aspirin—that yielded the improvement. In short, just because pain subsided after aspiring ingestion does not necessarily mean that we can be sure that the drug was the causal agent. The point here is that rival explanations exist, and it becomes important to consider whether the aspirin explanation is better than the others. The same logic applies to making decisions about policies. For example, you might be interested in implementing a new teaching technique, a new type of counseling procedure, or pay-for-performance compensation models. These all represent policy options, and making the best policy necessitates having data that result in reasonably valid inferences about program or policy impacts by ruling out rival explanations.

[3] A broader notion of internal validity can be conceptualized as the degree to which interpretations of a particular data set are reasonable inferences, causal or otherwise, without getting into the separate question of how well a set of findings applies to new settings outside of the study. But for now, we apply the narrower idea as used in an experimental framework, where one has a research question related to causation.

Once you infer that a given approach resulted in the desired outcome, the quality of this inference can be judged by assessing various threats to internal validity. Each threat is a form of alternative explanations—other than the treatment—for the cause of an observed outcome. The experimental validity framework identifies a number of common threats to internal validity; the following discussion draws on examples of these threats in order to illustrate the process of identifying and eliminating rival explanations (see Shadish et al., 2002, for a complete list and description of threats; see also Table 7.2). As an example, there is the so-called *history* threat, or the possibility that other events may have occurred during the duration of the experiment that could explain the change in outcome. In the case of aspirin, the fact that common headaches eventually subside on their own is an example of such a threat to the inference that it was aspirin that caused pain reduction. This threat, and many of the other internal validity threats, can be addressed by including a comparison (or *control*) group that does not receive the treatment being studied. If a treatment effect is observed by comparing performance across both groups (i.e., students who received counseling show better outcomes than those who did not), then it becomes the case that the independent variable—that is, *treatment exposure*—is the best overall explanation for the difference in scores between study groups. An important point here is that this and many other threats to internal validity can be addressed by adding a control group when the intent is to make a causal inference. Indeed, a basic quality indicator for studies that set out to address a causal question is to look for the presence of a control condition (Shadish et al., 2002).

By adding a control group, we also potentially introduce new threats to internal validity. One such threat is *selection*, which refers to how groups in a study were formed. There are many ways to form groups. People can volunteer to be treated, students might be picked by a teacher, a researcher may decide who is most in need of treatment, and so on. Of the many options, one approach to selecting who is treated and who is in a control group is to use random assignment to treatment and control groups. Such assignment is essentially based on chance: If you use random procedures (such as coin flips or computer algorithms, which are user-friendly in most statistical packages), you can expect that, on average, there will be no systematic differences between groups.

When assignment is nonrandom, such as when participants in groups are purposefully selected based on certain characteristics, there may be key differences between treatment and control groups. For example, due to legal restrictions about research, students in a treatment group may have to be volunteered by their parents/guardians; the requirement that participants actively volunteer and have parent permission introduces the possibility that the characteristics of students in the treatment group are systematically different than those in the control group. Thus, the manner by which selection was done can threaten later attempts to make causal inferences. Suppose the question at hand was whether a new counseling technique yielded better outcomes than typical treatment procedures. Furthermore, suppose that the counseled group did in fact show better outcomes than their control counterparts. One might infer that these improved outcomes were caused by the treatment, but the threat to this inference is the fact that treated students

TABLE 7.2
Overview of Internal Validity Threats

Threat	Definition	Example
History	Other events may have occurred during the duration of the study that could explain the improved behavior.	During the course of treatment, the children in the counseling treatment group were also assigned to a new teacher who is excellent at managing behavior concerns. In this scenario, was the improvement because of the treatment, the presence of a new teacher, or both?
Maturation	The fact that people, including study participants, change over time.	During a 1-year study, the treatment students could have simply outgrown their initial behavior problems; their personal development, unrelated to the counseling, may have contributed or even been solely responsible for the improved behavior scores.
Testing	The possibility that repeated exposure to a measurement instrument could, by itself, affect test-taking behavior and test scores.	The children who took the baseline measurement test reflected on what the test was measuring and at posttest offered socially desirable responses that resulted in higher scores; yet, their overall classroom behavior may not have actually improved. Instead, observed score change was the result of knowing from baseline the test questions and how to best respond.
Instrumentation	(1) A testing instrument may change or may be used in a way that does not correctly measure the treatment effect.	This might happen if there are two versions of a test (Form A and Form B) that are incorrectly assumed to be equivalent. In this case, it could be that behavior as measured by Form A looks more problematic as compared to Form B. If Form B was used at posttest (the second testing session), then any apparent improvement cannot be attributed to the treatment; rather, differences in the test drove the change.
	(2) There may be unknown contextual factors that can impact testing.	Perhaps baseline measurement was done in the morning and posttest measurement was done in the afternoon, and for some reason, the children in the study are more likely to report or demonstrate better behavior after lunch.

continued

TABLE 7.2 *(continued)*
Overview of Internal Validity Threats

Threat	Definition	Example
Statistical regression to the mean	The phenomenon that extreme scores tend to not be repeated.	Anyone might score unusually high on a psychological measure because of a string of positive but rare events, such as winning the lottery. Taking the test 6 months later may still result in a high score, but not extremely high. Over time, extreme scores—both positive and negative—tend to move closer to the average for that particular measure.
Researcher bias	Changes in research design or analysis that are a result of the researcher's subjective views regarding the study topic, participants, theory of change, or other relevant areas of design or execution.	The researchers may be so convinced that the new counseling approach works that they unintentionally modify aspects of the original study design in order to show that the treatment makes a difference.
Selection	The process of creating participant groups in a study. Nonrandom selection (e.g., participants in groups are purposefully selected based on certain characteristics) might yield two groups that are not equivalent at the beginning of a study. If there are key differences between two groups, one cannot know if any posttest differences are because of a treatment effect or because of such differences.	Due to legal restrictions about research, students in a treatment group may have to be volunteered by their parents/guardians; the requirement that participants actively volunteer and have parent permission introduces the possibility that the characteristics of students in the treatment group are systematically different than those in the control group. For example, parents who push for their children to be exposed to a new treatment might, on average, be more involved in their children's education than ones who do not. In such an example, if we see that children who were treated appear to perform better on a posttest, is such improvement because treated children have more involved parents (i.e., they would have been better off anyway), or is it because the treatment worked? The selection threat is not a concern if study participants are assigned randomly to study groups because, on average, there should be no differences between participants in treatment and control conditions.

continued

TABLE 7.2 *(continued)*
Overview of Internal Validity Threats

Threat	Definition	Example
Overall mortality (attrition)	Loss of members in the study sample.	A study compares pre- and posttest scores on an assessment in order to measure participant change. Some students drop out before completing the posttest. The loss of part of the sample creates the possibility that the students who remained in the study and completed the posttest have systematically different characteristics than the students who left the study. If this is the case, any average positive or negative change from pretest to posttest may simply be a function of the mortality threat, or the loss of students with particular characteristics.
Differential mortality	Members of sample groups (e.g., treatment and control) drop out in different rates, and nonrandomly, in one group as compared with the other(s).	Some students in the treatment sample drop out because they no longer wish to receive the counseling and miss out on other activities during the school day. These students may be differently motivated or have systematically different behavioral characteristics than the students who are willing or happy to miss other school activities.

came from families who desired and supported the treatment and students in the control group tended to come from families who were indifferent to the treatment. It is therefore possible that students in the treatment group might have done better than those in the control group even if the treatment was not the cause of the students' improvements. In other words, the way groups were formed may have made it look like the new counseling technique made a difference, even if it did not. Thus, selection processes can threaten the validity of any inference made about treatment. Researchers can prevent selection bias by using random assignment to treatment and control groups, when possible. When random assignment is not an option, researchers may statistically "control" for other variables such as socioeconomic status, gender, race, age, and disability status in statistical models. These types of controls can help tease out the differential impact of these contextual factors.

The broader point here is that, in general, not all designs are equal in terms of their inherent capacity to address the internal validity of target research questions. If you are conducting an experiment, it will behoove you to design studies to have stronger validity because doing so will result in better quality. Without certain design features put into place, the improvement in behavior may indeed be due to the treatment but may also be due to any of the threats listed. Fortunately, in the context of an experiment, as with many other methodologies, there is guidance that you can consult to assist with recognizing and addressing these threats. Unless these threats can be removed as plausible explanations, the study quality must be considered questionable (see Table 7.2).

External Validity of Experimental Findings

External validity—the extent to which findings hold true across contexts—and its threats are also major considerations in research design quality. Suppose you have produced a study with high internal validity; that is, none of the previously discussed threats are plausible explanations for observed improvements in a treated group. The best explanation for the outcome of the study is that the treatment worked. This high level of internal validity leads other researchers to want to know whether this finding has high external validity as well, or whether it is likely to hold true across other students, in other places, times, contexts, cultures, and so on. As with internal validity, there are several common threats to external validity (Shadish et al., 2002) (see Table 7.3).

To elaborate a little, one such threat is *treatment variation*; this type of threat addresses the degree to which observed treatment effects reflect variations in the treatment received by the study subject(s). Treatment variation can be a function of the human error of administering a treatment or a function of seemingly innocuous choices around implementing a program. One example would be inconsistency in dosage levels; consider two teachers ostensibly delivering the same intervention, but one teacher has excellent classroom management skills and the second teacher does not. The second teacher's students receive less of a dosage of the intervention because one-third of class time is spent on classroom management issues. Other examples include the time of day that treatment is delivered or failure to correctly implement some element of the treatment.

Threats to external validity present a number of concerns, and researchers must find ways to address these threats. There are two broad strategies for addressing threats to external validity. The first is to engage in thorough literature reviews and to build on previous, related studies. External validity can be strengthened by limiting the research focus and by comparing new findings to existing studies in the literature. A careful review can highlight gaps in the existing literature; these gaps then justify a specific focus that is situated within an existing framework of studies. For example, a specific counseling technique may have been thoroughly studied in residential treatment settings, and so your focus might be on the first effort to try it in a public school. The design of your study will be strengthened by the evidence available from other related studies, and the threats to external validity will be minimized by limiting the focus to a very specific area.

The second strategy is to think carefully about ways in which your findings may apply, or generalize, to other settings. Shadish (1995) listed a number of principles that can help you think about generalization when doing experiments, ethnographies, or other types of studies. You must consider how applicable the findings from your study might be to another setting, such as similarities in the sample and how it was obtained, measurements

TABLE 7.3
Overview of External Validity Threats

Threat	Definition	Example
Interactions of the observed causal relationship with sample units	The possibility that whatever was observed with one particular sample may not hold true for different samples.	Simply put, the treatment may work well with one type of student and not another.
Treatment variations	The effect of a treatment reflects variations in how it was administered, and so on, as opposed to the effect of the treatment itself.	Treatment variation can be a function of the human error of administering a treatment or a function of seemingly innocuous choices around implementing a program (e.g., dosage levels, time of day treatment is delivered, or failure to correctly implement some element of the treatment).
Types of outcome measures used	Treatment effects may be found with one kind of test but not another.	One might see an effect with a particular type of test but not another. If two tests measure approximately the same thing (e.g., SAT and ACT), this should be less of a concern, although when differences are found across similar but different tests, one has to wonder about the external validity of any observations from a study. Logically, concerns arise when thinking about the degree to which study findings might be externally valid when thinking about clearly different outcome measures.
Settings in which the treatment was delivered	The possibility that observed effects are due to contextual factors, as opposed to the treatment itself.	A simple example of this threat would be observing treatment effects in a school that is located in a high-income community; the same effects may or may not hold in more impoverished settings.
Context-dependent mediation	The influence of a mediating factor in one setting versus another setting.	A common mediating factor is treatment dosage; others may be factors such as staff skill or availability. For example, is it possible to fully implement an intended treatment in the form of intense counseling in an overcrowded school setting where there are extensive demands on a counselor's time?

Note. SAT = Scholastic Aptitude Test; ACT = American College Testing.

used, duration, and other treatment details. Above all, claims of generalizability are most appropriate when there is evidence that a very specific aspect of a treatment yields an exact outcome. Knowing what aspects of a study are likely to generalize and what aspects are likely to be highly context specific is the key to thinking through considerations that might threaten the generalizability of a finding to some new scenario.

To illustrate these issues around generalizability, we use an example from a study by Paino, Renzulli, Boylan, and Bradley (2014) that did not examine the effect of a treatment, but rather charter school closings in North Carolina (this is to show that generalization can and should be pondered not only when dealing with treatment effects but also when dealing with other issues, such as state policy). The authors performed a quantitative analysis of data on charter schools and the nearby public districts. Data included financial information, local market variables, density of charters in the area, school demographics, enrollment, age of school, and academic performance information. This quantitative analysis allowed the researchers to examine the probability of a charter school closing at a given point in time. The findings suggested that charter schools were less likely to close with increases in school enrollment, compliance with federal desegregation orders, and state and federal funding of charters. However, because the location of this study is in one state, its findings may not generalize well to another state that may have different policies. Here, the authors' inclusion of a qualitative case study analysis could help them to better understand the degree to which these findings might generalize to other states and contexts. Suppose a state has conditions similar to that of North Carolina—contextual conditions that have been rigorously analyzed in relation to the quantitative findings. As a reviewer of the study, you may feel more confident in applying the study findings to that new context. On the other hand, if the case studies in North Carolina show major differences in state charter policies, funding, or enrollment patterns, you may not feel confident in using the findings of this study to understand patterns in the other state.

Statistical-Conclusion and Construct Validity

There are two remaining types of validity from Cook and Campbell's (1979) framework. *Statistical-conclusion validity* refers to the degree to which researchers are correct about the relationship between two variables. This type of validity requires not only that researchers know which kind of statistical models or techniques are appropriate for a given data set and research question but also that they can accurately test those models and apply those techniques. Shadish and colleagues (2002) identified nine distinct threats that are helpful; if you are doing quantitative research, we highly encourage you to review this resource in depth. Other concepts and techniques that relate to statistical-conclusion validity include statistical power, data cleaning, and outlier analyses. Measurement reliability, or lack thereof, is classified as a threat to this form of validity.

Construct validity refers to the degree to which underlying ideas (e.g., treatments, behaviors, behavior problems, cooperative learning, and socioeconomic status) are properly conceptualized and operationalized in a study. Every study is based on a set of concepts that underlie the theory being tested. In our ongoing example, the theory being tested in the experiment is that a certain type of counseling intervention will improve problematic behavior issues. If the measurement of this improvement is completed through a student pre- and postintervention assessment, we must ensure that (a) the intervention addresses the behaviors under study and (b) the questions on the assessment correctly represent the behaviors under study. An intervention or measurement that does not accurately represent the constructs being studied cannot result in valid findings about the constructs.

Considerations in Qualitative Inquiry

Earlier, we presented aspects of the experimental validity framework to demonstrate the point that your design choices can affect the validity of inferences you make at the end of your study. We also demonstrated this point because causal questions tend to be of wide interest. Moving forward, we focus on another broad arena: qualitative research. Some of the challenges, or threats, to reaching validity and reliability in quantitative and qualitative research are similar, although they must be observed or measured using different techniques (e.g., Onwuegbuzie & Leech, 2007). For example, whereas quantitative researchers attempt to statistically control for variables that may influence the outcome, qualitative researchers attempt to understand the influence of variables through careful observation and recording of phenomena (Cook & Campbell, 1979; LeCompte & Goetz, 1982). In the next section, we provide an introduction to validity and reliability issues regarding qualitative research methods.

Trustworthiness

Trustworthiness is the qualitative term that is often used in place of the quantitative term *validity*. Trustworthiness is the degree to which you, as a researcher, can have confidence in your sources as well as the methods used to gather your sources. Steps taken in the earliest stages of research—study purpose and design—can help you decide which collection methods will result in the most relevant, trustworthy data for your questions under study. Ethnographic field notes, formal and informal interviews, formal and informal observations, video recordings, photographs, and archival records offer different strengths and weaknesses (LeCompte & Goetz, 1982). For example, Peräkylä (1997) discussed the specific benefits and drawbacks to tape-recorded and transcribed (audio and/or visual) data, as compared to ethnographic field notes. Field notes filter observations at the time of data collection through the researcher's particular frameworks; in contrast, audio/visual recordings capture all of the data from one particular angle and/or sense (e.g., visual vs. audio). Downsides to audio/visual recordings are the inabilities to see gestures and movements or to see the observation from multiple angles or perspectives, respectively. Ethnographers can take in an entire observation site through all of the senses, but they are limited in what they can record in words or pictures. Using a combination of these data collection methods would allow you to compare two or more data sources; such comparisons can highlight areas of inconsistency that need further inquiry or patterns/themes that have a high degree of consistency (i.e., they surface in multiple types of sources and in ways that do not conflict).

There are a variety of ways in which you as a qualitative researcher can check the trustworthiness of emerging themes in your data (Tracy, 2010). During data collection and analysis, researchers can attend to potential *observer effects*, employ *multiple researchers*, and use *member checks*. Also see Lincoln and Guba (1985) and Nastasi and Schensul (2005) for more in-depth discussion on trustworthiness.

Observer Effects. As a qualitative researcher, you must address the potential influence of *observer effects*, which is the possibility that collected data have been contaminated, or influenced, by your presence or your research instruments. One example of observer effects is a change in participant behavior during observations due to your presence (LeCompte & Goetz, 1982). Depending on the type of activity and individuals under observation, your demographic characteristics, and the methods by which you are recording data, participants may consciously or unconsciously change their behavior. If participants change their behavior, then you cannot report that their observations are typical of participants' natural or normal behavior.

We can use the example of a counseling intervention to illustrate this issue. Imagine this scenario: Suppose there is a qualitative observation element to the study, in which you might observe in a group counseling session for student participants. The majority of the students in the session speak English as a second language, and about half of them have parents who are not U.S. citizens. The majority of students in the group are also on free or reduced-price lunch. In comparison, you, the researcher, are White, well dressed, and speak only English. The demographic differences between you and the student participants include social class, first language, age, and, in some cases, citizenship. These differences may lead students to behave differently in front of you than they would behave with only the counselor present; additionally, they may behave differently in a group session with their peers than in a one-on-one session with the counselor. You can take two precautions against observer effects. First, you can note all of the potential effects that your presence may have on the participants or their behavior; getting a second opinion on these potential effects can dually strengthen this precautionary strategy. Second, you can follow up with members of the group—in this case, the counselor or one of the participants—to ask whether the observed session was typical or uncommon in any way. This type of context from a normal member of a group can help you put your observations in perspective.

Multiple Researchers. Although not always feasible in qualitative studies, using multiple researchers in data collection has benefits as well as challenges for validity. When multiple researchers collect data, they are able to demonstrate that they are recording data in comparable ways; this is vital to study validity. Similar to interrater reliability (see later discussion), multiresearcher data collection procedures must be uniform in order to collect valid and trustworthy data across an entire study. One example of aligning data collection procedures relates to level of detail; all researchers should know how much detail to include in field notes or observation rubrics. This is true for all methodologies; just as tests producing quantitative data must be administered and recorded consistently, interview and observation data must be recorded using the same techniques.

Member Checks. Similar to one of the strategies involved in minimizing observer effects at the data collection level, member checking involves sharing emergent patterns and findings with members of subject groups to get feedback on the accuracy of those findings. Although the purpose of independent research is to create and implement an unbiased research design that examines the input of all relevant stakeholders or participants, there are also limitations to using outside researchers. Outside researchers rarely have the insider cultural perspective or organizational knowledge that is needed to fully understand the phenomena being observed. Member checking allows the outside researcher to share his or her ideas with the views of an insider and develop an ongoing, increasingly accurate understanding of the phenomena (LeCompte & Goetz, 1982; Lofland et al., 2009). The dual use of insider and outsider perspectives is crucial to achieving this accuracy because both perspectives tend to have particular types of biases, such as ingrained cultural or social beliefs (Bloor, 1978; Turner & Coen, 2008). Such beliefs can include views on gender/sex, racial or ethnic groups, or age-appropriate behaviors; for example, a study participant who is a member of a diverse urban community may have different views on race and ethnicity than a White researcher working in a predominantly White, elite institution.

Reliability in Qualitative Research

The concept of reliability, sometimes called *dependability*, is relevant in some ways for qualitative methods and problematic in others. Specifically, the definition of reliability as "replicability" is problematic for qualitative, especially naturalistic, methodologies. As LeCompte and Goetz (1982) explained, "Because human behavior is never static, no study is replicated exactly, regardless of the methods and designs used" (p. 332). However, there are ways in which the larger concept of reliability has been adapted to apply to qualitative, naturalistic fields of study. Areas of focus within the umbrella of qualitative reliability include the replicability of data collection and analysis (e.g., understanding how much of the analysis is specific to an individual researcher's interpretations) and intercoder reliability or interrater agreement, which refers to the degree to which multiple researchers within the same study agree on how to describe and categorize the observed data in terms of the study's theoretical framework. These issues of reliability can be found in many qualitative studies, and the researcher subjectivity plays an important, if sometimes overlooked, role in these processes. The following sections examine challenges to and strategies for strengthening qualitative reliability.

Researcher Subjectivity. Researcher subjectivity refers to the unique perspective that each researcher brings to a given study; this uniqueness poses a reliability challenge for qualitative studies at the stages of both data analysis and data collection because the interpretations of two or more unique researchers are unlikely to be identical, or replicable (Carey & Gelaude, 2008). For example, in an empirical study of qualitative thematic coding, Armstrong, Gosling, Weinman, and Marteau (1997) found that a sample of trained, experienced experts in qualitative coding, looking at the same data set, did not reach the exact same conclusions about the data. The study demonstrated that when multiple researchers analyzed the same data, the themes that emerged were similar enough to be considered common, but different enough to highlight the role of researcher discipline, training, and cultural background. The findings of this study suggested that the inherent nature of subjective analysis in qualitative methods will result in some degree of agreement and some degree of disagreement. See Glaser and Strauss's (1967) description of the constant comparison method for a specific example of how to systematically code qualitative data.

Reflexivity. The findings of this study also point to the need for individual researchers to be *reflexive*, or transparent and forthcoming about their demographics, their discipline, their training, and any other characteristics that may influence their collection or analysis of data. Toward this end, you should reflect on your position in relation to the study and examine the potential for bias based on your cultural or socioeconomic background, nationality, ability status, and other factors (LeCompte & Preissle, 1993; Onwuegbuzie & Johnson, 2006). Your explanation of methodology should also include steps that you take to minimize the impact of your researcher bias on research design, data collection, and analysis (Guest & MacQueen, 2008).

Interrater Reliability. When multiple researchers are used to analyze qualitative data, reliability issues multiply as well. In addition to being reflexive about individual characteristics, the research team must also take steps to ensure that they are using the same criteria to collect or interpret the same data set. *Interrater reliability* refers to the rate of agreement among multiple research team members applying the same analytic methods to the same data set; these methods typically involve some degree of researcher subjectivity, such as coding text or rating observed behaviors. Additional benefits to determining interrater reliability are twofold: The process allows the research team to examine both the team's understanding of codes and concepts as well as individual team member accuracy in using the coding or rating system (Carey & Gelaude, 2008). Like many phases of research, interrater agreement is an iterative process. If the independently coded data samples end up with substantially different results, the coding system must be reviewed and clarified or individual coders must be trained further. Interrater reliability testing must continue until the desired level of agreement among researchers has been achieved (MacQueen, McLellan-Lemal, Bartholow, & Milstein, 2008).

Transferability. A final point, related to reliability in qualitative research, is to consider the concept of *transferability*. Transferability is the degree to which a set of findings from one study will transfer to another particular situation (Lincoln & Guba, 1985). The idea is largely associated with qualitative inquiry, but the principle can be applied to almost any kind of study. The general challenge in transferability is describing the setting of a study with sufficient clarity and detail so that readers of that study can make their own judgments about what does and does not apply to their particular scenarios.

Conclusion

Research quality is important in all disciplines and fields, including program development and implementation, because all knowledge—understanding human behavior, program designs, and effects of medical treatments—is influenced by the quality of the research on which it is based. If inaccurate research findings are used as the basis for products, program development, or policy improvements, these changes are unlikely to actually work as hoped, potentially wasting time and other valuable resources. Some areas of product or program development have a variety of parties with established financial or political stakes in the direction of development; here, it is especially important that cited research be independent and of high quality. Peer review is generally understood to be a hallmark in the research process, because it entails review by multiple experts in the field; the experts are looking for indicators of research quality that provide confidence in the findings. Even with basic knowledge of indicators of research quality, it is possible for a layperson to review the methodology of a given piece of research and decide for oneself whether the piece contains the necessary quality indicators.

It is also important to note that there are particular aspects of mixed methods research that lend it to increasing validity, such as the ability to take advantage of the strongest tools of each framework and discard the weaker tools (Onwuegbuzie & Johnson, 2006). One of the challenges of using mixed methods is figuring out which tools are strongest for which research questions, and Onwuegbuzie and Johnson (2006) discussed several sets of existing guidelines for making mixed methods research decisions (e.g., Collins, Onwuegbuzie, & Sutton, 2006; Greene, Caracelli, & Graham, 1989; Onwuegbuzie & Johnson, 2004). Some qualitative and quantitative methodologists, without purposefully using a mixed methods framework, have incorporated these tools organically in order to best answer their research questions (e.g., Reynolds et al., 2014; Wells et al., 2012).

In this chapter, we have offered an introduction to the range of quality issues that can arise in research studies. This introduction should help you understand that threats to validity and reliability can surface at any point of the research project: design, data collection, data analysis, or even results reporting. To handle validity and reliability concerns, you first need to be aware of them. At every step, you should be looking out for possible threats to research quality, making sure that their design minimizes these threats as much as possible, and clearly reporting the severity of existing threats. To facilitate this process, you should first have a clear understanding of your research question(s). Then, you should seek out methodological frameworks or guidance that promote thinking through designs and generating the highest quality inferences. Finally, you should identify design choices that have the capacity to answer the question well. Again, Table 7.1 is designed with that purpose in mind.

Given the space allocated for this chapter, our overriding advice for you is to appreciate the idea that design decisions can influence the quality of the data collected, later analyses, and overall inferences drawn from your work.

We recommend that you investigate further the wide range of specific strategies and techniques to address the threats to validity and reliability that were briefly introduced here. The following chapter, building on these quality concerns, examines ethical considerations in research projects. Many of the same rationales for research quality support the concern for ethics in research, such as the increasing focus on using research findings to make policy and program decisions.

KEY SOURCES

Brewer, C., Knoeppel, R. C., & Lindle, J. C. (2014). Consequential validity of accountability policy: Public understanding of assessments. *Educational Policy 29*, 1–35. doi: 10.1177/0895904813518099

LeCompte, M. C., & Preissle, J. (1993). *Ethnography and qualitative design in educational research* (2nd ed.). San Diego, CA: Academic Press.

O'Cathain, A. (2010). Assessing the quality of mixed methods research: Toward a comprehensive framework. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (2nd ed., pp. 305–338). Thousand Oaks, CA: Sage.

Shadish, W. R., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Tracy, S. J. (2010). Qualitative quality: Eight "big-tent" criteria for excellent qualitative research. *Qualitative Inquiry*, *16*(10), 837–851.

REFERENCES

American Education Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Armstrong, D., Gosling, A., Weinman, J., & Marteau, T. (1997). The place of inter-rater reliability in qualitative research: An empirical study. *Sociology, 31*(3), 597–606.

Bloor, M. (1978). On the analysis of observational data: A discussion of the worth and uses of inductive techniques and respondent validation. *Sociology, 12*(3), 545–552.

Borgers, N., Hox, J., & Sikkel, D. (2004). Response effects in surveys on children and adolescents: The effect of number of response options, negative wording, and neutral mid-point. *Quality & Quantity, 38*(1), 17–33.

Brantlinger, E., Jimenez, R., Klingner, J., Pugach, M., & Richardson, V. (2005). Qualitative studies in special education. *Exceptional Children, 71,* 195–207. DOI: 10.1177/001440290507100205

Brewer, C., Knoeppel, R. C., & Lindle, J. (2014). Consequential validity of accountability policy: Public understanding of assessments. *Educational Policy*, 1–35. DOI: 10.1177/0895904813518099

Carey, J. W., & Gelaude, D. (2008). Systematic methods for collecting and analyzing multidisciplinary team-based qualitative data. In G. Guest & K. M. MacQueen (Eds.), *Handbook for team-based qualitative research* (pp. 227–272). Altamira, CA: Lanham.

Collins, K. M. T., Onwuegbuzie, A. J., & Sutton, I. L. (2006). A model incorporating the rationale and purpose for conducting mixed methods research in special education and beyond. *Learning Disabilities: A Contemporary Journal, 4*, 67–100.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Boston, MA: Houghton Mifflin.

Creswell, J. W., & Plano Clark, V. L. (2010). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.

Crocker, L. M., & Algina, J. (1986). *Introduction to classic and modern test theory*. New York, NY: Holt, Rinehart, and Winston.

Cronbach, L. J. (1946). Response sets and test validity. *Educational and Psychological Measurement*, *6*(4), 475–494.

Denzin, N. K. (1989). *The research act: A theoretical introduction to sociological methods*. Englewood Cliffs, NJ: Prentice Hall.

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2005). The discipline and practice of qualitative research. In *The Sage handbook of qualitative research* (3rd ed., pp. 1–32). Thousand Oaks, CA: Sage.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail and mixed-mode surveys: The tailored design method* (3rd ed.). Hoboken, NJ: Wiley.

Field, A. (2013). *Discovering statistics using IBM SPSS Statistics* (4th ed.). Thousand Oaks, CA: Sage.

Fowler, F. J. (2009). *Survey research methods*. Thousand Oaks, CA: Sage.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New Brunswick, NJ: Transaction.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation & Policy Analysis*, *11*, 255–274.

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (2nd ed.). Hoboken, NJ: Wiley.

Guest, G., & MacQueen, K. M. (Eds.). (2008). Reevaluating guidelines in qualitative research. In *Handbook for team-based qualitative research* (pp. 205–226). Altamira, CA: Lanham.

Hamilton, L. S., Stecher, B. M., & Klein, S. P. (2002). *Making sense of test-based accountability in education*. Washington, DC: Rand.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York, NY: Academic Press.

Heubert, J. P., & Hauser, R. M. (1999). *High stakes: Testing for tracking, promotion and graduation*. Washington, DC: National Academy Press.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, *71*, 165–179.

Jacobson, M. F. (1998). *Whiteness of a different color: European immigrants and the alchemy of race*. Cambridge, MA: Harvard University Press.

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

Kane, M. (2013). The argument-based approach to validation. *School Psychology Review*, *42*(4), 448–457.

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. M. (2010). *Single-case design technical documentation*. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. M. (2013). Single-case intervention research design standards. *Remedial and Special Education*, *34*, 26–38. doi:10.1177/0741932512452794

Kratochwill, T. R., & Levin, J. R. (Eds.). (2014). *Single-case intervention research: Methodological and statistical advances*. Washington, DC: American Psychological Association.

LeCompte, M. D., & Goetz, J. P. (1982). Problems of reliability and validity in ethnographic research. *Review of Educational Research*, *52*(1), 31–60.

LeCompte, M. D., & Preissle, J. (1993). *Ethnography and qualitative design in educational research* (2nd ed.). San Diego, CA: Academic Press.

LeCompte, M. D., & Schensul, J. J. (2010). *Designing and conducting ethnographic research: An introduction* (2nd ed.). Plymouth, United Kingdom: AltaMira Press.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Lofland, J., Snow, D. A., Anderson, L., & Lofland, L. H. (2009). *Analyzing social settings: A guide to qualitative observation and analysis* (4th ed.). Belmont, CA: Wadsworth.

MacQueen, K. M., McLellan-Lemal, E., Bartholow, K., & Milstein, B. (2008). Team-based codebook development: Structure, process, and agreement. In G. Guest & K. M. MacQueen (Eds.), *Handbook for team-based qualitative research* (pp. 119–135). Altamira, CA: Lanham.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, *23*, 13–23. doi:10.2307/1176219

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*, 741–749. doi:10.1037/0003-066X.50.9.741

Moustakas, C. (1994). *Phenomenological research methods*. Thousand Oaks, CA: Sage.

Nastasi, B. K., & Schensul, S. L. (2005). Contributions of qualitative research to the validity of intervention research. *Journal of School Psychology*, *42*, 177–195. doi:10.1016/j.jsp.2005.04.003

O'Cathain, A. (2010). Assessing the quality of mixed methods research: Toward a comprehensive framework. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (2nd ed., pp. 305–338). Thousand Oaks, CA: Sage.

Onwuegbuzie, A. J., & Johnson, R. B. (2004). Mixed method and mixed model research. In B. Johnson & L. Christensen (Eds.), *Educational research: Quantitative, qualitative, and mixed approaches* (pp. 408–431). Boston, MA: Allyn & Bacon.

Onwuegbuzie, A. J., & Johnson, R. B. (2006). The validity issue in mixed research. *Research in the Schools*, *13*(1), 48–63.

Onwuegbuzie, A. J., & Leech, N. L. (2007). Validity and qualitative research: An oxymoron? *Quality & Quantity*, *41*, 233–249.

Paino, M., Renzulli, L., Boylan, R., & Bradley, C. (2014). For grades or money? Charter school failure in North Carolina. *Educational Administration Quarterly*, 50(3), 500–536.

Patton, M. Q. (2014). *Qualitative research and evaluation methods: Integrating theory and practice* (4th ed.). Thousand Oaks, CA: Sage.

Peräkylä, A. (1997). Reliability and validity in research based on tapes and transcripts. In D. Silverman (Ed.), *Qualitative research: Theory, method, practice* (pp. 201–220). London, United Kingdom: Sage.

Reynolds, J., DiLiberto, D., Mangham-Jefferies, L., Ansah, E. K., Lal, S., Mbakilwa, H., . . . Chandler, C. I. (2014). The practice of "doing" evaluation: Lessons learned from nine complex intervention trials in action. *Implementation Science*, *9*(75), 1–12.

Shadish, W. R. (1995). The logic of generalization: Five principles common to experiments and ethnographies. *American Journal of Community Psychology*, *23*, 419–428.

Shadish, W. R., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Spillane, J. P., Pareja, A. S., Dorner, L., Barnes, C., May, H., Huff, J., & Camburn, E. (2010). Mixing methods in randomized controlled trials (RCTs): Validation, contextualization, triangulation, and control. *Educational Assessment, Evaluation and Accountability, 22*(1), 5–28.

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

Tracy, S. J. (2010). Qualitative quality: Eight "big-tent" criteria for excellent qualitative research. *Qualitative Inquiry, 16*(10), 837–851.

Turner, S., & Coen, S. E. (2008). Member checking in human geography: Interpreting divergent understandings of performativity in a student space. *Area, 40*(2), 184–193.

Wells, M., Williams, B., Treweek, S., Coyle, J., & Taylor, J. (2012). Intervention description is not enough: Evidence from an in-depth multiple case study on the untold role and impact of context in randomised controlled trials of seven complex interventions. *Trials, 13*(95), 1–17.

Wolcott, H. F. (1990). On seeking—and rejecting—validity in qualitative research. In E. W. Eisner & A. Peshkin (Eds.), *Qualitative inquiry in education: The continuing debate* (pp. 121–152). New York, NY: Teachers College Press.

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Thousand Oaks, CA: Sage.

8

Ethical Considerations

Kimberley A. Cox, PhD

Introduction

Research ethics will have far-reaching implications on decisions you make throughout the various stages of your research including, but not limited to, considerations about participant recruitment, measurement, data storage and analysis, and dissemination. Although the research designs covered in previous chapters on quantitative, qualitative, and mixed methods research each carry unique ethical implications, they share some common considerations regarding the ethical treatment and protection of research participants. It is these commonalities that are the focus of this chapter.

In this chapter, I address ethical issues of particular relevance to and most commonly encountered by student researchers in the social, behavioral, management, health, and educational disciplines. This chapter is also appropriate for experienced researchers in these disciplines who seek a refresher in research ethics. The modest coverage of ethics adopted here acknowledges that ethical issues are often context dependent. Therefore, you are encouraged to discuss your proposed study with your institution's ethics review board and supervisory committee early in the stages of research planning.

I begin with a brief historical overview of codes and regulations governing ethical conduct in research followed by the role of professional associations in research ethics. I then follow with coverage of the essential components of ethical research planning, such as assessing risks and benefits, the role of institutional review boards, and the informed consent process. Then, I discuss common ethical challenges, such as conducting research with vulnerable populations, maintaining privacy, and using deception and debriefing techniques. Last, I address the unique ethical implications of Internet-based research.

Ethics Codes and Regulations

The development of the first international code of protections for human research participants dates back to the first part of the 20th century, most notably with the 1946–1947 Nuremberg medical trial held in Nuremberg, Germany. The *Nuremberg Code* arose as a result of this trial, during which it was revealed that German doctors and others conducted involuntary human medical experimentation in Nazi concentration camps and clinics (Weindling, 2001).

The Nuremberg Code contains a set of basic ethical principles that has influenced subsequent research ethics codes. These principles emphasized certain conditions to protect research participants, including voluntary informed consent, avoidance of harm, assessment of risk, right to withdrawal, and the researcher's responsibility to terminate an experiment if its continuation may pose harm to the participant ("Trials of War Criminals," as cited in U.S. Department of Health and Human Services [HHS], 2005).

In the United States, a history of egregious research abuses,[1] such as the Tuskegee and Willowbrook studies, led to the 1974 National Research Act (Public Law 93-348), which directed establishment of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (HHS, n.d.) and included the requirement for review boards at research institutions. This commission, through publication of the Belmont Report (U.S. Department of Health, Education, and Welfare, 1979), identified three ethical principles—respect for persons, beneficence, and justice—and established guidelines for conducting research to ensure those principles are observed.

The ethical principle of *respect for persons* refers to respecting autonomy, which involves acknowledgment of an individual's autonomy to make personal choices regarding research participation and protections from harm for individuals with diminished or impaired ability to exercise autonomy. *Beneficence* refers to an obligation to *do no harm* and to the need for research to maximize potential benefits and minimize possible harms. *Justice* refers to treating individuals in a way that is morally right, such that the selection of research participants and any potential benefits and burdens of participating in research are equitable among groups (U.S. Department of Health, Education, and Welfare, 1979).

The Belmont Report influenced current protections in place for human research participants in biomedical and behavioral research in the United States. The three ethical principles were translated in 1991 by the work of over a dozen U.S. federal departments and agencies, such as the HHS[2] and the U.S. Food and Drug Administration, into the *Federal Policy for the Protection of Human Subjects*, commonly referred to as the *Common Rule* (HHS, n.d.). The Common Rule describes the regulations for protecting individuals who participate in research, including the requirements for informed consent and functions of institutional review boards.

On September 8, 2015, several U.S. federal departments and agencies published a notice of proposed rulemaking (NPRM)[3] to communicate proposed changes to the Common Rule. This NPRM seeks to modernize the U.S. federal regulations and in doing so has potential implications for researchers in the social, behavioral, management, health, and educational disciplines. Most notable are (a) the proposed changes to the informed consent process, such that it is more transparent and stringent in its requirements for communicating information

---

[1] For a brief historical overview of the most widely cited abuses in research, see Hardicre, J. (2014). An overview of research ethics and learning from the past. *British Journal of Nursing*, *23*(9), 483–486. doi:10.12968/bjon.2014.23.9.483.

[2] The HHS regulations are published as 45 C.F.R. § 46; see http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.html.

[3] See https://federalregister.gov/a/2015-21756.

about research to prospective participants, and (b) the proposed expansion of exclusion and exemption categories of low-risk research, which would be determined by researchers themselves using a web-based "decision tool" rather than through the mechanisms of an institutional review board review.

Professional Codes of Ethics

In addition to the U.S. federal regulations mentioned earlier, another source of guidance for designing ethically sound research is the ethical code and guidelines in your discipline. Many disciplines with affiliated professional associations have developed their own codes of ethics that reflect the common issues faced by researchers in their respective discipline and set standards for the protection of research participants. Box 8.1 provides examples of some professional association codes of ethics and where to locate them.

Many professional associations outside of the United States have also established ethics codes and guidelines, and over 100 countries have established regulations, laws, and/or guidelines for the protection of research participants.[4] Box 8.2 provides examples of some international codes of ethics and where to locate them.

Assessing Risks and Benefits

Early in the research planning stages, you should assess the possible risks and potential benefits of conducting your study. This assessment of possible costs or risks relative to the potential benefits or contributions of a proposed research study is commonly referred to as the *cost–benefit ratio* or *risk–benefit ratio*. The nature of this assessment is often subjective, so it is especially helpful at this early stage in research planning to consult ethical codes and regulations.

A research study may provide potential benefits to participants, such as a sense of satisfaction, self-insight, or new knowledge or skills acquired through participation. In terms of costs or risks, some research methods, such as observations of public behavior, typically hold no risk. Other methods, however, such as surveys and experiments, may place research participants *at risk* or at *minimal risk*. The distinction[5] between at risk and minimal risk is determined by considering the extent to which prospective participants will experience situations during the research study that are similar to those "ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests" (HHS, 2010; see 45 C.F.R. § 46.102.i). That is, minimal risk means that the discomfort or harm participants may experience does not exceed that which they might routinely experience in daily life or from physical or psychological tests considered routine in nature. For example, participants might experience minor

[4] For an extensive catalog of international regulations, laws, and guidelines, see http://www.hhs.gov/ohrp/ international/index.html for the latest edition of the International Compilation of Human Research Standards.

[5] For some case examples in assessing risk, see http://www.hhs.gov/ohrp/sachrp/sachrpminrisk20080131.html.

BOX 8.1

Professional Associations' Codes of Ethics

- **The American Anthropological Association**
  American Anthropological Association. (2012). *AAA statement of ethics*. Retrieved from http://ethics.americananthro.org/category/statement/
- **The American Counseling Association**
  American Counseling Association. (2014). *ACA code of ethics*. Retrieved from http://counseling.org/knowledge-center/ethics
- **The American Educational Research Association**
  American Educational Research Association. (2011). *Code of ethics*. Retrieved from http://www.aera.net/AboutAERA/AERARulesPolicies/CodeofEthics/tabid/10200/Default.aspx
- **The American Evaluation Association**
  American Evaluation Association. (2008). *Guiding principles for evaluators*. Retrieved from http://www.eval.org/p/cm/ld/fid=51
- **The American Political Science Association**
  American Political Science Association Committee on Professional Ethics. (2012). *A guide to professional ethics in political science* (2nd ed.). Washington, DC: Author. Retrieved from https://oldapsa.apsanet.org/ content_9350.cfm
- **The American Psychological Association**
  American Psychological Association. (2010). *Ethical principles of psychologists and code of conduct*. Retrieved from http://www.apa.org/ethics/code/index.aspx
- **The American Public Health Association**
  Thomas, J. C., Sage, M., Dillenberg, J., & Guillory, V. J. (2002). A code of ethics for public health. *American Journal of Public Health*, *92*(7), 1057–1059. doi:10.2105/AJPH.92.7.1057
- **The American Sociological Association**
  American Sociological Association. (1999). *Code of ethics.* Retrieved from http://www.asanet.org/about/ethics.cfm
- **The Project Management Institute**
  Project Management Institute. (n.d.). *Code of ethics and professional conduct*. Retrieved from http://www.pmi.org/About-Us/Ethics/Code-of-Ethics.aspx
- **The Society for Human Resource Management**
  Society for Human Resource Management. (2014). *Code of ethics*. Retrieved from http://www.shrm.org/about/pages/code-of-ethics.aspx

frustration or stress in a survey study that asks about their spending habits or in an experimental study that involves a timed performance task under observation. The discomfort that may arise from such frustration or stress in these examples would likely be considered minimal risk because an individual might experience similar levels of minor frustration or stress in the course of daily life. If, however, the participants' experience during a research study is such that the degree of discomfort may exceed what might be expected in daily life or from physical or psychological tests considered routine in

BOX 8.2

International Codes of Ethics

- **Australia**
  National Statement on Ethical Conduct in Human Research: http://www.nhmrc.gov.au/guidelines/publications/e72
  Australian Code for the Responsible Conduct of Research: http://www.nhmrc.gov.au/guidelines/publications/r39
- **Canada**
  Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans, second edition: http://www.pre.ethics.gc.ca/eng/policy-politique/initiatives/tcps2-eptc2/Default/
- **Cyprus**
  The Safeguarding and Protection of Patients' Rights Law: http://www.bioethics.gov.cy/Law/cnbc/cnbc.nsf/All/6960B7A5AA76C4A3C22571C9002B99F0?OpenDocument
- **India**
  Ethical Guidelines for Research in Social Sciences and Health: http://www.cehat.org/publications/ethical.html
- **South Africa**
  Ethics in Health Research: Principles, Structures, and Processes: http://www.health.uct.ac.za/usr/health/research/hrec/links/Department_of_Health-Ethics_in_Health_Research_Principles_Structures_and_Processes_2004.pdf
- **United Kingdom**
  Research Governance Framework for Health and Social Care: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4108962

nature, then prospective participants are considered to be at risk. These risks include research that may pose psychological and/or physical harm to participants. For example, participants might experience embarrassment or anxiety during a research study that requests disclosure of sensitive aspects of their private lives. Although physical harm is less likely in research conducted in the social, behavioral, management, health, and educational disciplines, stressful conditions could potentially lead to physical discomfort.

It is your responsibility, as the researcher, to assess the cost–benefit ratio of your study and clearly articulate the ethical safeguards in place to protect participants from possible risks when submitting your research proposal to your institution's ethics review board or committee. In the next section, I discuss the role of the ethics review board or committee and highlight the process you can expect when submitting a research proposal for review.

Institutional Review Boards

U.S. federal regulations require that institutions (including colleges and universities, private research firms, and governmental agencies) that receive federal funding for biomedical or behavioral sciences research with human participants have an *institutional review board* (IRB) in place to assess the ethical compliance of proposed research studies and to oversee the conduct of research to protect participants, researchers, and institutions (HHS, n.d.). There are many institutions that neither seek nor receive federal funding for research, but that voluntarily apply federal regulations when reviewing and overseeing research. Many institutions located outside of the United States also have external bodies that review proposed research, such as Australia's Human Research Ethics Committee, Canada's Research Ethics Board, and the United Kingdom's local Research Ethics Committees.

U.S. federal regulations stipulate that an IRB be composed of at least five members from various disciplines representing varied backgrounds. These members must include at least one individual from the outside community who is not affiliated with the institution, one from the sciences, and one who is considered a nonscientist or whose background falls outside of scientific areas (HHS, 2010; see 45 C.F.R. § 46.107).

Before your data collection can begin, you must submit your proposed study to your institution's IRB for review. This review process typically requires your completion of an application that consists of a series of questions about your research accompanied by copies of your research proposal, informed consent form, and related study materials, such as survey instruments and letters of agreement or cooperation from community research partners. Research partners are outside entities, such as a school or agency, involved in participant recruitment or data collection. Because an IRB may impose requirements that expand upon federal regulations, you should seek information on the IRB review process at your institution early in the research planning stages. Most IRBs will require confirmation of the researcher's completion of ethics training. Box 8.3 provides resources that you can use to locate ethics training if you need to satisfy this requirement; it is highly recommended that you complete this training even if it is not required. Check with your institution's IRB for any training preference or requirement.

In conducting its review, the IRB is particularly concerned about the possibility of harm or risk to participants. The IRB therefore weighs the potential benefits or contributions of the proposed research in relation to its possible risks. The IRB evaluates the safeguards that you have proposed to minimize risk to your participants, such as the informed consent process, which is covered in the next section.

There are various types of IRB reviews, including full, expedited, and exempt. The distinctions and determinations rest with the IRB and can vary by site. It is worth mentioning that although U.S. federal regulations establish the research categories that may qualify for an exempt or expedited review (see 45 C.F.R. § 46.110), an IRB can set its own standards for what is exempt and expedited, and it may require more review than what is minimally required by the federal regulations. Therefore, check with your institution's IRB to determine the type of review that will apply to your proposed research study.

If you are in a discipline such as management, education, or psychology and you are proposing to conduct your research in your place of employment or collect your data from your own employees, students, or clients, you may face potential ethical dilemmas unique to these settings. These ethical dilemmas primarily involve confidentiality concerns and conflict of interest. One example

BOX 8.3

Resources for Research Ethics Training

- The **Collaborative Institutional Training Initiative (CITI)** offers a series of modules on human subjects research that is geared toward disciplines in the social, behavioral, management, and educational sciences at no cost for users affiliated with a CITI program–subscribing institution. Unaffiliated users must pay a fee. Registration is required for login access. A report is generated upon course completion.

  https://www.citiprogram.org/

- The **Family Health International 360** offers a no-cost research ethics training curriculum in English, French, Portuguese, and Spanish. A certificate is generated upon successful completion of the test.

  http://www.fhi360.org/en/RH/Training/trainmat/ethicscurr/index.htm

- The **Global Health Training Center** offers several no-cost short courses and modular courses in research ethics. Registration is required for login access. A certificate is generated upon successful completion of all quiz sections.

  https://globalhealthtrainingcentre.tghn.org/elearning/research-ethics/

- The **National Institutes of Health** offers no-cost training in protecting human research participants. Registration is required for login access. A certificate is generated upon successful completion of all quiz sections.

  https://phrp.nihtraining.com/users/login.php

Training opportunities also exist for specific research designs and target populations. For example, **CITI** (https://www.citiprogram.org) offers training on topics related to children in research, cultural competence, international research, and Internet-based research and the **HHS** (http://www.hhs.gov/ocr/privacy/hipaa/understanding/training/index.html) provides training materials for researchers proposing to access protected personal health information.

of this is when you hold dual roles, such as teacher and researcher. There is also the possibility that participants may feel unduly pressured to participate or may respond in a biased manner, both of which can compromise the validity of data collected. Two common methods for mitigating these risks are to collect data anonymously or to perform a secondary analysis on existing records, such as employee assessments or therapeutic records, which have had all information that could identify a participant removed.

The outcome of an IRB's review can include notice of approval, required modifications, or disapproval (see 45 C.F.R. § 46.109). If your proposed study involves a community research partner, the research proposal will likely need to be reviewed by that partner's review board or committee before your institution grants approval. Notice of approval from an IRB to begin a research study does not end your responsibility to comply with ethical standards and guidelines. You also may be required to renew your approval on a periodic basis, so consult with your institution on policies regarding renewals.

Informed Consent

Informed consent is the ongoing process of the communication of information between you, the researcher, and prospective participants. It begins at the recruitment stage of your study and may continue for its duration (HHS, 2013). Informed consent is based on the notion that it is an individual's right to make an informed and voluntary decision about participating in a research study.[6]

As you approach and manage the process of informed consent, it is important to consider different cultural perceptions and beliefs that may exist about research and decision making. For example, "In the United States, individuals are expected to make decisions for themselves. . . . However, in many places throughout the world, decisions are not necessarily made by the individual, but instead by family members or community representatives" (Citro, Ilgen, & Marrett, 2003, p. 96).

Informed Consent Form

A *consent form* is a common method for informing prospective research participants about their potential involvement in a research study and documenting (often by written signature) their voluntary willingness to participate. An informed consent form should reflect the three ethical principles identified in the Belmont Report: respect for persons, beneficence, and justice. For example, the principle of respect for persons is supported in an informed consent form that discloses sufficient information about the study so that prospective participants can freely decide, without coercion, whether or not to participate (HHS, 2013). The principle of beneficence is supported with a description of any anticipated risks and benefits associated with participation. Finally, the principle of justice is supported when the consent form explains who will benefit from the research and discloses any alternative treatments or procedures, if applicable.

Many institutions provide a consent form template for their researchers. If a template is not available for you, the ethical principles found in your discipline's professional code of ethics are a good source to review for guidance. The following key components represent the content that is found in a typical consent form; an informed consent form should

- Briefly explain the purpose of the research
- State the expected duration of participation in the study
- Describe the study's procedures (i.e., what the participant will do)
- Describe any experimental procedures and disclose any alternative treatments or procedures, if applicable
- Describe any anticipated discomforts or risks associated with participation

[6] For a background on the history of informed consent, see Faden, R. R., & Beauchamp, T. L. (1986). *A history and theory of informed consent*. New York, NY: Oxford University Press; and Weindling, P. (2001). The origins of informed consent: The international scientific commission on medical war crimes, and the Nuremberg Code. *Bulletin of the History of Medicine*, *75*(1), 37–71. doi:10.1353/bhm.2001.0049.

BOX 8.4

Tips on Writing the Informed Consent Form

- Avoid scientific jargon. Write clearly using language that is easily understandable by the target research population (HHS, 1993).
- Write at an appropriate reading level. It is common practice to write the consent form at an eighth-grade reading level. Word processing programs typically have tools to check the grade level and readability (i.e., reading ease) of a document.[7] For example, in Microsoft Word, these tools are accessible by enabling the readability statistics.
- Use headings in bold type to differentiate the main sections of the consent form.
- Use an active voice to address the prospective participant directly. For example, write, "You will be asked to complete a survey . . . ," rather than "A survey will be completed . . . ."
- Show sensitivity to cultural differences by using language that is sensitive to the target research population (American Psychological Association, 2010). Consult with your institution's IRB about when language translation and back-translation may be necessary.

- Describe any anticipated benefits of participation
- Describe any incentives (e.g., monetary compensation) for participation, if applicable
- State that participation is voluntary and refusal to participate or discontinuation of participation at any time will pose no negative consequences
- Describe how the confidentiality of data will be maintained
- State the names and contact information of at least two individuals—one who can answer questions about the research and another who can answer questions about rights as a participant

Most of these components are self-explanatory, but some warrant further explanation, and some will require tailoring to your particular study more than others, as described below.

In describing your study's procedures, it is best to keep the description focused on what participants will actually do in language that is clear and easily understood to nonscientists. A detailed discussion or even mention of, for example, your study's research questions, hypotheses, theoretical framework, or statistical analysis is typically omitted to avoid potentially biasing participants' responses or weakening the effectiveness of a study. Box 8.4 provides some recommendations on writing the informed consent form.

Any benefits of participation should be realistically described within the bounds of your study. You should avoid making claims or setting expectations

---

[7] For more on readability of consent forms, including examples of consent content ranging from fourth-grade to college reading level, see Paasche-Orlow, M. K., Taylor, H. A., & Brancati, F. L. (2003). Readability standards for informed-consent forms as compared to actual readability. *New England Journal of Medicine*, *348*(8), 721–726. doi:10.1056/NEJMsa021212.

for a benefit that participation cannot genuinely deliver (Rosenthal, 1994). Often, the benefits for research participants fall within the domain of satisfaction and, in some cases, self-insight, such as when the results are shared with each participant.

If an incentive such as monetary compensation will be offered to participants, you should give consideration to an appropriate amount that is not excessive in order to avoid unduly influencing prospective participants' decision to voluntarily participate. If there are conditions (because of the study's design) where participants might receive no incentive or a partial incentive, you should clearly explain these conditions in the informed consent form. For example, if your study involves participation across multiple points of time, it is best to provide compensation at each time of participation to avoid unduly influencing participants' decision to continue.

If your proposed research will involve students as participants, it is a common and generally acceptable practice to offer extra credit, if the amount is reasonable and an alternative to research participation is also offered to obtain the extra credit. This is done to avoid unduly influencing students' decision to voluntarily participate. Subject pools at colleges and universities are growing in popularity as a means to recruit students who are willing to participate in research. It is important to remember that registration with such a pool does not qualify as informed consent because its registrants do not yet have the information they need to make an informed decision about whether or not to participate in a particular study. Unless consent has been waived by your institution's IRB, you must obtain informed consent from prospective participants who are recruited from subject pools.

In stating the voluntary nature of participation, include that an individual can refuse to participate or discontinue participation at any time without negative consequences in such a way that it is tailored to your research setting. For example, for a research study proposed in a workplace setting, this section would include a statement to the effect that one's decision to participate or not participate will have no effect on employment or treatment by the employer.

You should address privacy measures in the informed consent form with a statement that describes how participants' identity and data will be protected. Anonymity and confidentiality are two methods for safeguarding privacy. *Authentic anonymity* (L. Endicott, personal communication, August 13, 2015) is only possible when the identities of participants are not known to anyone, including you—this means that no identifying information is collected from participants. This type of authentic anonymity differs from anonymizing the data so that it no longer contains identifying information (L. Endicott, personal communication, August 13, 2015). *Confidentiality* refers to the situation when the identities of participants are known, but identifying information is not shared without their consent or the reporting of data does not reveal participants' identities or permit their identities to be inferred. You should state in the consent form whether there are any limits or risks to confidentiality. For example, in some disciplines, such as psychology, confidentiality may be broken in some situations when mandated by law, such as to protect a participant from self-harm or harm to others (American Psychological Association [APA], 2010; see Standard 4.05). You should check for any local and state mandating laws that may apply to your discipline. Box 8.5 describes some common methods for safeguarding participants' privacy.

BOX 8.5

Common Methods for Protecting Privacy

When a study's design allows for it, anonymity protects the privacy of participants' data. The simplest way to do this is to not request any identifying information from participants. For example:

- For a survey study conducted online, consent can be implied by participants' completion of the survey in some situations. If approved by an IRB, the initial webpage that greets participants would include components of an informed consent form, but no written or electronic signature would be obtained that could reveal participants' names. To further safeguard participants' privacy, settings can be disabled within the survey program so it does not collect participants' unique Internet protocol (IP) addresses.
- For a survey research study conducted by postal mail, in some situations, participants' consent can be implied by their response to the survey questions and return of the survey by mail. If approved by an IRB, a written statement provided with the study materials would include components of an informed consent form, but no signature would be obtained that could reveal participants' names. The return envelope and study materials would not contain any information that could potentially identify a participant.

When anonymity is not possible due to the study's design, the following examples provide some common methods for ensuring confidentiality:

- For a survey study conducted online that requires multiple points of access for completion and, thus, the matching of participants' data across time, participants could be instructed to create a unique identifier only known to them to repeatedly gain access to the study.
- For a survey research study conducted by postal mail that calls for matching participants with their data, the participants could be assigned an identification number. The master file with the identification numbers and the signed consent forms should be stored in a secure (i.e., locked) location separately from the data. This same method could also apply to studies conducted in person.
- Computer data files should be password protected, and knowledge of the password and access to the data should be limited to authorized persons only.
- If your research is on topics deemed highly sensitive, such as drug use or illegal behavior, you may be able to obtain a Certificate of Confidentiality to protect the release of identifying information about participants as a result of a court order or subpoena.[8]

[8] For more information on the Certificate of Confidentiality, see http://grants1.nih.gov/grants/policy/coc/index.htm and http://www.hhs.gov/ohrp/policy/certconf.html.

The names and contact information for at least two individuals should be provided on the consent form. You, as the researcher, are typically listed as the contact for questions about the research study, and an ombudsperson, IRB, or administrative representative is typically listed as the contact for questions about the rights of research participants.

Under certain conditions, your institution's IRB may allow you to modify components of or waive informed consent, per federal regulations (see 45 C.F.R. §§ 46.116.c, 46.116.d, and 46.117). These conditions are applicable to some research commonly conducted in social, behavioral, management, health, and educational disciplines. IRBs may also require the documentation of additional information that is not listed in the federal regulations (HHS, 1993; see 45 C.F.R. § 46.109.b). Therefore, you should check with your institution for any requirements or guidelines on writing the informed consent form.

Research With Vulnerable Populations

Special protections are necessary during the informed consent process for research involving members of vulnerable populations, such as children, residents of nursing homes, prisoners, persons with mental disability, or persons who are economically or educationally disadvantaged (HHS, 2010; see 45 C.F.R. § 46.111.b). Because informed consent implies the ability to understand information about a research study and freely decline participation, individuals who lack either of these are afforded extra protections during the recruitment and data collection processes. When an individual from a vulnerable group is not legally capable of informed consent, you must obtain informed consent from that individual's legal guardian while also obtaining the individual's *assent*. Assent refers to a verbal (or nonverbal) expression of agreement from the prospective participant to participate in the study (Roth-Cline & Nelson, 2013).

Children, for example, do not have the legal capacity to provide informed consent (see 45 C.F.R. Part 46, Subpart D). Therefore, the child's parents or legal guardian must provide informed consent on their child's behalf and the child must have an opportunity to provide or withhold assent. More complex informed consent procedures exist for research with prisoners and persons with mental disability that are beyond the scope of this chapter. Consult with your institution's IRB and applicable regulations for guidance on conducting research with any vulnerable population.

Deception in Research

In some circumstances, it may be permissible to intentionally withhold some information about a study from participants—a research technique that is referred to as *deception*. Deception involves not fully informing participants about a study's purpose or misleading participants about the true nature of a study's procedures.[9] When it is used, deception is typically justified on

[9] For a taxonomy of deception methods, see Sieber, J., Iannuzzo, R., & Rodriguez, B. (1995). Deception methods in psychology: Have they changed in 23 years? *Ethics and Behavior*, *5*, 67–85. doi:10.1207/s15327019eb0501_5.

methodological grounds as a means to obtain unbiased data on attitudes and behaviors that will be more true to real life than what could be expected in an artificial research setting if the participants were fully informed (Fisher & Fyrberg, 1994). In conjunction with the IRB, you might consider including a forewarning statement in the informed consent form to let participants know in advance that the study's purpose or some of its procedures will not be accurately described so they can consider whether or not to participate.

When participants are misled or not fully informed about a study's true purpose or procedures, this is typically done through a *cover story*, which refers to the presentation of a false description to lead participants into thinking that the study's purpose and/or procedures are different than their true intention. One of the most famous examples of a study that used deception and a cover story is Stanley Milgram's (1963) behavioral study of obedience. Participants were told that the purpose of the study was to examine the effects of punishment on their memory, and they were intentionally led to believe that they were delivering electric shocks of increasing intensity to a fellow participant when, in reality, there were no shocks delivered. The fellow participant was a research confederate, and the true purpose of Milgram's study was to examine obedience to authority. Another example of deception is giving participants false feedback on a task, such as a quiz or activity, regardless of their true performance.

Deception has a history of generating considerable controversy within the research community (see Baumrind, 1964; Milgram, 1964) and public, and it remains ethically controversial (Fisher, 2005; Kimmel, 2012). If you are proposing a study that involves deception, you must assume considerable responsibility to weigh the possible risks associated with the deception in relation to the potential benefits of conducting the study. The APA's (2010) code of ethics (see Standard 8.07) permits the use of deception in research if the study lacks the means for an alternative nondeceptive technique and the deception is not expected to cause significant distress or pain, is justified by the study's importance, and is explained to participants as soon as possible. This latter condition, of explaining the deception to participants, is referred to as *debriefing*. To mitigate any possible negative effects, you should provide participants with an explanation of the deception used and its rationale during a debriefing session. When handled successfully, a debriefing session achieves two outcomes: postdeception dehoaxing and postdeception desensitizing (Holmes, 1976a, 1976b). Dehoaxing refers to explaining the true purpose of the study and correcting any misleading or false information, and desensitizing refers to reducing any negative effects participants may have experienced, such as stress. You should also allow time to address any questions or concerns in the debriefing session. When research is conducted solely online, informed consent, deception, and debriefing can pose unique challenges. These challenges are described in the next section on Internet research.

Ethical Implications of Internet Research

According to a January 2014 survey of a nationally representative sample of American adults, about 80% reported using a computer at school, work, and/or home, and similar rates were found for Internet usage (Pew Research Center, 2014). The widespread use of computers and access to the World Wide Web have contributed to the growing use of the Internet as a medium for participant recruitment and data collection. Indeed, Internet-based research has been growing in popularity because of the advantages and benefits it can afford researchers, such as access to large, diverse populations, reduced risk of researcher bias, and cost effectiveness, to name a few (Allen & Roberts, 2010; Kraut et al., 2004; Nosek, Banaji, & Greenwald, 2002). With these advantages and benefits come unique ethical challenges and responsibilities for researchers, including managing the online informed consent process, protecting participants' privacy, debriefing, and maintaining data security (Allen & Roberts, 2010; Emery, 2014; Hoerger & Currell, 2012).

Internet-based research is most commonly used for studies that employ online observations, surveys, and experiments, all of which require upholding the ethical principles of respect for persons, beneficence, and justice. The cost–benefit ratio or assessment of possible risks relative to potential benefits should be conducted. The possibility of risks is often unique for any particular online study and, thus, must be considered in the context of the proposed research (Kraut et al., 2004) with safeguards tailored toward minimizing any possible harm.

Informed Consent

The ethical principle of respect for persons remains of upmost importance in Internet-based research given the unique challenges that exist with the online informed consent process. Managing this process effectively requires thoughtful consideration of the design of the study and the format of the online materials including, for example, how participants will proceed from one question or activity to the next and how they will contact you, the researcher, with any questions or concerns.

The informed consent form is often presented as the first webpage when a prospective participant gains access to the study's website. If the signature requirement is waived by the IRB, information about the study should still be presented on this introductory page with sufficient description of the study's purpose, followed by links or buttons for participants to click to indicate their decision to participate, such as *I agree* and *I decline*. When a signature is not required, a participant's consent is implied by his or her decision to complete the study.

For Internet-based research that may place participants at risk, additional safeguards should be put in place to ensure participants have read and understood all of the information presented in the informed consent form. Examples of safeguards include such strategies as presenting a series of buttons that read *I agree* or *Click to accept* following each section within the form or including a brief quiz at the end of the form to confirm comprehension (Kraut et al., 2004). Another strategy is to include a frequently asked questions (FAQs) document that covers possible concerns and questions with the consent form (Nosek et al., 2002).

If you intend for child or adolescent participants to gain access to your online study, you need to be familiar with relevant laws and regulations. For example, Internet-based research with minors is subject to the federal regulations of the Children's Online Privacy Protection Act (1998). When minors are not the intended participants of an online study, you must consider how to minimize and screen their potential access. One strategy, for example, is the use of a short quiz as described earlier (Hoerger & Currell, 2012; Kraut et al., 2004). Another strategy is to include a question that asks prospective participants if they are 18 years of age or older or asks for their date of birth with a response that indicated minor status resulting in termination of the study (Alessi & Martin, 2010). A more reliable and robust technique is to have prospective participants use a trusted technology company's identification system to verify age.

When participants gain access to the online study materials, they should be permitted to skip any questions or items they wish rather than being presented with forced answer choices (Fox, Murray, & Warm, 2003). It is also recommended that a link or button to exit the study be provided on each page to reinforce the voluntariness of participation (Emery, 2014; Hoerger & Currell, 2012; Keller & Lee, 2003; Nosek et al., 2002). It is a good practice to design the online site to lead to a webpage containing contact information when a participant exits at any point within the study (in the case of early termination) or completes his or her participation (Emery, 2014). This can be done on a "Thank you" page that is presented upon exit. You can also include a statement to encourage participants to print the page or write down the contact information should they have any questions or concerns. The abovementioned techniques are important in Internet-based research because you are limited in your ability to know whether a participant had a negative experience. Thoughtful planning and designing of the online study site and materials serve as good preparation to maximize benefits and minimize negative outcomes.

Debriefing

Debriefing participants in an online study holds the same importance as it does in research that is not conducted online. You can use a variety of strategies for debriefing, such as (a) emailing participants a link to a webpage with debriefing content, (b) emailing participants debriefing content in the body of the message, or (c) directing participants to a debriefing webpage after exiting or completing the study.[10] The study's design, including whether participants' identities are anonymous or not, influences the type of strategy to use. For all possible strategies, the debriefing content should be updated as necessary and tailored to participants when appropriate (Kraut et al., 2004).

Privacy

Safeguarding online research participants' privacy poses unique ethical challenges. For example, you must consider the ways in which websites may identify and store information that could be used to potentially

---

[10] For other debriefing techniques, see Nosek, B., Banaji, M., & Greenwald, A. G. (2002). E-research: Ethics, security, design, and control in psychological research on the Internet. *Journal of Social Issues*, *58*(1), 161–176. doi:10.1111/1540-4560.00254.

detect participants' identities, such as IP addresses (Nosek et al., 2002). Researchers have also drawn attention to the potential for threats to privacy when external online survey providers are used to collect and store (often temporarily) participant data (Allen & Roberts, 2010).

Most online survey software programs allow for disabling IP address tracking, and you can ask participants to create a unique code to be used for tracking, if necessary, instead of collecting IP addresses (Hoerger & Currell, 2012). If you intend to use a commercial survey provider to host, collect, and store your data rather than an internal web server at your institution, you should familiarize yourself with the provider's policies, including in what situations (e.g., a court order) the provider might disclose identifying participant information. If you are planning to conduct international research, familiarize yourself with any possible regulations that pertain to online privacy and use of the Internet to collection data in the country of interest.

Conclusion

I began this chapter with a brief historical background to provide context to the current regulatory frameworks in place to protect research participants. I then highlighted the role of professional associations' ethics codes as a source of guidance as you plan your research and write your informed consent form.

I presented the distinction between levels of risk that research may pose to participants, as well as the importance of weighing the possible risks of the study in relation to its potential benefits. The role of IRBs was discussed, including their membership and function. I encouraged you to discuss your proposed study with your institution's ethics review board and supervisory committee early in the research planning stages to ensure a smooth review.

I emphasized informed consent as a process to draw attention to your responsibilities to maintain ongoing communication with participants throughout the duration of your study. The key elements of the informed consent form were presented, including its relation to the three ethical principles identified in the Belmont Report—respect for persons, beneficence, and justice—and the special protections that must be afforded to members of vulnerable populations, such as children, were noted. I presented tips for writing the consent form as best practices for clearly, appropriately, and respectfully conveying information about the study to the target population. At the time of this publication, several U.S. federal departments and agencies published notice of proposed changes to the informed consent process and consent form. Therefore, I recommend that you consult with your institution's IRB regarding any new requirements that may pertain to informed consent in your proposed research study.

I then offered some common safeguards for maintaining participant privacy as a source of guidance on methods you can use to protect participants' identities and data. In addition to understanding how federal regulations may apply to your research, you were also encouraged to check local and state laws that may pertain to your particular study.

I briefly discussed deception, including why it may be ethically justifiable in some research, and mentioned the important role of a debriefing session to mitigate any possible negative repercussions stemming from its use. Last, I briefly discussed the unique ethical implications of Internet-based research.

Although I did not provide an exhaustive review of the ethical issues that you may encounter in quantitative, qualitative, or mixed methods research, I nevertheless presented an overview of some essential topics in research ethics to support your efforts to design and implement ethically sound research. The suggestions and best practices offered here are meant to serve as the start of a conversation among you and your institution, supervisory committee, and colleagues.

Acknowledgment

I thank Leilani Endicott, PhD, for her helpful comments on an earlier draft of this chapter.

KEY SOURCES

All online resources presented here were current at the time of printing.

Institute of Medicine. (1995). *On being a scientist: Responsible conduct in research* (2nd ed.). Washington, DC: National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=4917

National Research Council. (2014). *Proposed revisions to the Common Rule for the protection of human subjects in the behavioral and social sciences*. Washington, DC: National Academies Press.

Sigma Xi. (1999). *The responsible researcher: Paths and pitfalls*. Retrieved from https://www.sigmaxi.org/programs/ethics-and-research/ethics-publications

Sigma Xi. (2000). *Honor in science*. Retrieved from https://www.sigmaxi.org/programs/ethics-and-research/ethics-publications

Steneck, N. H. (2007). *Introduction to the responsible conduct of research*. Retrieved from http://ori.hhs.gov/ori-introduction-responsible-conduct-research

REFERENCES

Allen, P. J., & Roberts, L. D. (2010). The ethics of outsourcing online survey research. *International Journal of Technoethics*, *1*(3), 35–48. doi:10.4018/jte.2010070104

Alessi, E. J., & Martin, J. I. (2010). Conducting an Internet based survey: Benefits, pitfalls and lessons learned. *Social Work Research*, *34*(2), 122–128. doi:10.1093/swr/34.2.122

American Psychological Association. (2010). *Ethical principles of psychologists and code of conduct*. Retrieved from http://www.apa.org/ethics/code/index.aspx

Baumrind, D. (1964). Some thoughts on ethics of research: After reading Milgram's "Behavioral Study of Obedience." *American Psychologist*, *19*(6), 421–423. doi:10.1037/h0040128

Children's Online Privacy Protection Act, 13 U.S.C. §§ 1301–1308 (1998). Retrieved from http://www.ftc.gov/ogc/coppa1.htm

Citro, C., Ilgen, D., & Marrett, C. (2003). *Protecting participants and facilitating social and behavioral sciences research*. Retrieved from http://www.nap.edu/catalog.php?record_id=10638

Emery, K. (2014). So you want to do an online study: Ethics considerations and lessons learned. *Ethics & Behavior*, *24*(4), 293–303. doi:10.1080/10508422.2013.860031

Fisher, C. B. (2005). Deception research involving children: Ethical practices and paradoxes. *Ethics & Behavior*, *15*(3), 271–287. doi:10.1207/s15327019eb1503_7

Fisher, C. B., & Fyrberg, D. (1994). Participant partners: College students weigh the costs and benefits of deceptive research. *American Psychologist*, *49*(5), 417–427. doi:10.1037/0003-066X.49.5.417

Fox, J., Murray, C., & Warm, A. (2003). Conducting research using web-based questionnaires: Practical, methodological, and ethical considerations. *International Journal of Social Research Methodology*, *6*(2), 167–180. doi:10.1080/13645570210142883

Hoerger, M., & Currell, C. (2012). Ethics issues in internet research. In S. J. Knapp, M. C. Gottlieb, M. M. Handelsman, & L. D. VanderCreek (Eds.), *APA handbook of ethics in psychology: Vol. 2. Practice, teaching and research* (pp. 385–400). Washington, DC: American Psychological Association. doi:10.1037/13272-018

Holmes, D. S. (1976a). Debriefing after psychological experiments: I. Effectiveness of postdeception dehoaxing. *American Psychologist*, *31*(12), 858–867. doi:10.1037/0003-066X.31.12.858

Holmes, D. S. (1976b). Debriefing after psychological experiments: II. Effectiveness of postdeception desensitizing. *American Psychologist*, *31*(12), 868–875. doi:10.1037/0003-066X.31.12.868

Keller, H. E., & Lee, S. (2003). Ethical issues surrounding human participants research using the Internet. *Ethics & Behavior*, *13*(3), 211–219. doi:10.1207/S15327019EB1303_01

Kimmel, A. J. (2012). Deception in research. In S. J. Knapp, M. C. Gottlieb, M. M. Handelsman, & L. D. VanderCreek (Eds.), *APA handbook of ethics in psychology: Vol. 2. Practice, teaching and research* (pp. 401–421). Washington, DC: American Psychological Association. doi:10.1037/13272-019

Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., & Coupler, M. (2004). Psychological research online: Report of Board of Scientific Affairs' Advisory Group on the conduct of research on the Internet. *American Psychologist*, *59*(2), 105–117. doi:10.1037/0003-066X.59.2.105

Milgram, S. (1963). Behavioral study of obedience. *The Journal of Abnormal and Social Psychology*, *67*(4), 371–378. doi:10.1037/h0040525

Milgram, S. (1964). Issues in the study of obedience: A reply to Baumrind. *American Psychologist*, *19*(11), 848–852. doi:10.1037/h0044954

Nosek, B., Banaji, M., & Greenwald, A. G. (2002). E-research: Ethics, security, design, and control in psychological research on the Internet. *Journal of Social Issues*, *58*(1), 161–176. doi:10.1111/1540-4560.00254

Pew Research Center. (2014). *The Web at 25 in the U.S.* Retrieved from http://www.pewinternet.org/2014/02/27/the-web-at-25-in-the-u-s/

Rosenthal, R. (1994). Science and ethics in conducting, analyzing, and reporting psychological research. *Psychological Science*, *5*(3), 127–134. doi:10.1111/j.14679280.1994.tb00646.x

Roth-Cline, M., & Nelson, R. M. (2013). Parental permission and child assent in research on children. *The Yale Journal of Biology and Medicine*, *86*(3), 291–301. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3767214/

U.S. Department of Health, Education, and Welfare. (1979). *The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research.* Washington, DC: National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, Office of the Secretary, U.S. Department of Health, Education, and Welfare. Retrieved from http://www.hhs.gov/ohrp/policy/belmont.html

U.S. Department of Health and Human Services. (1993). *Tips on informed consent*. Retrieved from http://www.hhs.gov/ohrp/policy/ictips.html

U.S. Department of Health and Human Services. (2005). *The Nuremberg Code*. Retrieved from http://www.hhs.gov/ohrp/archive/nurcode.html

U.S. Department of Health and Human Services. (2010). *Code of federal regulations*. Retrieved from http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.html

U.S. Department of Health and Human Services. (2013). *What is informed consent and when, why, and how must it be obtained?* Retrieved from http://www.hhs.gov/ohrp/policy/faq/informed-consent/what-is-informed-consent.html

U.S. Department of Health and Human Services. (n.d.). *Federal policy for the protection of human subjects ('Common Rule')*. Retrieved from http://www.hhs.gov/ohrp/humansubjects/commonrule/index.html

Weindling, P. (2001). The origins of informed consent: The international scientific commission on medical war crimes, and the Nuremberg Code. *Bulletin of the History of Medicine*, *75*(1), 37–71. doi:10.1353/bhm.2001.0049

9

Critiquing and Synthesizing the Literature

Annie Pezalla, PhD

Introduction

Understanding the basics of research methods will help you to critique the articles you are reading and then write a literature review that reflects synthesis rather than reporting descriptions of individual studies. That might be surprising news: You might have assumed that the lessons you have learned thus far about research methods are only applicable to a methods chapter in a master's thesis or doctoral study, but that is not true. Understanding all the lessons that have come before this chapter—the terminology of research, the philosophical underpinnings of research, the roles of theory, the various designs of research (qualitative, quantitative, and mixed methods), and the considerations of quality and ethics—will help you become a more critical consumer of information and, in turn, a better writer and thinker overall.

This chapter is organized into two major sections that reflect two important and parallel tasks in writing literature reviews: Critiquing the Literature and Synthesizing the Literature. Within both sections are four subsections: *Definitions*, where the basic meaning of each skill is elaborated and clarified; *Important Considerations*, where tips and guidelines are provided to facilitate the development of these skills; *Common Mistakes*, where examples are given of typical challenges and oversights from novice readers and writers; and *Ways to Stay Organized*, where strategies are offered to help keep you organized.

The important thing to note is that all the guidelines in this chapter are related. Being able to synthesize the literature, for example, requires that you have the ability to critique it in the first place. Similarly, being able to critique the literature requires that you understand how the author(s) synthesized the work to present a coherent narrative. The important thing to know is that all of these skills are acquired through an abundance of *reading*, not just on the information of this chapter but also in the world of published literature. Many blossoming researchers assume that they will become proficient writers by practicing their writing. They spend the bulk of their time writing, but less time reading. Although writing is certainly an important skill to hone, the more important skill to practice is that of reading the literature and of analyzing it for all its various parts. I start this chapter on doing just that.

Critiquing the Literature

Definitions

Critiquing the literature means becoming an active rather than a passive reader and learning how to *interrogate* the literature. Acquiring these skills does not mean using harsh or personal criticism; rather, it requires an impersonal scrutiny of a piece of work using a balanced and objective approach. The purpose of critiquing the literature is to highlight both a source's strengths and weaknesses in order to identify whether a piece of research is trustworthy and unbiased. As a scholar-practitioner, it is important that you have the ability to critically appraise the research in your field in order to identify and apply the best practices in your work.

Important Considerations

As a scholar-practitioner, you will be expected to read widely, but it is rarely possible to read *everything* that applies to your interests. To focus your attention on the most important, well-written, and credible sources, it is important to have some kind of reading process in place, one that "progresses from the general to the particular" (Hart, 2011, p. 53). One stage that is commonly recommended is a *preevaluation stage* that allows you to very quickly assess the utility of a source. A good place to start in this stage is to read the source's abstract, which is the short summary of the full article. The abstract can help determine if the study is relevant to your interests; it might also reveal the methods used in the study, the basic findings, and the recommendations for future research. If the abstract looks relevant, you can proceed.

After reading the abstract, note the author(s) who wrote the study. Some authors carry a great deal of credibility in their work, so seeing a familiar author name can validate the quality of a source. That does not suggest that an article should be dismissed because it was written by an unknown author, but it would behoove you to become well read on the *giants* of your field, those who are heavily published and well respected. If the source is written by a well-known author, or even if you recognize some of the names in the source's references section, the source is likely credible.

Another consideration in this preevaluation stage is to note the date of the publication. Unless the source is considered a *seminal* source—one that serves as a pivotal piece in your field, by reporting a major breakthrough or insight or a new synthesis of ideas—if it was published more than 5 years ago, it may be less valuable to you and your research. Our understanding of trends, methods, and theoretical applications is continually changing, so reading the most recently published work will generally benefit you the most.

The type of publication is also important to note in this stage. You will likely be required to read and report primarily peer-reviewed literature. This requirement is set in place to ensure that the materials you are reading have been vetted by an established panel of professionals who abide by a common set of standards in accepting submissions. Journal editors are charged to look for manuscripts that "(a) contribute significantly to the content area covered by the journal, (b) communicate with clarity and conciseness, and (c) follow style guidelines" (American Psychological Association, 2010, p. 226).

The peer reviewers of any peer-reviewed journal typically perform a masked review of the manuscript submissions to their journals; this means that they remain unaware of the author(s) who contributed the submission. This masked review process protects against any potential favoritism for or unfairness against authors who are familiar to the reviewers; it also helps mitigate any potential gender bias. The important thing to remember is that if you are reading a peer-reviewed article, the odds are in your favor that you are reading a high-quality piece of literature, one that has withstood a lot of scrutiny and is, subsequently, viewed as credible.

You may be tempted to read other types of resources for your research, such as popular journals. Popular sources could include magazines like *Business Week*, *Time*, or *People*. Those resources may be enjoyable to read because they tend to use more casual language with flashier claims than what you would find in a scholarly source. Walsh, Pezalla, and Marshall (2014) illustrated the language differences between popular sources and peer-reviewed sources, contrasting a popular source claim, “Grammar is dead! A new poll shows that text language is becoming more acceptable in the workplace, and we couldn’t be happier! LOL!” with a claim from a scholarly source, “The correct use of grammar in early college assignments has shown downward trends among young adults” (p. 61). Both sources share the same broad topic, but the popular source uses much different language, often with colloquial phrases to grab the reader’s attention. The scholarly source, in comparison, uses more formal language and precise claims.

Beyond the language differences between popular and peer-reviewed sources, the ideas within those sources may be held up to less rigorous standards. The content of a popular source is unlikely to have withstood the same high-intensity scrutiny of a peer-reviewed source. The authors of such sources may have used suspect methods to arrive at their claims, for example, or may have overlooked the work of seminal researchers in the field. A variety of considerations may have been overlooked in popular sources, and because of those omissions, such sources may not be credible.

After you have preevaluated a source, it is time to dig deeper, to review the value of that source in terms of its integrity. Assessing the integrity variables of a source requires asking: How robust is the research method? How appropriately and accurately did the researcher follow the steps in the research process? The answers to those questions will help you decipher the veracity of a source.

Determining the robustness of the research method requires, first and foremost, confirming that a method has been presented. Sometimes this important bit of information is omitted altogether, especially in sources that are not peer reviewed. Perhaps the author provided no information about the sample that was recruited, the instruments that were used, or the inferential statistics that were employed to achieve the results. If that information has been omitted, the source lacks integrity, and you should use caution in including such a source in your research.

Most often, your assessment of a source’s methods will be a little more ambiguous. An author may provide *some* information about his or her method but may never specify, for example, the instrument that was used to collect the data; alternatively, the author might list the questions in an instrument but never specify whether the instrument was tested for reliability or validity. Note these bits of detail that are provided and ask yourself whether the information is sufficient to justify the claims made in the Results section. Also note whether the approach is a logical one to address a problem and whether the information is clearly presented.

As you read a source, you will also need to decipher how appropriately and accurately the author follows the steps of the research process. All claims should be substantiated in the study, either with information from well-respected sources or from the author's own careful data collection or analysis. If you see any grandiose claims in a source (e.g., "All women desire to stay at home with their children") with no substantiating evidence or a citation for support, you should question the credibility and integrity of that source. Noting all these limitations does not mean personally attacking a source; rather, it means critiquing it in a scholarly way.

Common Mistakes

There are a variety of common mistakes in critiquing the literature. One of the most common mistakes is using emotion instead of evidence to "fuel" the critique. For example, you might attack a researcher because you dislike his or her results, not because of the way in which the research was conducted. Perhaps you are conducting a study on the utility of a particular educational intervention for struggling readers, and you find a study that shows no impact of this intervention. Maybe the source even shows that it has a harmful effect. If you personally believe this intervention is effective and are hoping to generate support for such an intervention through your study, you may be tempted to dismiss this article as "bad" and may choose not to include it in your work. Use caution before you entertain such a response. Reflect on what the article really says and focus in on the methods of that study. Who was involved in the sample? How was the intervention tested? If you examine that study's methods, you will likely understand the study's reported results and why the findings of your proposed study—*in conjunction with the findings of that contradicting study*—would provide a more robust understanding of your phenomenon of interest. For example, if you were conducting a study on a particular literacy-based intervention and your hope was to find positive results from that intervention in your own population of interest, you might be hesitant to read anything (let alone to include it in your literature review) that would argue against that intervention's utility. But if there was such evidence, look into it a bit more. Perhaps the researchers in that particular study tested the utility of this intervention on a much different sample. Perhaps they used a much smaller sample. Perhaps they only used a portion of the intervention. The point is that there is a lot to learn in the methods of any study. The important thing to remember is to not dismiss an article simply because it tells a different story than the one you are hoping to tell in your own research.

Another mistake involves dismissing complex studies as "no good" simply because they are complex. Sometimes, the findings in research studies are indeed overly complex (e.g., "Central aspects of Bronfenbrenner's theory are structurally represented in the schematic representation of the study's findings wherein an explicit set of relations between variables in the interest domain are specified"), and the findings could benefit from a better presentation (e.g., "Bronfenbrenner's theory is supported by the findings of this research"). However, dismissing a long or complex article as unusable simply because it contains vocabulary that you are unaccustomed to or statistics that you have not yet learned is not appropriate. Trust yourself. Spend time with those articles and do your best to interpret them. You will likely be able to interpret more than you would originally think at first glance of the study.

Ways to Stay Organized

It should be clear now that there are many factors to consider with any source you read. Below are some elements of a publication to consider during your reading. These elements are intended to organize your thoughts as you critique the integrity of a source, whether that source is a quantitative, qualitative, or mixed methods study.

Research Problem. What is it that the author(s) actually studied? The research problem is often presented after a short introduction, where the general topic has been introduced and the social problem has been presented. The research problem should be distinguished from the social problem. A social problem should be seen as an undesirable situation in one's community or some other specified context (e.g., adolescents who affiliate with deviant peer groups tend to engage in risky, self-harming behavior). A research problem, on the other hand, is a topic that needs to be studied (e.g., little is known about how adolescents identify with those groups). Although the social problem might be the reason why you are interested in a particular study, take note of the research problem and consider whether it has been clearly presented.

Literature Review. Is the review logically organized? Does it offer a balanced critical analysis of the literature? Does it explain the philosophical underpinnings of the study? All of those questions are important to address because the robustness of a literature review is an important indicator of the entire study's credibility. A sparse literature review that is poorly organized and that has little cohesion and flow should alert you that the foundation of the study might be questionable. You should look for a variety of authors and sources being used as well as for use of current sources. If dated sources (i.e., those that are older than 5 years old) are being used, investigate their merit. Make sure they are from seminal research studies and ones that are highly respected within the field. This consideration is important because research is continually evolving. Any source that has an abundance of excessively old citations is likely not an accurate reflection of the current state of knowledge on a topic.

Theoretical or Conceptual Framework. Has a conceptual or theoretical framework been identified? Is the framework adequately described? A theoretical or conceptual framework provides an overall orienting lens that shapes the types of questions asked, informs how data are collected and analyzed, and/or provides a call for action or change. Although some authors contest the utility of a theoretical or conceptual framework (e.g., see G. Thomas, 1997), arguing that frameworks provide unnecessary structure or constrain thought, most scholars encourage the use of a framework to help bring focus and a commonly understood set of tenets to research (Hart, 2011; also see Dr. Linda Crawford's chapter in this book on conceptual and theoretical frameworks). Take note of a framework, if there is one, and note its major tenets.

Research Question. Has a research question been identified? If so, is that research question clearly stated? Sometimes, a research question is very explicit: It is given its own space in an article and is clearly presented as a question, sometimes prefaced with the overt "Research Question" heading (e.g., "What is the relationship between adolescent peer group affiliation and the proclivity toward risk behavior?"). Other times, the research question is not a question at all but instead a declarative statement of intent (e.g., "The purpose of this study was to address the relationship between adolescent peer group affiliation and the proclivity toward risk behavior."). Research questions will also differ depending on the design of a study. Quantitative studies will examine the relationship between two or more quantifiable variables, whereas qualitative studies will examine the experience of a phenomenon with the intent to develop a complex picture of the problem or issue under study (Creswell, 2014, p. 186). For example, a quantitative study might address the relationship between peer group affiliation, defined by an adolescent's self-identified peer group among a finite number of group options, and his or her self-reported risk behavior. A qualitative study, in comparison, may study the lived experiences of adolescents who identify with a particular peer group. Whatever the study design of the resource you are reading, make note of where the research questions are presented and how they are presented. Ideally, the constructs that are being studied have been clearly defined so that you know exactly what is being examined.

Methods. Is the research design clearly identified? Has the instrument used to gather data been described? Is the instrument appropriate? The answers to those questions might look different, depending on the method used in the study. A quantitative study should describe the instruments that were used, including their reliability and validity. Ideally, the authors will provide a few example items from the instrument, too, to give you a glimpse into how the items corresponded to the constructs that were studied. The same level of detail should be provided in a qualitative study, too, but the language will likely be different. Instead of using language like *reliability* and *validity*, qualitative researchers are more likely to describe the rigor of their methods in terms of a variety of *trustworthiness* indicators (E. Thomas & Magilvy, 2011). The important thing to look for, simply, is the level of detail the authors provide. The more transparent and clear the authors are in presenting their methods, the more rigor you can assume their study to have.

You should also note the details of the sample that was used. Ask yourself: Are the sampling method and sample size identified? Is the sampling method appropriate? Were the participants suitable for informing research? The sample size will almost certainly differ depending on the study design. Quantitative studies tend to use larger, more diverse samples, and qualitative studies tend to user smaller, more homogeneous samples, but the differences are occasionally more subtle. What is important to note is the amount of detail. The more you know about the sample, which should include not just the size of the sample but also the inclusion criteria and the method of recruitment, the better equipped you will be to determine the generalizability or transferability of the results.

Results. Relatedly, look for information on the findings or results. Ask yourself the following questions: Are the results clearly connected to the hypotheses and methods of analysis? Are all the major research questions answered in the results? There should be a clear sense of parallelism in the source: Each research question or hypothesis should correspond with its own findings report, and those findings should make sense in relation to the method of analysis that was used to achieve it. Each result should be presented in the same order in which its corresponding question was presented, and no superfluous details should be included. That is, no results should be presented if they are tangential to the original research questions, and no interpretive commentary should be included—for example, about whether the findings were interesting or surprising. Any such commentary should be reserved for the very end of the study, where the authors provide some of their own reflections about the study.

Cohesion. There are other factors to consider in evaluating the credibility and integrity of a source, and those will be loosely bundled here under the overarching category of *cohesion*. The cohesion of a study speaks to how all the elements hang together. Once you have read through an entire study, take a moment to pause and reflect on its cohesion. Ask yourself: Does the research report follow the steps of the research process in a logical manner? Do these steps naturally flow, and are the links clear? The introduction should lead into the problem statement and research purpose, which should lead into the methods, results, conclusions, and interpretations. These elements should flow together, much like a good story. The writing should be clear and accessible and should abide by the guidelines provided in the writing style manual adopted by your discipline or program.

In critiquing the aspects of a research study, these questions should foster more than just a simple *yes* or *no* response. They should stimulate you to consider the implications of what the researcher has done, and how you might build upon their research and/or improve it. Explore different ways to keep your ideas organized as you read. You may choose to highlight or underline these various aspects in every article that you read. You might take handwritten notes in the margins. Alternatively, you might create an entirely separate document to catalogue these important details. I recommend a literature review matrix (Table 9.1), something that will structure your note taking so that you are able to note the details within, and across, the various articles you are reading. Doing so will allow you to move from critiquing the literature to synthesizing it. I turn to synthesis next.

Synthesizing the Literature

It should be noted that these critiquing skills should be used for every article that you read. If that thought seems overwhelming, do not fret: You will get much faster at the process as you read more sources, and you will start to

TABLE 9.1
Example of a Literature Review Matrix on Adolescent Drug Abuse and Cognitive-Behavioral Therapy

Author/Date	Overview	Research Questions/ Hypotheses	Methodology	Analysis	Results
Williams and Ricciardelli (1999)	Investigated restrained drinking and self-control in relation to alcohol consumption and problem drinking in a sample of adolescents.	Baumeister and Heatherton's (1996) model of self-regulation failure: Lack of cognitive control will be associated with high levels of alcohol consumption in adolescents.	N = 198 adolescents. Completed questionnaires that assessed levels of alcohol consumption, problem drinking, restrained drinking, and cognitive self-control	Correlation analyses were employed to examine the association between self-control and alcohol consumption.	Low levels of cognitive self-control predicted high levels of problem drinking.
Myers and Brown (1990)	Examined coping strategies of 50 adolescents treated for drug and alcohol abuse.	Relapsers will generate fewer coping strategies than abstainers in hypothetical high-risk situations.	N = 50 adolescents. Structured interviews were given to both adolescents and their parents at pre- and posttest.	Multivariate analysis of variance (MANOVA) was employed to determine group differences in coping skills.	Low problem-solving coping strategies were associated with the poorest drug use outcomes.

TABLE 9.1 *(continued)*

Example of a Literature Review Matrix on Adolescent Drug Abuse and Cognitive-Behavioral Therapy

Author/Date	Overview	Research Questions/ Hypotheses	Methodology	Analysis	Results
Wagner, Myers, and McIninch (1999)	Examined stress coping and temptation coping as potential contributors to substance abuse.	Stress coping and temptation coping will contribute to the most variance in adolescent substance abuse.	N = 332 adolescents. Completed questionnaires over 2-day period.	Regressions were employed to examine the degree to which stress and temptation coping predict adolescent substance abuse.	High stress was the strongest predictor of adolescent substance abuse.
Botvin, Baker, Dusenbury, Tortu, and Botvin (1990)	Examined the effectiveness of cognitive-behavioral skills training program on adolescents in New York.	Adolescents who receive cognitive-behavioral skills training will report lower cigarette, alcohol, and marijuana use than those who do not receive the training.	N = 4,466 adolescents. Students were randomly assigned to one of three groups. Groups were given a pre- and posttest.	MANOVAs were employed to examine group differences in substance use and cognitive skills.	Those who received cognitive-behavioral treatment reported lowest levels of cigarette smoking, marijuana use, and alcohol use.
Longabaugh and Morgenstern (1999)	Examined current status of cognitive-behavioral therapy (CBT) for substance abuse treatment.	What determines the effectiveness of cognitive-behavioral skills therapy? Are certain CBT approaches superior to others?	N = 10 studies. Studies were identified that examined the hypothesized mechanisms of action of CBT.	Qualitative content analysis	CBT has been associated with a decrease in drug abuse but *no conclusions have been made for the underlying mechanisms of CBT*.

do it implicitly, so you will not need to write everything down. Note taking is always encouraged, but it can be a fast and informal process, jotting down notes in margins or on a piece of scratch paper or whatever is necessary to capture the major ideas of a source. The important lesson to remember is that you will get more adept at critiquing each article as you gain experience in reading. It is also encouraging to remember that the more you read, and the more you hone your ability to critique the literature, the better you will write. Better readers make better writers.

As you develop your skills in critiquing the literature—essentially, in analyzing or breaking down each source into its constituent parts—you will become better equipped to synthesize what you read into a coherent whole. The second half of this chapter provides some guidance on engaging in that process.

Definitions

The term *synthesis* is often misused to describe an annotated bibliography, one that includes various sources presented alongside one another. Although an annotated bibliography is certainly a helpful start, it is not, in itself, a synthesized account of a particular topic. Instead, synthesis is defined as "an integrated analysis of different subtopics that help the writer and reader to come to a greater understanding of the state of knowledge on a larger issue" (Hart, 2011, p. 110). The unified, singular entity is the important feature in a synthesized narrative because the source provides *a level of analysis* to relate the ideas of different sources together.

Important Considerations

A synthesized paper is an integrated, critical essay on the most relevant and current published knowledge on the topic. Hence, not everything needs to be included in a synthesized paper. The type of instrument used in every study you have read, for example, may not be necessary to report in your synthesis, but a general overview of methods for your topic would be. Similarly, an individual account of sample sizes for each study you read is probably not a helpful detail but a summary of the samples sizes (e.g., large or small, homogeneous or diverse) would be informative. It will be up to you to decide on the important details to include or omit, but the key thing to remember is that the information that is shared should be relevant to the larger story you are telling.

Beyond considering the level of detail in a synthesized paper, you must consider the ways in which those details are organized. Typically, the ideas of a synthesized paper are organized around major ideas or themes, including a discussion of theories, models, background, and trends. Once you identify those major ideas or themes, put them into an outline. Your outline can (and likely will) change as you continue to gather more sources, but the simple act of outlining major ideas will help you to stay focused on the bigger, overarching story of your study, rather than on individual studies. Your initial synthesized draft might look very simple, something like this:

I. Main Idea A
   a. Supportive evidence (Author name, year)
   b. Supportive evidence (Author name, year)

        c. Supportive evidence (Author name, year)
            i. Analysis that compares and contrasts the above evidence and relates it all to Main Idea A
            ii. Transitional sentence to guide the reader into Main Idea B

II. Main Idea B
    a. Supportive evidence (Author name, year)
    b. Supportive evidence (Author name, year)
    c. Supportive evidence (Author name, year)
        i. Analysis that compares and contrasts the above evidence and relates it all to Main Idea B
        ii. Transitional sentence to guide the reader into Main Idea C

III. . . . and so on

Each main idea might be related to a theme or a topic on which you would like to expand from the literature; alternatively, each main idea might be on a different theory or a particular methodology that has been used in the literature to explain or explore your topic. However you choose to organize your outline, make sure that the overarching ideas—the ones that you specify through the biggest roman numerals—are *ideas*, not individual articles or studies.

Some writers who have not done a lot of synthesis struggle to identify what synthesis looks like and wonder what sort of language typifies a synthesized paper. To help you get started, there are some phrases that alert the reader to a synthesized paper. Phrases such as, "In line with these findings . . . , " "The results of X are consistent with Y . . . , " and "Although X suggests . . . , Y reaches a different conclusion" are all illustrative of synthesis, where two or more ideas are being analyzed alongside each other. In addition, the American Psychological Association's (2010) *Publication Manual* provides some excellent guidance on the use of transitional phrases to demonstrate a synthesized analysis of two or more sources. For example, some phrases are helpful to demonstrate time links (*then, next, after, while, since*); some are used to demonstrate cause–effect links (*therefore, consequently, as a result*); some are used to demonstrate additions (*in addition, moreover, furthermore, similarly*); and others are used to illustrate contrasts (*but, conversely, nevertheless, however, although*) (p. 65). Consider those phrases as you begin to consider how your sources are related and feel free to use them if they frame the relationship between your sources appropriately.

It is important to note, however, that a synthesized paper is more than one that includes those synthesizing phrases. A synthesized paper should leave a reader with a holistic sense that the writer has conveyed his or her own new ideas and has drawn on a chorus of support. Synthesis gives you the opportunity to make your voice heard.

Common Mistakes

Synthesis is a developed skill, and it may take some time before you feel proficient at it. Be patient with yourself and continue practicing. While you do, keep in mind these common mistakes so that your work represents a true synthesis.

A common mistake in the attempt to synthesize is to include two or more sources in a paragraph. Here are some lighthearted examples:

> Dark chocolate has been found to affect mood (Hershey, 2014); namely, the more dark chocolate one eats, the more cheerful one tends to be. White chocolate has little effect on mood (Ghirardelli, 2013).

Synthesis does not mean simply having two or more sources. Having more than one source is an excellent first step, but real synthesis is not occurring unless a relationship between the sources is apparent. In the examples above, the two sentences are broadly related (i.e., they both discuss chocolate), but there is no analysis given about how, exactly, they are connected. Something like this would be necessary to truly analyze the two sources together:

> The *type of chocolate* may be an important consideration in the study of the chocolate–mood connection.

Another common mistake is to "throw in" random synthesis words for good measure but to stop there, without providing any level of description or analysis about how the various ideas are related. Here is an example:

> One reason that chocolate should be on every elementary school lunch menu is that it makes students feel good. Rita Dove, president of Dove Chocolate, agrees, stating, "Chocolate is the perfect end to a satisfying luncheon" (Hershey, 2012, p. 18).

Did Rita Dove actually *agree* to the assertion that chocolate should be on every elementary school menu? She probably did not. The important lesson to learn here is to use words deliberately. A word like "agrees" should be reserved for occasions when a real relationship has been explored. A more appropriate synthesis of those sources would be something like this:

> One reason that chocolate should be on every elementary school lunch menu is that it makes students feel good. Rita Dove, president of Dove Chocolate, provided a statement recently about the pleasant aftereffects of chocolate, saying that "Chocolate is the perfect end to a satisfying luncheon." Such a statement would likely support the argument of chocolate's inclusion in the lunch menu.

Yet another common mistake in synthesized drafts is the overuse of quotes. Below is an example:

> "Dark chocolate contains antioxidants" (Ghirardelli, 2013, p. 3); "antioxidants protect cells from free radical damage" (Oz, 2015, p. 27). "Dark chocolate can be incorporated into many snacks or meals" (Crocker, 2014, p. 22).

Always ask yourself this question: Are you interpreting quotes and furthering your own ideas, or are you just piling quotes one on top of the other? A more appropriate synthesis of those quotes would be something that takes the *ideas* of those quotes and translates them into a more coherent whole:

> Dark chocolate is seen as a wise accompaniment to one's diet. This perspective is shared by those who have touted the antioxidants in dark chocolate

> (Ghirardelli, 2013), the ability of dark chocolate to protect cells from free radical damage (Oz, 2014, p. 27), and the ease with which one can incorporate dark chocolate into regular snacks or meals.

Using one source too often is an additional mistake that writers often make. Below is an example:

> Hershey (2015) has found that chocolate can boost one's mood . . .
>
> Hershey (2015) has discovered a relationship between happiness and consumption of chocolate . . .
>
> Hershey's (2015) study revealed that women eat significantly more chocolate than men . . .

You would be well advised to make sure that one source is not driving the study. If every paragraph contains material from one particular source on the reference page or if whole paragraphs are not synthesized but rather offer this single, particular source alone, a reader could become suspicious that you have copied another author's structure and are relying too heavily on that source's argumentative structure, rather than coming up with your own. Be on the alert for this problem.

An additional mistake is to start or end a paragraph with a direct quote. Below is an example:

> "Potato chips are tasty treats but leave your fingers feeling greasy" (Lays, 2015, p. 267). Potato chips are a favorite snack food among Americans (Pringles, 2014, p. 23), but researchers have documented how they tend to leave your fingers, arteries, and intestinal tract with a Vaseline-like substance (Ruffles, 2015). "There is a big problem for people who eat potato chips while typing on a laptop" (Utz, 2013, p. 19).
>
> "Another snack food with potential problems is the corn chip" (Jones, 2013, p. 45).

The first sentence of a paragraph should be a topic sentence written in your own words. Paragraphs that begin with a quote leave the reader confused as to where the paragraph is going. Similarly, paragraphs should end in your own words and not with a quote. The last sentence of a paragraph should recap the main point of the paragraph and transition into the ideas of the next paragraph.

Ways to Stay Organized

Synthesizing is difficult, and it can be especially difficult to organize the information in a way that makes the writing process simpler. One tool that may be helpful in organizing literature reviews is the literature review matrix (see Table 9.1), an organizational tool that was also recommended in the Critiquing the Literature section of this chapter. The literature review matrix is a chart that allows you to sort and categorize the different arguments presented on an issue. Across the top of the chart are the spaces to record different aspects of each study, and along the side of the chart are the spaces to record each separate article. As you examine your first source, you will work horizontally across the first row, recording as much information as possible about each significant aspect of the study. Follow a similar pattern for your following sources. As you find information that relates to an article you have already recorded, group those articles together in your matrix. If you are adept at Microsoft Excel, you can even sort your articles on keywords that allow you to group similar ideas. As you gather more sources, you will start to be able to identify different subheadings or thoughts in the literature. You may organize your paper by these subheadings. Table 9.1 illustrates a literature review matrix that has been populated with a variety of related articles about adolescent drug abuse and cognitive-behavioral therapy, followed by an example of synthesis, based on the information within the matrix.

Below is a synthesis of articles on adolescent drug abuse and cognitive-behavioral therapy, using the articles from the literature review matrix in Table 9.1 to pull together a synthesized narrative. Before looking at the example, briefly review Table 9.1 to familiarize yourself with the articles to be synthesized.

Synthesis Example

> Drug abuse may likely be a function of one's cognitive style. This idea has been supported in a variety of studies on cognition and drug abuse in adolescence. In comparison to adolescents who do not abuse drugs, past research has found drug-abusing adolescents to report lower levels of cognitive self-control (Williams & Ricciardelli, 1999), fewer problem-solving coping strategies (Myers & Brown, 1990), and higher levels of stress (Wagner, Myers, & McInich, 1999).
>
> Because of these empirical findings, cognitive-behavioral therapy (CBT) has grown in popularity as an effective approach to reduce or prevent adolescent drug use. The effectiveness of CBT to reduce drug use, such as cigarette smoking, marijuana use, and alcohol use, has been demonstrated in a variety of clinical trials (e.g., Botvin, Baker, Dusenbury, Tortu, & Botvin, 1990). Despite the promising findings from these studies, a recent content analysis on CBT studies found no explanation of the underlying mechanisms linking CBT to behavioral change (Longabaugh & Morgenstern, 1999).

If a literature review matrix is not an effective approach for you, an alternative tip is to use some sort of data management software program, such as Zotero or NVivo, in which you are able to create various folders, each one containing summaries and hard copies of the articles you are reading. Alternatively, you can code by hand the sources you are collecting via highlighters, Post-it Notes, note cards, or any other hard copy method to lay out your articles in front of you. Create piles, sort by theme, take notes, and highlight what seems important. The important lesson is to do what works for you.

Conclusion

In this chapter, I brought together various considerations on critiquing and synthesizing the literature, including the definitions of each skill, the important considerations of each skill, the common mistakes for each skill, and the ways to stay organized. The multifaceted nature of critiquing and synthesizing the literature should underscore two key points. First, these skills are hugely important. They are necessary not only to become a better reader but also to become a better writer. Second, honing these skills takes time. As you work on developing the ability to critique and synthesize, be patient with yourself. Being able to write an elegant, synthesized paper is not something that novice scholars are able to do at their first attempt. It takes practice, and just like any other practiced skill in life—such as playing the piano, throwing a curve ball, or cooking a soufflé—the skills to critique and synthesize require regular practice. Set aside some time every day for this practice. Read often and write often. The more you practice your skills in critiquing and synthesizing the literature, the more proficient you will become at doing so, and you will reap the rewards. With these skills, your work will be welcomed into the scholarly community because you are able to engage in a sophisticated discussion with your peers.

KEY SOURCES

American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.

Walsh, M. L., Pezalla, A. E., & Marshall, H. (2014). *Essential guide to critical reading and writing*. Baltimore, MD: Laureate International Universities.

REFERENCES

American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.

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

Hart, C. (2011). *Doing a literature review: Releasing the social science research imagination*. London, United Kingdom: Sage.

Thomas, E., & Magilvy, J. K. (2011). Qualitative rigor or research validity in qualitative research. *Journal for Specialists in Pediatric Nursing*, *16*(20), 151–155. doi:10.1111/j.1744-6155.2011.00283.x

Thomas, G. (1997). What's the use of theory? *Harvard Educational Review*, *67*(1), 75–104.

Walsh, M. L., Pezalla, A. E., & Marshall, H. (2014). *Essential guide to critical reading and writing*. Baltimore, MD: Laureate International Universities.

10

Writing the Research Proposal

Linda M. Crawford, PhD, Gary J. Burkholder, PhD, and
Kimberley A. Cox, PhD

Introduction

A research proposal is a thorough plan for conducting a study. Specifically, a dissertation or thesis proposal usually contains chapters, or some other organizational structure and presentation order, that include the rationale for the study and a description of how the study will be carried out. In this chapter, we will discuss the elements of a research proposal, the development of a plan for conducting the study, and alignment among the research design, research questions, and methods. This chapter is oriented to those writing thesis or dissertation proposals; however, the principles easily apply to many kinds of research proposals. We saved this chapter for the end of Part 1 of this book because it draws on information presented in each of the preceding chapters. Our intent is that this chapter serves as the practical organizing principle for the research proposal.

Constructing the proposal is an iterative process that does not proceed linearly. When writing the proposal, you will often go back and forth among the elements, filling in some parts and moving to other parts until you are assured that all the parts are fully aligned. In that light, although we discuss elements of a research proposal in this chapter in what we believe to be a logical order, we do not expect that you will write a proposal sequentially, nor do we imply that every institution will require every element in the order discussed. Understanding each of these proposal elements, though, will help you to conceptualize your proposal and prepare it in the manner required by your institution.

Writing the Dissertation/Thesis Research Proposal

A research proposal generally contains two main elements: a study rationale and a study implementation plan (Figure 10.1). The rationale for a study is rooted in an exhaustive literature review that builds context for and establishes the need for the study. The rationale is further specified in the problem, purpose, research questions, and conceptual framework. The implementation plan identifies the overall framework, or design, of the study and the methods to be used for selecting participants, collecting data, and analyzing data. The implementation plan also delineates the study in terms of definitions, assumptions, scope and delimitations, limitations, and significance.

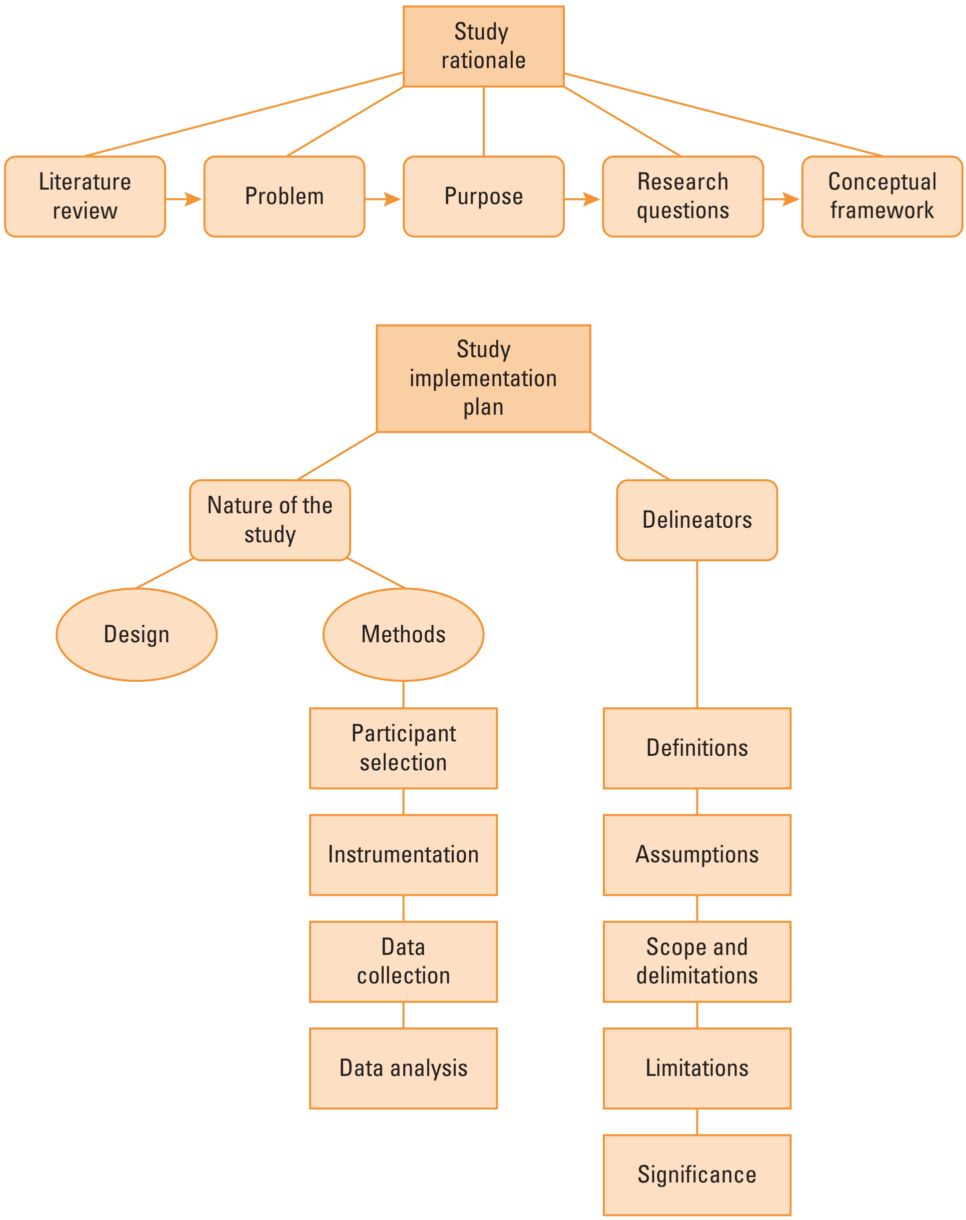


FIGURE 10.1 Elements of a research proposal.

Study Rationale

The rationale begins with the literature review, which establishes context and the need for the study by presenting what is already known and what yet needs to be known or understood about a topic. The literature review may be summarized in a background and gap statement, which leads to a problem that the study is designed to solve. The rationale is further specified in a statement of purpose for the study, research questions and/or hypotheses, and the conceptual framework of the study.

Literature Review

When approaching a study, you begin with some wonderment in your mind concerning a question about some issue or topic of interest. For example, you may be wondering how decisions are made in collegial groups, what motivates employee loyalty to organizations, or why some leaders generate more productivity among workers than other leaders. As the researcher, you then seek to determine if this point of interest is of concern, not only to you but also to the profession at large, and you also seek to determine what is already known and not known about the topic. To ascertain what is known and not yet known or understood about the topic, you will need to delve into the related literature.

The literature review is a systematic synthesis of prior research and knowledge on your topic. There are a number of purposes that are served by the literature review, and it is good for you to be aware of these as you begin the process of reading and synthesizing. These purposes include (a) providing evidence for what researchers know and do not know in a particular discipline or topical area and how addressing what researchers do not know will advance discipline-specific knowledge; (b) learning the types of approaches to inquiry and research designs that have been used in the discipline to address the phenomenon of interest; (c) understanding the seminal thinkers in the discipline and how that thinking shapes your own understanding of the nature of knowledge; and (d) gathering evidence to support your choice of methods. Thus, a literature review is much more than just a summary of the research; rather, it represents a critical understanding of the relevant scientific discourse in your discipline.

Note also that when we use the word *synthesis*, we mean something different from summarization. For example, a summary of an article reiterates the contents of an article, but in a briefer form. Synthesis requires you to find similarities and differences among articles that reveal understanding. For example, you might summarize the point of one article but find that three other articles reiterate that idea and one article contradicts the idea. A literature review combines, or synthesizes, these findings to create context for the study and to identify what is not yet known or understood about the topic that is of concern to the profession.

Literature reviews may be organized in different ways, but typically, a literature review presents the seminal content and theoretical literature on the topic; the literature related to the independent variables, dependent variables, or phenomenon under study; studies conducted on the topic with various populations; and qualitative, quantitative, and mixed methods designs that have been used previously to study the topic. Based on the literature review, you determine what is lacking in knowledge or understanding of your topic that people in your profession are wondering about. This determination leads to the problem statement.

Problem Statement

A problem statement is a description of what is not known or understood about your topic that is of interest to the profession (Booth, Colomb, & Williams, 2008). Although you might begin thinking about the topic based on personal experience, personal experience alone is insufficient to support a research study. The problem that you pursue in your study may begin with your personal experience but must be reflected in the literature. The problem statement is one of the three most important elements of a research proposal, along with the purpose statement and research questions/hypotheses. In writing the problem statement, you should use precise and parsimonious language that clearly signals the problem to your reader. Although some institutions might require additional elements in a problem statement, there are three essential parts of a problem statement. These three elements are

1. The research problem itself
2. For whom and why studying the research problem is important
3. How the current study will address the research problem

It is best to begin the problem statement with a clear declarative sentence, such as "The problem to be addressed by this study is that [something is not understood or not known]." If the sentence reads "The problem to be addressed by this study is to . . . , " then the statement is probably a purpose rather than a problem statement because it is indicating what your study will do rather than what problem it will address. Be sure to indicate what is not understood or not known rather than stating what your study will do.

Also, recognize the difference between a *situation*, which often comes from personal experience, and a *problem*, which is a literature-based concern across the profession. For example, "The problem to be addressed by this study is that principals do not use fair means of evaluating teachers." This is a statement of a situation and is likely rooted in the researcher's personal experience. The statement does not state what is not understood or not known based on the research literature. Compare that statement with the following: "The problem to be addressed by this study is that the means of establishing equity in teacher evaluations conducted by principals are not known." This second statement reflects that the literature has not yet provided sufficient research related to equity in teacher evaluation.

The problem statement indicates for whom and why studying the problem is important. This element is a precursor to stating the study significance, described later in this chapter. In the problem statement, you need to indicate why the profession considers an investigation of the topic important and also how addressing the problem may affect the lives of stakeholders to the problem. For example, for a study on voting patterns in rural areas, this part of the problem statement might state, "Understanding voting patterns in rural areas might reveal issues of importance to politicians and the general public that impact accessibility to voting sites."

The problem statement also states how the current study will address the research problem in general terms. For example, for a study on minority parents' satisfaction with screening procedures for acceptance into a talented and gifted program for their children, this part of the problem statement might state, "This study will describe minority parents' perceptions of the procedures for admittance into talented and gifted programs."

The problem statement can be written in one concise paragraph. Below is a sample problem statement from a completed dissertation. The dissertation topic was the use of one social skills curriculum strategy, video self-modeling (VSM), with twice-exceptional students. *Twice exceptional* refers to those students recognized as intellectually gifted who also have a special needs diagnosis (Allard, 2014).

> The problem addressed by this study is the underuse of VSM as part of a comprehensive social skills curriculum for twice-exceptional students. Understanding the reasons for underuse of this potentially viable intervention for twice-exceptional students is important for school personnel invested in ensuring the social skills success for this population. This study addressed the problem of underuse of VSM by gaining an understanding of the experiences of school personnel in relation to the use of VSM as an intervention to improve social skills in twice-exceptional students, including the incentives and disincentives for its use in the school setting. Themes generated from interviews of the team members contribute to a more informed understanding of VSM as an intervention with this student population, as well as better informed practice by understanding the lived experiences of staff who are actually carrying out the VSM intervention. (Allard, 2014, pp. 6–7)

Notice that the problem statement begins with a clear identification of what is of issue—the underuse of a model for a certain population. Next, Allard (2014) explains why understanding the underuse of this intervention is important and for whom it is important. Finally, the reader learns that the study will focus on gaining insight based on the experiences of school personnel who have used the intervention with the defined student population.

Purpose Statement

The final section of a problem statement, how the current study will address the problem, is the purpose statement. The purpose statement describes the study (a) design; (b) theory being tested (quantitative), central phenomenon investigated (qualitative), or both (mixed methods); (c) intent; (d) variables (quantitative), definition of phenomenon (qualitative), or both (mixed methods); (e) participants; and (f) site or context. That seems like a lot of information, but it can actually be presented in one or two concise sentences.

We provide here a format for writing a purpose statement that includes the items listed above; the format is an adaptation of a purpose statement script provided by Creswell (2009, pp. 114, 118). Although your purpose statement does not need to follow this somewhat formulaic presentation, your purpose statement does need to include all the items identified earlier. Examples of purpose statements for qualitative and quantitative studies are provided. Mixed methods studies would combine the two statements shown below.

> *Qualitative purpose statement*. The purpose of this ____ (design: case study, ethnography, etc.) study is to ____ (understand, describe, explore, develop, etc.) the _____ (central phenomenon) for _____ (participants) at _____ (site). The central phenomenon is generally defined as _____ (general definition).

> *Quantitative purpose statement*. The purpose of this ____ (design: experimental, quasi-experimental, descriptive) study is to ____ (test or describe) the theory of _____ that ____ (describes, compares, or relates) the _____ (independent variable) to _____ (dependent variable), controlling for _____ (name control variables, if appropriate) for _____ (participants) at _____ (site). The independent variable(s) are generally defined as _____ (general definition). The dependent variable(s) are generally defined as _____ (general definition). The control and intervening variable(s), ____ (identify, if appropriate), will be statistically controlled in the study.

As an example, following from the Allard (2014) problem statement cited earlier, her purpose statement was

> The purpose of this qualitative, phenomenological study was to explore the lived experiences of special education team members who have incorporated VSM as a social skills intervention for twice-exceptional elementary students. I examined the nature of these experiences in an attempt to understand the underuse of VSM and to explore incentives and disincentives for the use of VSM in the context of an elementary school setting with this population. (p. 7)

Notice how Allard (2014) identified the design as a qualitative, phenomenological study. The central phenomenon is incorporation of the VSM intervention, which is defined as a social skills intervention. She used an infinitive verb form, *to explore*, which signals study intent. Participants are specified as special education team members who have incorporated VSM for twice-exceptional students. The site is an elementary school setting.

In this example, all six elements of a strong purpose statement are clearly identifiable, and extraneous information is absent. We recommend that, when writing your purpose statement, you notate your content beside each of the six items listed earlier prior to formulating a purpose statement paragraph. Then, double check to see if your paragraph incorporates all six items. Furthermore, we recommend that your purpose statement begin with a clear declarative sentence, such as "The purpose of this [design] study is [to verb] . . . ." Writing your sentence in this way offers the reader a precise understanding of your study purpose.

In summary, we have presented recommendations for the flow of development of your research proposal. You should begin by writing your literature review, which is a synthesis of prior research and knowledge on your topic. You summarize the literature review in the background section and end that section with a gap statement that identifies what, of relevance to your study, is not yet known or understood about your topic. The gap statement serves as a precursor to the problem. The problem contains three elements: the problem itself, for whom and why studying the problem is important, and how the study will address the problem. The last part of the problem statement, how the study will address the problem, is a precursor to the purpose statement. The purpose statement specifies study design, theoretical context, intent, variables or phenomenon, participants, and site. The next step is the development of research questions.

Research Questions and Hypotheses

The purpose of a study shapes the development of research questions and/or hypotheses. An understanding of the body of knowledge on your topic, including what is known and not known, also shapes research questions and hypotheses. Research questions and hypotheses should be a direct extension of the research problem that you intend to study and the purpose of your research. The decisions you will subsequently need to make about your research design and method largely rest on understanding the type of data needed to answer your study's research questions and/or hypotheses.

Research Questions. Research questions are interrogative statements that serve three roles (Onwuegbuzie & Leech, 2004, p. 478). First, research questions frame the focus of the study. Second, research questions set the boundaries of the study; this describes the scope of the study. Third, the research questions point you toward the data that will need to be collected to answer them (Onwuegbuzie & Leech, 2004). In the process of developing research questions, your ideas about the concepts under study become more clarified as you consider how to operationally define them and link research questions with a particular approach to inquiry and research design (White, 2013).

There are several possible types of research questions, with the most apparent distinction being by the type of approach to inquiry—qualitative, quantitative, or mixed methods. Within each type of design, various types of research questions are also possible, such as those that speak to the relationship between variables or those that compare groups. Your choice of wording is important in this regard in that the words you use to write research questions should be aligned with the approach to inquiry and research design that will be used. For example, a quantitative research question that uses the word *relate* or *relationship* typically signals the use of a correlational design; an example is the question, "What is the relationship between affiliation and loyalty among employees in mid-level managerial positions in Northeastern nonprofit organizations?" This approach, where the research question informs design decisions, is referred to as a *question-led* strategy (White, 2013). This is the preferred strategy we typically recommend for research projects such as a thesis or dissertation. A *methods-led* strategy, on the other hand, is characterized by initially focusing on a particular method from which the research questions are then developed (White, 2013). According to White (2013), however, "an effective research design can only be developed after a clear set of research questions have been formulated" (p. 222), as reflected in the question-led strategy.

A common mistake in developing research questions is to write questions too broadly without mention of the context (e.g., the population or setting) for which the study is situated (Agee, 2009); for example, "What is the relationship between affiliation and loyalty among employees?" This research question fails to indicate the type of employees being studied or the setting, leaving readers unclear about who specifically will be studied and where. A better question might be, "What is the relationship between affiliation and loyalty among contract (nonpermanent) employees working in a technology innovation company?"

Another common mistake is to confuse *interview* or *survey* questions with *research* questions. Research questions guide your overall study. Interview or survey questions are those questions you ask the participants in order to obtain the data to answer the research question. Following from the earlier example, the research question is, *What is the relationship between affiliation and loyalty among contract (nonpermanent) employees working in a technology innovation company*? An interview or survey question might be *How would you describe your employer's expression of loyalty to you as an employee*? As a way to check your questions, notice whether you use the word *you* in the research questions. If you use the word *you*, then you are likely writing an interview or survey question rather than a research question.

In addition to providing context, another characteristic of a good research question is its alignment within the context of the problem and purpose of the study. As one example, let's return to the earlier hypothetical study and follow the development from research problem to purpose and finally to research questions to demonstrate the importance of this alignment.

Problem Statement: "The problem to be addressed by this study is that the relationship between affiliation and loyalty among contract workers is not fully understood. Understanding that relationship may influence the opportunities employers provide for contract employees to build a sense of affiliation in the workplace. This study will address the problem by correlating contract employee sense of affiliation with their self-reported loyalty to the company."

Purpose Statement: "The purpose of this correlational study is to describe the theory of *X* by determining the relationship between affiliation and loyalty among contract (nonpermanent) employees working in a technology innovation company. Affiliation is defined as *Y*, and loyalty is defined as *Z*."

Research Question: "What is the relationship between affiliation and loyalty among contract (nonpermanent) employees working in a technology innovation company?"

In reviewing the example, notice the similarity in wording that helps provide internal coherence of each of the key components of the study—problem, purpose, and research question—and the consistency of ideas threading through each of them.

Qualitative Research Questions. Agee's (2009) analogy of a research question as a camera lens that brings focus to a study is a helpful way to think about writing questions in qualitative research:

> A question can be thought of as a tool that is much like a steady-cam lens used to document an event or a journey. In the initial stages of study design, the researcher uses the steady-cam to frame an ever-changing broad landscape and then narrows the focus to frame and follow a specific set of events or actions in the broader terrain. However, that terrain is not just any place; it is a specific place with a dense, rich history. (pp. 441–442)

Qualitative research questions are typically characterized by their intent to describe, discover, or explore an experience or process (Onwuegbuzie & Leech, 2004). For example, Creswell (2014) recommended using the word *what* or *how* to start a qualitative research question and verbs that reflect the exploratory nature of qualitative research designs, such as *discover* and *explore*. We recommend using an approach that starts with a broad question, called a *central question*, followed by subquestions that further narrow the scope of the study (Creswell, 2014, p. 140). For example, using the general study involving the relationship between affiliation and loyalty in contract employees, the broad question might be "What are the lived experiences of contract employees working in a technology innovation company in regard to affiliation and loyalty?" Some subquestions might include (a) "How do contract employees experience their affiliation to their company?"; (b) "How do contract employees experience their affiliation to other employees?"; and (c) "What does loyalty mean to contract workers who may not have the company as his or her primary source of income?" We recommend three to five, and no more than seven, subquestions.

Qualitative research questions may evolve as data are analyzed during the conduct of the study. In addition, analysis of transcripts of interviews or other artifacts may lead the researcher to generate new questions that allow for continued exploration of a particular phenomenon. Thus, research questions may change somewhat to reflect new understandings that emerge during analysis. Revised research questions, though, need to remain within the bounds of the study, as stated in the problem and purpose statements and approved by the institutional review board (IRB). This tendency for qualitative research questions to evolve as the study is carried out is a characteristic that distinguishes qualitative research questions from quantitative research questions, which are typically developed before the study is conducted (Agee, 2009).

Quantitative Research Questions. Quantitative research questions are commonly characterized by their intent to make comparisons or examine relationships between variables. Creswell (2014) identified three common categories of quantitative research questions. First, groups may be compared on an independent variable to examine the effect on a dependent variable. Second, one or more predictor variables may be related to one or more criterion variables. Third, responses to variables may be described, such as frequency. For quantitative research questions, the use of words such as *relate*, *influence*, and *affect* are appropriate (Creswell, 2014).

Onwuegbuzie and Leech (2004) provided a typology of quantitative research questions that is similar to Creswell's (2014) categories. Their typology includes three categories of research questions: descriptive, comparative, and relationship. Descriptive questions are those that quantify a response on a variable(s), and they typically start with words such as *what is*. For example, a question might be "What is the retention rate among contract employees in a technology innovation company?" Comparative questions are those that compare groups, and they typically use words such as *compare* and *differ*. For example, a question might be "What is the difference between contract and full-time technology innovation company employees in regard to affiliation with and loyalty to their company?" Relationship questions are those that speak to the relationship between or among variables, and they typically use words such as *association*, *relationship*, or *relate*. For example, a question of this type might be "What is the relationship between affiliation and loyalty among contract employees working in a technology innovation company?" We agree with Onwuegbuzie and Leech that researchers should not begin quantitative research questions with words that will yield only a yes or no answer. Examples include "Is there a relationship between *X* and *Y*?" or "Does variable *X* predict *Y*?" This advice can also apply to writing qualitative and mixed methods research questions. Questions should provoke more complete understanding.

Mixed Methods Research Questions. There are different approaches to the development of mixed methods research questions. For example, you might write one or more quantitative research questions and one or more qualitative research questions (Onwuegbuzie & Leech, 2004). Creswell (2014) advocated for the use of three distinct types of questions in mixed methods research—qualitative, quantitative, and mixed methods questions—with the mixed methods question explicitly speaking to the mix of the qualitative and quantitative approaches of the study. Tashakkori and Creswell (2007) referred to this type of question as an integrated or hybrid research question. Based on a literature review of common practices that researchers used to write mixed methods research questions, Tashakkori and Creswell offered guidance on possible writing approaches:

> (a) Write separate quantitative and qualitative questions, followed by an explicit mixed methods question (or, more specifically, questions about the nature of integration); (b) Write an overarching mixed (hybrid, integrated) research question, later broken down into separate quantitative and qualitative subquestions to answer in each strand or phase of the study; or (c) Write research questions for each phase of a study as the study evolves. If the first phase is a quantitative phase, the question would be framed as a quantitative question or hypothesis. If the second phase is qualitative, the question for that phase would be framed as a qualitative research question. (p. 208)

For example, a mixed methods research question to complement the qualitative and quantitative examples provided earlier might be "How do the findings of the qualitative phase of the study complement the understanding of the nature of the relationship between affiliation and loyalty among contract employees in a technology innovation company?" Remember that the mixed methods questions should provide a direction for integration of approaches to understanding the phenomenon.

Hypotheses. A hypothesis is written as a statement, as opposed to a question, to reflect the proposed relationship between variables and the study outcome before data are collected. Two types of hypotheses are commonly used: the null hypothesis and the alternative (sometimes called *research*) hypothesis. The null hypothesis states that there is no significant difference (or relationship) between two or more variables, and it is typically designated H_0. For example, in a quantitative research study that intends to compare groups, the null hypotheses would speak to no difference on the variable(s) under study between the groups. The alternative hypothesis states that there is a significant difference or relationship among two or more variables, and it is typically designated H_1.

Hypotheses can be written as either nondirectional or directional. A nondirectional hypothesis states that a difference exists but does not indicate the nature of that difference. A nondirectional null hypothesis example might be "There is no significant correlation between employee affiliation, as measured by the *X* scale, and employee loyalty, as measured by the *Y* scale, among contract employees at Company *Z*." Using another example that reflects a group difference test, a nondirectional null hypothesis could be "There is no significant difference between contract employees and permanent employees at Company *Z* on loyalty, as measured by the *Y* scale." In a directional hypothesis, the nature of the predicted difference or relationship is explicitly stated. This difference could be stated in terms of more than or less than. Directional hypotheses are typically used when you have some knowledge of how the variables are related based on prior research. An example directional alternative hypothesis might be "There is a significant negative correlation between employee affiliation, as measured by the *X* scale, and employee loyalty, as measured by the *Y* scale, among contract employees at Company *Z*." In this example, the appropriate null hypothesis would be "There is no correlation or significant positive correlation between employee affiliation, as measured by the *X* scale, and employee loyalty, as measured by the *Y* scale, among contract employees at Company *Z*." An example of an alternative hypothesis for a group difference test might be "Permanent employees will report higher levels of employee loyalty than contract employees as measured by the *Y* scale among employees of Company *Z*." The corresponding null hypothesis would be "Permanent employees will report lower levels of employee loyalty, or no statistical difference in scores, as measured by the *Y* scale, compared with contract employees at Company *Z*."

A good hypothesis clearly states the variables under study based on their operational definitions, that is, how they are defined for the study. You should use similar phrasing when writing research questions and hypotheses, including the order in which the variables are presented, such that the independent variable is stated first followed by the dependent variable (Creswell, 2014). A good hypothesis permits empirical testing such that the data collected and statistical analysis performed will indicate whether the null hypothesis is supported or not. In hypothesis testing, if the null hypothesis is rejected, the alternative hypothesis is accepted. Also note in the example in the previous paragraph that the hypothesis contains information indicating how the variables were operationalized in the research study; this is also a characteristic of a complete and clearly stated hypothesis.

To summarize, hypotheses are predictive statements, or educated guesses, of outcomes of statistical tests designed to answer the research question in quantitative studies wherein the research question is looking at differences between groups or relationships between variables. Hypotheses are written to cover all possible answers to the research question. The null hypothesis states that the data will not support the prediction. The alternative hypothesis states that the data will support the prediction. Hypotheses can be either nondirectional or directional. A nondirectional hypothesis predicts that any difference or relationship between groups or variables can be either more than or less than. A directional hypothesis assumes that any difference or relationship is only more than or less than and tests only for the expected direction.

Conceptual Framework

The conceptual framework argues for the importance of studying a topic and why the chosen design and methods are appropriate to the study in compliance with rigorous standards. A conceptual framework ordinarily contains a theoretical framework that explicates how the study either generates or tests theory and situates the study in the theoretical dialogue of the profession.

Summary of Study Rationale

The rationale for a dissertation study is typically presented as part of the introduction section of the research proposal. The background section summarizes the more extensive literature review and ends with a statement of what is not yet known or understood about the topic, a gap statement. The gap statement is the basis for statement of the problem, which incorporates at least three elements: (a) the problem (what is not known or understood), (b) for whom and why studying the problem is important, and (c) a general statement of how the study will address the problem. How the study will address the problem is further specified in the purpose statement, which describes study design, theory or phenomenon of focus, intent, variables/phenomenon, participants, and site. The research questions turn the purpose statement into operable interrogatives that will be answered through the study implementation. A conceptual/theoretical framework presents both justification for and explanation of the study.

Study Implementation Plan

As shown in Figure 10.1, one of the purposes of the research proposal introduction is to summarize the design, methods, and other delineators that specify how the study will be implemented. In this section, we will briefly define how the proposal presents those elements in the nature of the study and further delineations of the study.

Nature of the Study

The Nature of the Study section in the introduction of a proposal may be termed differently by different institutions, but, regardless of title, it contains a summary of the detailed description of the research design and methods, which is usually presented in a major section, or chapter, of the proposal. As such, it summarizes the research design and the methods.

Design. The research design includes the selected approach to inquiry (quantitative, qualitative, or mixed methods), the specific design selected within the approach with appropriate rationale, and other approaches to inquiry or designs considered and the rationale for why they were rejected. The methods include participant selection, instrumentation, data collection procedures, and data analysis strategies.

You need to identify the research approach chosen for the study and then identify the specific design you have selected within the approach. For example, you might select a qualitative phenomenological design or a quantitative pre- and posttest quasi-experimental design. You will then need to explain how the selected design best serves the study purpose.

Methods. In the Nature of the Study section, you will need to summarize the methods for the study by describing from whom you will collect data, how you will select participants, what instruments you will use for data collection (such as surveys, tests, observations, and interviews), the procedures for data collection, and the data analysis plan. Each of these methodological elements (instrumentation, participant selection, data collection, and data analysis) are discussed in prior chapters in this text. The point here is to clearly present each element in your research proposal.

After reading this section, you should have a clear idea of how you will conduct your study.

Study Delineators

In addition to the Nature of the Study section, the implementation plan portion of the proposal specifies a number of delineators that clarify the study, including term definitions, assumptions, scope and delimitations, limitations, and significance.

Definition of Terms. Definition of terms is critical for positioning a study within the professional conversation on a topic. To converse, professionals need to know what terms mean, and the definitions should leave no ambiguity. Dictionary definitions are irrelevant for a research study. For example, it is easy to get a definition for *depression* from the dictionary. However, this does not tell the reader how depression is being operationally defined for your study; this is what is important to other researchers. You need to provide definitions that explicate how you will use the term in the study (Leedy & Ormrod, 2013). You should provide definitions to specify the participants in the study, variables (quantitative) or phenomena (qualitative), and any other terms that are used in your study.

In relation to defining participants, for example, the dictionary definition of *teacher* is insufficient to inform a reader of how the researcher identifies teachers as part of the recruitment strategy. Is the researcher including anyone with an educational license, such as a school counselor or librarian, or is the researcher including only those licensed individuals who instruct students in a classroom? Is the researcher including specialists, such as special education teachers? Are the teachers eligible for the study only those who work in public schools, or are teachers who work in private schools also eligible? For example, a definition of *teacher* for a study might read, "Individuals licensed as an elementary teacher by the state of *X* who currently hold positions teaching in K-3 regular education classrooms in public schools in the state of *X*."

For a concept that can have various definitions and manifestations, operational definitions specify the bounds around the concept. An operational definition states what will be recorded as evidence of the concept contained in the term, that is, what can be seen or heard to evidence presence of the concept. Given the earlier example of employee motivation, you might define employee motivation as frequency of contribution to group goal attainment. Your definitions of variables and phenomena must be rooted in the professional literature. In other words, you cannot create your own definition of, say, motivation, but must cite one or more sources in the professional literature that support your definition of the concept.

Assumptions. In a proposal, you must state any assumptions on which your study is based. "An assumption is a condition that is taken for granted without which the research project would be pointless" (Leedy & Ormrod, 2013, p. 5). For example, in a study of a new method of teaching arithmetic to second graders, it would be important that the teachers using the new method had been trained in the method. An assumption of the study, then, would be that teachers had been trained in the method. If they had not been trained in the method, then the study integrity comes into question. There needs to be a basis for an assumption, a reason why it is valid to hold the assumption. For example, a basis for the assumption that teachers had been trained in the new arithmetic teaching method may be that the districts have documentation for the training.

Crawford (2015) identified three critical attributes of an assumption:

1. The assumption must relate to a critical, not trivial, condition of the study,
2. There is a basis for the researcher making the assumption, and
3. The assumption must relate to a procedure of the study that is not fully within the control of the researcher.

The first critical attribute of an assumption requires that the assumption relate to a critical condition of the study. For example, to state an assumption that all participants in the study can read English when all participants in the study are English teachers may be considered trivial, because reading English would be a characteristic of the population. However, it may be important to state an assumption that teachers in the study can read English if the study is being conducted in a country where Spanish is the first language.

The second critical attribute requires statement of a basis for the assumption. For example, following from above, if the study is being conducted in a country where Spanish is the first language and there is the assumption that the teacher participants can read English, the basis for assuming that the members of the population can read English must be stated. Perhaps all teachers in the population must pass an English reading examination in order to obtain teaching credentials or there is some other documentation of the English reading competency of the participants.

Finally, assumptions must relate to procedures of the study that are not fully within your control. For example, to write an assumption that the participants in the study meet study requirements is not valid. Selection of participants is within your control, so it cannot be an assumption; you control participant selection. However, if a school district provides professional development for a certain teaching method, that provision is not fully within your control, and you can assume that the professional development has been delivered based on documentation of delivery of the professional development. Furthermore, that provision does relate to a procedure of your study—participant selection.

A well-written assumption statement will incorporate all three attributes: a critical condition, a basis for making the assumption, and a study procedure that is not fully within your control. For example, for a study on principal leadership and strategic planning, the following assumption statement meets all attributes: "Principals in this study are familiar with strategic planning, because the state has required all districts to develop strategic plans at the building level." It is critical, not trivial, for the study that principals are familiar with strategic planning. The basis for the assumption is that the state has required all districts to develop strategic plans at the building level, so principals are likely familiar with strategic planning. Finally, the assumption is related to a study procedure, participant selection, but the degree to which principals participate in the state-required strategic planning initiative is not within your control.

Scope and Delimitations. The scope of a study frames the group to which the study might be applicable. For example, a study might examine friendship patterns among highly mobile (defined as having moved five or more times within the past 3 years) female military spouses within the European and Asian military sectors. This study, therefore, would be applicable to that population; that is the scope of the study, or the broad group for whom the study might be applicable.

The delimitations narrow the study in terms of participants, time, and/or location by stating what the study will not include. For example, the study of friendship patterns of highly mobile military spouses might be delimited by not studying male military spouses and by defining highly mobile to exclude those female military spouses who have moved fewer than five times within the past 3 years. The study might further be delimited by not studying female spouses who have moved beyond the European and Asian military sectors.

Beyond setting the boundaries of the study, stating the scope clearly helps in writing the study title, and stating the delimitations clearly helps in specifying the study population. Therefore, the study title and scope must be aligned, and the study delimitations and population definition must be aligned.

Limitations. Delimitations and limitations are often confused. Delimitations narrow the study by stating what the study does not include. Limitations identify weaknesses in the study design or methods. All studies have design and/or methodological weaknesses; the point is to declare known weaknesses and describe what has been done to overcome them and/or what has prohibited overcoming them. For example, a weakness (limitation) of a study may be that parents in the population typically work several jobs and availability for interviews is limited. To overcome that limitation, you might extend a broad range of hours and locations for your interviews. A study of adoption practices may be limited if the study is conducted in only one state or region, with unknown applicability to other states or regions. If you found yourself in that situation, you may declare that the limitation may not be overcome because of a feasibility issue in accessing participants beyond the stated geographic area.

Significance. As discussed earlier, the problem statement suggests for whom and why the study might be important. That part of the problem statement can evolve into the significance section of the proposal. Studies are significant for one or more of three reasons: (a) They can influence practice, (b) they can affect policy, and/or (c) they can generate future research. For example, a study might be of practical importance to therapists in that it identifies a population for whom a particular therapeutic strategy is effective. Another study might impact policymakers considering budget allocations for preschool education by identifying long-term benefits of preschool education. In terms of generating future research, a study might provide new questions for understanding how parental behaviors influence the development of self-esteem in children.

In your significance section, you will need to anticipate how your study will be important for one or more of those reasons and expand on for whom it will be important. This section is often called the "so what?" section. If the study is not important to someone for some reason, then so what if it is done?

Summary of Study Implementation

After providing a rationale for the study, you summarize how the study will be conducted, including design and methods, definitions, assumptions, scope and delimitations, limitations, and significance. The rationale and implementation sections of study introduction must be consistent within each section and also aligned between and among sections. We turn our attention to the issue of alignment of design and each methodological element.

Aligning Research Design, Research Questions, and Methods

The issue of internal alignment is an important aspect of study design and methods. In the earlier parts of this chapter, we described each of the key sections of the research proposal with technical information on how you can craft them.

In this section, our focus is on how to go back, when the proposal is completed, to ensure that you have consistency between all of the key proposal sections. For example, you must take care to ensure that the research question(s) and the methods, which include the analytic strategies, are consistent with the research design that will best help advance knowledge in a particular field. In addition, the problem statement, which describes a gap in the literature, and the purpose statement both must align with the research questions, design, and methods. Many programs require students to craft a concept paper, prospectus, or some other form of an abbreviated proposal that allows for discussion of the study with members of a research team; one specific purpose of this discussion is to ensure that the overall design leads to answering the study research questions.

Aligning research design, research questions, and methods is important and does much to convey the scientific rigor of the research. Choices regarding research design and methods always begin with the right research questions. As a review of more detailed information presented earlier in this chapter, the first step in the process is to carefully review the literature to understand the state of development of the topic of interest, what is known as well as what is not known, and the key themes that surface only in an exhaustive review of the literature. This information is then used to formulate a problem, a study purpose, and precise research questions that define the scope of the study. The choice of research design and methods clearly follows from the research question(s). You likely have a worldview that orients you to a particular philosophical orientation or paradigm (positivist, postpositivist, or constructivist). These interests will naturally lead you to preferred approaches to inquiry, including qualitative, quantitative, or mixed methods approaches. These preferences then lead to particular types of research questions. However, the final selection of approach to inquiry and methods must still follow from the choice of research question. Once you formulate the research question, the choices you make for the appropriate approach to inquiry and methods will be validated by members of your dissertation, thesis, or research committee. The committee helps ensure that question, approach, and methods are logically consistent and make sense.

Research Design

Within qualitative, quantitative, and mixed research approaches to inquiry, you have several design options. The research design is the particular strategy you will choose to answer the study questions. Alignment starts with the approach to inquiry. If you are primarily looking to generate theory or to explore a topic in more depth and through an interpretive lens (both of which involve a more inductive approach to knowledge generation), a qualitative research design is likely most appropriate for you. On the other hand, if you seek more stringent tests of theory to ascertain cause and effect, you would likely be choosing a quantitative research design in the quantitative approach to inquiry. The correct choice of paradigm or philosophical orientation is the result of a thorough investigation and analysis of the literature that yields a clear understanding of the state of development of the discipline related to the topic of interest. For example, it is important to understand whether a phenomenon of interest has been the subject of extensive research or whether the research on that topic is still underdeveloped. Such knowledge leads to the best choice of approach to inquiry based on whether the particular topic area requires more descriptive development to support theory generation (qualitative) or whether theory development is advanced to the point that strategies to ascertain cause and effect are needed (quantitative).

Research Questions

Research questions define the key objective for the study. The way research questions are worded usually suggests a particular approach to inquiry and, within that approach, a particular research design. For example, from the perspective of approaches, questions that include words such as *describe* and *explain* signify a qualitative approach, whereas words such as *compare*, *correlate*, and *difference* tend to be associated with a quantitative approach. A question such as “What are the lived experiences of female chief executive officers of Fortune 500 companies?” would best be addressed by a phenomenological research design because the focus on lived experience is consistent with a phenomenological epistemology. Such a question embedded in a grounded theory design would likely not make sense because the purpose of grounded theory is to generate theory rather than to describe lived experiences. In that case, the research questions and design would not be aligned.

Methods

Methods include components of the study that describe how a particular research design will be achieved. The Methods section involves explication of how you plan to achieve answers to the research questions. Methods include aspects of the study such as context (setting) as well as how participants are selected; these are usually described as the sampling or participant selection methods. The Methods section also includes data collection procedure and instrumentation and the methods used to structure, clean, and analyze data. Together, these aspects of the study define its methods.

Each particular design has its own characteristic methods that are derived from its ontological and epistemological roots. For example, grounded theory, a qualitative design, comes from an ontologically relativist perspective with a subjectivist epistemology. Grounded theory designs require a relatively small (purposive) sample. The sample (as in all designs) is constrained by the delimitations and the setting of the study. The instrumentation used in a grounded theory design includes semistructured questions meant to address aspects of the phenomenon under consideration, and the research questions can change as the interviews progress; in grounded theory, analysis starts with the very first interview, and questions are shaped by that interview that guides the next interview. This changing of research questions as the study progresses is an aspect of the methods that is a unique but essential feature of grounded theory. This feature also reinforces the relativist and subjectivist orientations of this kind of design.

A randomized controlled trial, on the other hand, begins with a positivist philosophical perspective that is ontologically realist and epistemologically objectivist. There is a sampling method that is specified; for example, the sample may be obtained randomly or by convenience. The sample size is prescribed in advance of the study and is based on specific formulas that support achieving statistical power necessary for generalizability of results. The method of data collection and analysis is always articulated. Data are then collected in their entirety, and analyses are specified in advance by the research questions and hypotheses. The analyses selected must be consistent with those that can provide answers to the research questions. Thus, quantitative designs tend to be much more prescribed in advance of data collection and analysis. It is important to note, however, that once data are collected and analysis begins, other analyses of emerging questions may be conducted that complement results from tests of the null hypotheses. As part of the analysis of the a priori hypotheses, you might find results that suggest other hypotheses that can be tested that will provide a richer understanding of the phenomenon. You are encouraged to explore the data; just take care that a priori hypotheses are distinguished from those that emerge in the process of data analysis.

Each of the examples describes key components of the specific approaches to inquiry, design, and methods that require alignment. Ontological and epistemological choices drive interest in the approach to inquiry. The designs best suited for the research questions are selected, and tools and settings appropriate to those designs are selected and described. In the grounded theory example, semistructured interview protocols are developed that serve as the instrumentation. We noted that specific interview questions change as interviews progress, which is a defining feature of grounded theory. In the case of the randomized controlled trial, specific, structured questionnaires (or other data collection tools) are used to gather data for analysis of a priori hypotheses. This, also, is a defining characteristic of quantitative research in general and randomized controlled trials in particular. To summarize, you need to ensure that all aspects of the approach to inquiry, design, and methods are each internally consistent and aligned with other aspects of the proposal.

Table 10.1 provides some questions that can be used to help guide the alignment process. In addition, members of your research committee or team can be invaluable in providing verification that the proposal makes logical sense.

Table 10.2 provides information that you may find helpful in aligning various aspects of the quantitative, qualitative, and mixed methods proposals. Within a given paradigm, you can see aspects of the scientific approach, philosophical foundation, nature of the study, research design, instrumentation, data, analysis strategies, software, and reliability/validity considerations that would be consistent within that paradigm. During proposal writing, if one of these components you are describing falls into one of the other paradigms, it can be a sure sign that misalignment may be occurring.

Table 10.3 provides another tool that many new as well as seasoned researchers find useful. The Historical Alignment Tool (HAT) can be used in the earliest stages of proposal writing to track in a linear way the development of the proposal's key components. By reading across the columns, it is easy for you or those reviewing the research project to get a quick snapshot of the elements described earlier—research questions, approach to inquiry, and methods—to check for alignment. In addition, you can track any changes based on conversations with other committee members or research team members and identify changes made to ensure overall alignment once the adjustment is made. This tracking of changes results in a document that shows the evolution of your study.

TABLE 10.1
Aligning Research Question, Research Design, and Methods

Before Writing the Proposal
• Have you conducted a thorough review of the literature in the field that provides evidence for the state of advancement of knowledge? • From the literature review, can you clearly identify what is known as well as what is not known? • Have you framed knowledge of what is known and not known into a research question that is of sufficient scope that it can be answered in the context of your study? • Have you thoroughly reviewed the literature related to the design, including the study of seminal resources in the field, which provides its theoretical and philosophical underpinnings? • Have you reviewed research studies that are similar to your own to generate evidence to support decisions on methods, including population, sample, context, instrumentation, and analysis?
Writing the Proposal
• Does the research question make sense (i.e., can it be answered) in the context of the research design? Does it make sense in terms of the chosen approach to inquiry? • For quantitative studies, are the statistical hypotheses consistent with the research questions (i.e., will answering the hypotheses adequately address the research questions)? • For qualitative studies, are the interview questions and subquestions consistent with the research questions? Will the interviews likely generate data that can answer appropriately the research question? • In qualitative and quantitative studies, is the required sample size consistent with the research question and design? • Are the analysis strategies consistent with the chosen research design?

Summary

A precisely written proposal will serve you well as you conduct your study. There are three major sections of a research proposal, the organization and content of which may vary by institution: literature review, study introduction, and design and methods. The first section of a proposal provides rationale for the study and summarizes the plan for conducting the study.

Your careful review of the literature situates the study rationale by surveying what is already known about a particular phenomenon under study. You summarize your review of the literature in a background section of the proposal. The review results in statements that articulate the research gap (what is known and what is not known in your area of study). The research gap statement focuses the study on some aspect related to the topic that, prior research notwithstanding, still is not fully known or understood. That gap statement forms the kernel of the problem to be addressed by the study. A proposal contains a clearly written problem statement that contains (a) what is not yet known or understood about the topic, that is, the

TABLE 10.2
Aligning Study Components Within Research Paradigms

	Qualitative	Quantitative	Mixed Methodologies
Scientific approach	***Inductive*** (Generate meaning or understanding from data.)	***Deductive*** (Test hypotheses derived from theories.)	***Inductive and deductive approaches***
Philosophical foundations	***Postmodernism/constructivism*** (knowledge is shared or constructed; focus on individual meanings, points of view)	***Positivism*** (There is an objective reality, or truth, that can be measured.)	***Includes aspects of both*** but mixed methodologists likely would embrace more of a postmodern/constructivist perspective.
Nature of studies	***Descriptive***	***Correlational, quasi-experimental, experimental***	***Descriptive, correlational, quasi-experimental, and experimental***
Research designs	***Grounded theory*** (Develop theory in an area where theory is lacking or existing theory is inappropriate.) ***Phenomenology*** (understand the meaning a phenomenon has for participants; lived experience) ***Case study*** (examines a particular event, occurrence, phenomenon in its natural context; understand the factors influencing members of the case) ***Ethnography*** (in-depth analysis of a culture) ***Narrative*** (an approach that analyzes stories of life experiences and the meaning people make of those experiences) ***Participatory action research*** (in which those being researched are coresearchers and work together to solve socially relevant problems)	***Surveys*** (self-administered; interviewer administered; computer assisted [CASI and ACASI systems]; Internet based - Surveymonkey) ***Observations*** (counting behaviors; counting and reporting behaviors in experiments) ***Experiments*** (quasi-experiments in which groups are predetermined; randomized controlled trials in which participants are randomly assigned to groups, as in drug efficacy studies)	***Parallel qualitative and quantitative approaches*** ***Sequential studies*** (qualitative → quantitative or quantitative → qualitative) ***Example of qualitative → quantitative:*** Interview people in a community, and from this, develop instruments that are appropriate to that community. ***Example of quantitative → qualitative:*** Participants complete an instrument/scale and then participate in in-depth interviews to understand more about why they answered the way they did.

Instrumentation	***Interview protocols*** (typically unstructured or semistructured interviews that used more open-ended questions) ***Observations documents*** (e.g., school records, meeting minutes, mission statements, organizational planning documents, annual reports, historical records, government documents, journals, diaries)	***Scales or instruments*** (require responses on some kind of forced scale of measurement, such as yes/no or Likert scales) ***Observation and counting protocols*** (e.g., counting specific behaviors on a playground that provide quantitative data for analysis)	Combination of interviews and scales/instruments, depending on the objectives/goals of the mixed method study.
Data	***Text*** (interview transcripts, observation notes, documents, etc.)	***Numerical***	***Text and numerical***
Analysis strategies	***Open coding*** (initial coding of data into blocks) ***Axial coding*** (Emerging concepts are dimensionalized in a grounded theory approach.) ***Constant comparative analysis*** (comparing emerging codes across participants) ***Thematic analysis*** (searching for themes) ***Narrative analysis*** (analyzing stories and narratives by structure, function, or oral performance)	***Correlation and regression*** (Pearson *r*, Spearman rho, multiple and logistic regression, path analysis and structural equation modeling) ***Group comparison tests*** (analysis of variance, analysis of covariance, multivariate analysis of variance) ***Instrument development*** (factor analysis, item response analysis)	***Qualitative and quantitative*** (analysis methods chosen as appropriate to the mixed method study)
Software examples	***NVivo*** ***AtlasTI*** ***Ethnograph*** (Use of software is not required in qualitative research, but many use it to help manage data.)	***SPSS*** ***SAS*** ***R*** ***Stata***	***Uses both***

continued

TABLE 10.2 *(continued)*

Aligning Study Components Within Research Paradigms

	Qualitative	Quantitative	Mixed Methodologies
Reliability and validity	***Trustworthiness, including credibility, dependability, confirmability, and transferability*** (Lincoln & Guba, 1985; strategies include prolonged engagement with data; persistent observation of phenomenon under study; triangulation of findings; debrief findings and analyses with peers; member checks [checking results with those who provided original data]; thick description) Note: This is just one set of criteria for evaluating the validity of qualitative research. There are other sets of criteria proposed by other authors.	***Test-retest reliability*** (measures the same construct in the same person at two different times) ***Internal consistency reliability*** (Cronbach's alpha; assesses how well items in a scale are correlated with each other) ***Split-half reliability*** (split scale in half and correlate the items in each half) ***Concurrent validity*** (Measure correlates with an existing measure that assesses the same construct.) ***Discriminant validity*** (Measure does not correlate with a measure to which it is not theoretically expected to be related.) ***Construct validity*** (face validity [the instrument appears to measure what it is expected to measure] and content validity [experts review items]) ***Predictive validity*** (Construct predicts another construct to which it is theoretically expected to be related.)	***Approaches appropriate to the needs for the study*** (Triangulation is popular in mixed methods research [multiple sources of evidence converge on the same findings].)

Note. Table 10.2 used with permission of G. Burkholder and M. Spillett. CASI, computer-assisted self-interview; ACASI, audio computer-assisted self-interview.

TABLE 10.3
Historical Alignment Tool

Planned Research Focus:	Enter your planned research focus here.							
	Dissertation Research Components							
	Research Framework					Research Design and Methodology		
Dissertation Planning & Revision History	Problem Statement	Purpose	Significance	Theoretical Framework	Research Question(s)	Nature of Study/ Approach to Inquiry	Possible Sources of Information or Data	Possible Analytical Strategies
Original Plan								
Rationale for Change								
Revised Plan								
Rationale for Change								
Revised Plan								
Rationale for Change								
Revised Plan								
Rationale for Change								
Revised Plan								

problem itself; (b) for whom and why studying the topic is important, which is a precursor to significance of the study; and (c) how the study will address the problem, which implies the purpose of the study. The purpose statement informs the reader of study (a) design, (b) related theory or central phenomenon, (c) intent, (d) variables of phenomenon definition, (e) participants, and (f) site. Research questions and hypotheses follow from study purpose and design; they focus the study, set study boundaries, and inform data collection.

With the study rationale fully explicated, the study introduction summarizes the design and methods for the study, as well as other delineations that frame the study. The design (overall study structure) and methods (participant selection, instrumentation, data collection procedures, and data analysis plan) need full discussion in a section devoted to design and methods but may also be summarized in a Nature of the Study, or similarly titled, section of the proposal introduction. The proposal must also delineate the study in terms of (a) definitions, (b) assumptions, (c) scope and delimitations, (d) limitations, and (e) significance.

Alignment of study elements must flow from gap statement through to methods, with focus on answering the research questions. Alignment is enhanced with an orderly progression from gap statement to problem, to purpose, to research questions, to design, and finally, to methods. If the research questions are not clearly crafted, alignment with the design and methods will falter. Use the information provided in Figure 10.1 to help you to see the broad ways approach to inquiry, design, and methods are aligned. Use the HAT provided in Table 10.3 as a practical guide for tracking alignment as you change and fine-tune key aspects of the proposal. Following these steps will maximize the probability that you will have a proposal that shows strong internal consistency and alignment.

Acknowledgment

The authors wish to acknowledge the contributions of Laura Schindler, PhD, and Ronald Paige, PhD, who provided sound critique of this chapter from student and faculty perspectives.

KEY SOURCES

Krathwohl, D. R., & Smith, N. L. (2005). *How to prepare a dissertation proposal: Suggestions for students in education & the social and behavioral sciences.* Syracuse, NY: Syracuse University Press.

Leedy, P. D., & Ormrod, J. E. (2013). *Practical research: Planning and design* (10th ed.). Boston, MA: Pearson.

Simon, M. K., & Goes, J. (2013). *Dissertation & scholarly research: A practical guide to start & complete your dissertation, thesis, or formal research project.* Dubuque, IA: Kendall/Hunt.

REFERENCES

Agee, J. (2009). Developing qualitative research questions: A reflective process. *International Journal of Qualitative Studies in Education*, *22*(4), 431–447. doi:10.1080/09518390902736512

Allard, K. E. (2014). *Special education team member perceptions of video self-modeling with twice-exceptional students* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3645160)

Booth, W. C., Colomb, G. G., & Williams, J. M. (2008). *The craft of research* (3rd ed.). Chicago, IL: University of Chicago Press.

Crawford, L. (2015). *Assumptions*. Unpublished manuscript.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.

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

Leedy, P. D., & Ormrod, J. E. (2013). *Practical research: Planning and design* (10th ed.). Boston, MA: Pearson.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.

Onwuegbuzie, A. J., & Leech, N. L. (2004). Linking research questions to mixed methods data analysis procedures. *The Qualitative Report*, *11*(3), 474–498. Retrieved from http://www.nova.edu/ssss/QR/QR11-3/onwuegbuzie.pdf

Tashakkori, A., & Creswell, J. W. (2007). Exploring the nature of research questions in mixed methods research [Editorial]. *Journal of Mixed Methods Research*, *1*(3), 207–211. doi:10.1177/1558689807302814

White, P. (2013). Who's afraid of research questions? The neglect of research questions in the methods literature and a call for question-led methods teaching. *International Journal of Research & Method in Education*, *3*, 213–227. doi:10.1080/1743727X.2013.809413

11

Grounded Theory

Annie Pezalla, PhD

Introduction

In this chapter, I provide an overview of the concepts associated with grounded theory approach and analysis. First, I introduce you to some of the origins and philosophical underpinnings of grounded theory. These are followed by examples of how grounded theory has been used within several disciplines, information on appropriate sample size, and matters related to data collection and analysis. I will discuss constant comparative analysis, which is somewhat unique to grounded theory and results in interviews whose focus can change over the course of the study. I then clarify the role of the literature review in grounded theory research and its importance to crafting a proposal that has a defensible argument. Finally, I will discuss the benefits of grounded theory research.

Background

Grounded theory is a qualitative research design that researchers use to develop theory inductively from data. In grounded theory, you do not begin with a theory and then attempt to support or disprove it. Instead, you begin with an area of study and allow what is relevant within that area to emerge into a theory.

Origins of Grounded Theory

Grounded theory design emerged from two sociologists, Barney G. Glaser and Anselm L. Strauss, and their 1960s research on dying in hospitals—a topic that was rarely studied at that time. Glaser and Strauss (1967) observed dying patients and how an awareness that they were dying influenced their interactions with relatives and hospital staff. As they constructed their analyses of dying, they developed systematic methodological strategies that social scientists could adopt for studying other topics. Their book, *The Discovery of Grounded Theory: Strategies for Qualitative Research* (Glaser & Strauss, 1967), first articulated these strategies and advocated for developing theories from research grounded in data rather than deducing testable hypotheses from existing theories.

Philosophical Underpinnings of Grounded Theory

Grounded theory marries two contrasting and competing traditions in sociology, influenced by differences in the backgrounds of Glaser and Strauss. Glaser came from Columbia University, where quantitative methods were a strong influence. Glaser's positivist training resulted in his focus on codifying qualitative methods, generating "middle-range" theories (i.e., abstract renderings of specific social phenomena that were grounded in data), and maintaining a dispassionate and empirical process. Glaser saw the need for a well-thought-out, explicitly formulated, and systematic set of procedures for both coding and testing hypotheses generated during the research process. His Columbia background also emphasized empirical research in conjunction with the development of theory.

Strauss came from the University of Chicago, which has a long history in qualitative research. In contrast to Glaser, symbolic interactionist and pragmatist writings heavily influenced Strauss. Symbolic interactionism emphasizes understanding the world by interpreting human interaction, which occurs through the use of symbols, such as language. Strauss's background in symbolic interactionism contributed to the following aspects of grounded theory:

- The nature of experience and undergoing as continually evolving
- The active role of persons in shaping the worlds they live in
- An emphasis on change and process and the variability and complexity of life
- The interrelationships among conditions, meaning, and action (Strauss & Corbin, 1990)

Pragmatism emphasizes practical consequences as the primary criteria in determining meaning or truth. Strauss's pragmatist background contributed to the following aspects of grounded theory:

- The need to get out into the field, if one wants to understand what is going on
- The importance of theory, grounded in reality, to the development of a discipline

Grounded theory's emphasis on humans as active, rather than passive, agents is a result of Strauss's influence. Strauss also emphasized the notions of emergent processes on social and subjective meanings.

Grounded Theory: Qualitative or Quantitative?

Researchers traditionally view grounded theory as a qualitative design because it often uses qualitative data sets, such as interview transcripts or observational notes. However, grounded theory can also employ quantitative data sets. For example, you could use grounded theory to investigate how medical students learn and employ ethical behavior in their interactions with patients. You could create a quantitative data set from students' self-reported numerical rankings of the value of ethics-based courses and probability of behaving ethically in specific contexts. You could collect qualitative data through observations of medical students' interactions with patients. Using analysis of both data sets, you could then generate a grounded theory about the relationship between medical students' experiences with ethical behavior in their classrooms and in their interactions with patients. Whatever the method employed, the main goal in grounded theory studies is to understand the meanings behind human behavior. Despite the possibility of either quantitative or qualitative sources of data for a grounded theory study, because of the preponderance of qualitative-based grounded theory studies, this chapter focuses on qualitative grounded theory (see Box 11.1).

Grounded Theory: Alignment and Design

In this section of the chapter, I provide information on the types of research topics and questions explored with grounded theory, along with examples illustrating the use of grounded theory with other methodologies. In addition, I discuss the role of the researcher, as well as sample size and ethical considerations, in doing grounded theory.

Research Topics for Grounded Theory

Grounded theory research topics should aim to discover emergent ideas or connections and increase understanding of meaning, context, process,

BOX 11.1

Discipline-Specific Applications of Grounded Theory

Any discipline can use grounded theory as an approach to a research question. What matters in grounded theory are the procedures, and they are not discipline bound. For example:

- In psychology, therapists have used grounded theory to understand the experiences of parents who have difficult-to-treat teenagers with attention deficit hyperactivity disorder (ADHD) (Brinchmann & Sollie, 2014).
- Sociologists have used grounded theory to discover the meaning of spirituality among cancer patients and how spiritual beliefs inform patients' attitudes toward cancer treatments (Landmark, Strandmark, & Wahl, 2001).
- Public health researchers have used grounded theory to examine nursing home preparedness needs through the experiences of Hurricane Katrina evacuees sheltered in nursing homes (Laditka et al., 2008).
- In business, managers have used grounded theory to explain the ways in which top-level corporate managers can be politically successful (Adjognon, 2014).
- Engineers have used grounded theory to explore the relationship between employee perceptions of innovation and aspects of that innovation among product-focused firms (Braddy, Meade, & Kroustalis, 2006).
- Information technology researchers have used grounded theory to investigate older adults' use of computer technology (White & Weatherall, 2000).

how things happen, and the ways in which *X* affects *Y*. For example, the Ellingson (2002) study increased understanding of how companions of geriatric oncology patients affected interactions between the patients and their health care providers.

Other examples of research topics explored with grounded theory include

- The meaning of chronic illness (Charmaz, 2006b)
- The context of successful leadership among senior executives (Parry, 1998)
- The process of mothering in a neonatal nursery (Fenwick, Barclay, & Schmied, 2001)
- The experiences and consequences of professional nurses' moral distress (Nathaniel, 2004)
- The role of school-based snacking policies on the consumption of snack foods among 9- and 11-year-old children (Freeman, Ekins, & Oliver, 2005)
- How middle-aged women cope with menopause and with issues of hormone therapy (Ekström, Esseveld, & Hovelius, 2005)
- The issues surrounding various ability grouping models in middle-level mathematics classes (Johnston, 2008)

In exploring topics such as those listed above, research questions should always be open-ended (i.e., they cannot be answered with a simple yes/no response). For example, the research questions from the Ellingson (2002) article are "What roles do companions play in the geriatric oncology patient-interdisciplinary team member interaction?" (p. 366) and "How do companions' roles shift or remain stable across discipline of the geriatric oncology team members?" (p. 368). Note that both of these questions are open-ended questions and are nicely amenable to a qualitative grounded theory approach.

Can You Use Grounded Theory With Another Design?

You could use grounded theory with another research design in a mixed-methods study as a first step toward identifying testable constructs and hypothesizing about their relationship to one another. You could also employ a quantitative, cause–effect design after you carried out a grounded theory design to further test the grounded theory. For example, Corlett's (2000) action research study used grounded theory to help a group of nurses understand (a) why they were having conflict in their clinical training and clinical placements and (b) the ways they could alleviate that conflict. As another example, Carter and Dresner (2001) employed a case study research project within a grounded theory design to examine, and then create a theory of, the ways in which government regulations impact successful environmental projects. Both examples illustrate that different designs can be used within the grounded theory approach. The important thing to remember is that, no matter the design, if a grounded theory approach is involved, the end product of the study must be a theory, one that is grounded in the data.

Determining Sample Size for Grounded Theory

Determining an appropriate sample size for grounded theories has been an elusive process. The main goal in any grounded theory study is to gather enough data until theoretical saturation has occurred—or the point at which no new properties emerge from the identified categories in one's theory. This goal focuses more on the amount of descriptive data than it does on the number of people to recruit.

Despite grounded theory's focus on saturation, rather than sample size, qualitative researchers have continued to discuss the general parameters of when saturation *might* occur within particular sample sizes. Qualitative researchers have recommended sample sizes ranging from as few as 6 participants to as many as 30 for a grounded theory study (Creswell, 1998); however, no rationales exist for those recommendations. Only one study has currently investigated saturation in a quantifiable way, by examining the point in data collection at which codes were created and the point at which the distribution of code frequency stabilized (Guest, Bunce, & Johnson, 2006). In this study, no new codes emerged after 12 interviews; in fact, basic codes were present as early as 6 interviews. Although Guest et al. (2006) did not intend to generalize their results to all grounded theory studies, their findings could be helpful for researchers who hope to establish an initial goal for the number of participants to recruit. Based on this information, a sample size ranging from 12 to 30 might be appropriate for a grounded theory study.

A smaller sample size of 10 to 12 is appropriate if your sample is relatively homogeneous, your participants possess "cultural competence" (Guest et al., 2006, p. 74), or a certain degree of expertise about the domain of inquiry, and your objectives in your research are fairly narrow. A larger sample size of 25 to 30 is appropriate if your sample is diverse, your participants possess varying levels of expertise or familiarity about your domain of inquiry, your research objectives are fairly broad, or you are looking for disconfirming evidence or trying to discover maximum variability in your phenomenon of interest.

Data Collection and Analysis

"All is data" (Glaser, 1998, p. 8). This statement means that grounded theory can use any kind of data. As a grounded theory researcher, it is up to you to figure out what data to gather. All kinds of data can be the building blocks of a grounded theory, including field notes, interviews, audio recordings, observations, and information in records and reports. The kind of data you pursue depends on the topic and access available.

You should also gather "rich, thick data" (Geertz, 1973, p. 10) for a grounded theory study. Obtaining these data means writing extensive field notes of observations, collecting respondents' written personal accounts, and/or compiling detailed narratives (e.g., transcribed tapes of interviews).

Data Collection

Interviews, audio recordings, observational research, memos, and the collection of open-ended questionnaires are common types of data collection in grounded theory studies, but almost any data collection procedure—within ethical limits, of course—is permissible in grounded theory. What is important to remember is that the data collection procedures may change throughout the course of the study, as the grounded theory is refined and sharpened in focus. Charmaz (2006a) likened this process to that of a photographer, honing in on a scene: "Like a camera with many lenses, first you view a broad sweep of the landscape. Subsequently, you change your lens several times to bring scenes closer and closer into view" (p. 14).

Grounded theory researchers often use semistructured, open-ended interviews in their studies, but these may change throughout the course of the study as the topic becomes more refined. With grounded theory, data collection procedures evolve as the theory becomes more refined. Uncertainty is common in grounded theory studies, and flexibility is an absolute must. For example, Ellingson (2002) gathered her data first through ethnographic field notes of interactions between medical professionals, patients, and patients' companions during visits. Then, in the later part of her 2-year period of research, she supplemented these field notes with nine audio recordings of initial patient visits. She transcribed the audio recordings and used them as data for her study. Ellingson's flexibility and adaptability to gather and analyze the most appropriate data as her study progressed is a hallmark of the grounded theory approach.

Data Analysis

The first step in any grounded theory is to become very familiar with your initial data. You can establish this familiarity by reading and rereading your data and/or by listening to the audio recordings of your interviews (if you collected interviews). Transcribing your own interviews is also an important way to become familiar with your data.

The second step is to start identifying codes (a code is a word or phrase that serves to indicate the meaning of a segment of data) and to engage in a process of constant comparison, which is defined as comparing the codes you have created for other segments of data to determine if they are similar or different in meaning.

If you are conducting your first grounded theory study, it is wise to transcribe all of your materials. However, the actual transcribing may be selective. At first, you may decide to completely transcribe and analyze your very first interviews or field notes before going on to the next interviews or field observations. This early coding gives guidance to the next field observations and/or interviews. Later, as your theory develops, you may wish to listen to the tapes and transcribe only those sentences, passages, or paragraphs that relate to your evolving theory. Regardless of whether you transcribe all or part of your tapes, it is still important to listen to the tapes in their entirety. Listening as well as transcribing is essential for full and varied analysis.

The third step is open coding, which is the process of identifying, labeling, examining, comparing, and categorizing your codes into larger, conceptual categories, which encompass a variety of similarly themed codes.

The fourth step is the development of hypotheses, where you speculate about the relationship between the categories. In doing this step, you should also seek examples that contradict those hypotheses.

Throughout this process, *you should be collecting more data—as data collection and analysis are intertwined in grounded theory*—and you should continually refine your hypotheses until all examples are accounted for and explained. For example, as Ellingson (2002) collected her data, she used the constant comparative analysis process, where she constantly compared one interaction among a patient, companion, and health care provider with another interaction among a patient, companion, and health care provider, and assessed their similarities and differences. Once she had developed some hypotheses about the patterns she observed, she collected and analyzed more data with which to test her hypotheses and develop more patterns, continually refining them in the process.

Role of the Literature Review

Anyone proposing a grounded theory thesis or dissertation must conduct a thorough literature review—not a scanty one, or one that excuses careless coverage. Some grounded theorists' statements in their literature reviews reflect a cavalier attitude toward earlier works. Give earlier works their due. Completing a thorough literature review strengthens your justification of the importance of the research problem and thus credibility of your project.

Glaser and Strauss (1967) proposed that data collection and analysis should occur before conducting a literature review because they did not want researchers to see their data through the lens of earlier ideas and become biased by any "pet theoretical ideas" or "received theories" (Strauss & Corbin, 1990). This perspective has been heavily disputed (Charmaz, 2006a), particularly among funding agencies and those who work with doctoral students who must, by the requirement of their programs, demonstrate a sophisticated knowledge of leading studies and theories in the field.

An important approach is to use your literature review without letting it stifle your creativity or "strangle your theory" (Charmaz, 2006a, p. 166). Equally as important is the understanding that you may need to rewrite portions of your literature review after you have identified and developed your grounded theory in a way that adequately positions your grounded theory and clarifies its contribution.

For those of you working on a grounded theory thesis or dissertation, your results chapter may be significantly longer than your literature review chapter. That is because your contribution will be a substantial theory, one that is grounded in an abundance of rich, thick data. To even collect your data, though, you will need to demonstrate what is known based on existing research as well as what is not known. You cannot demonstrate that level of understanding without an adequate literature review. Just remember this very practical point: The requirement for a thorough literature review for grounded theory studies is similar to any other major research project or any application for a grant. No one will approve a thesis or dissertation proposal or fund any study if a literature review is not part of the proposal.

Use of Memos in Data Analysis

Memos are an important aspect of data analysis in any qualitative study. Memos are particularly important to a grounded theory study because they stimulate thinking and provide you, the researcher, with opportunities to focus on emergent categories. There are different types of memos that you should write during the process of analysis. Memos common to all qualitative methods include textual, observational, and reflexive. Memos specific to grounded theory include conceptual/theoretical and operational.

Memos Common to Qualitative Methods

Textual Memos. These memos describe how you are thinking about a code you have assigned to a data segment or how you have assigned names and meanings to data. These are also known as *preliminary* and *orienting* memos. Their purpose is to stop you (briefly) from coding so that you can capture interpretive ideas and hunches. Example: "Participant A mentioned feeling overwhelmed with the idea of retiring when she felt so passionate about helping people in her career. I'm going to label that segment of our interview 'retirement as a moral conflict.'"

Observational Memos. These memos describe the context of interviews and observations. They focus on what your other senses are telling you in the research: what you have seen, felt, tasted, or experienced while doing research. Example: "Participant B's home was filled with the smells of freshly baked bread and coffee. I also noticed that she had cleared off a space for me to rest my bag on her typically crowded dining room table. The baked goodies, the coffee, the clean house—all of it gave me the impression that she wanted me to stay awhile for today's interview."

Reflexive Memos. These memos could potentially cut across all others. They are essentially observations of yourself—your voice, impact, and/or challenging roles throughout the process. These memos help you to pay attention to your rising values, feelings, mistakes, embarrassments, and personal insights and to reflect on how those qualities might contribute to how you are making sense of the data. Example: "I can't help but wonder if my participants have been offering socially desirable responses to my questions about moral behavior in the workplace. Maybe I should complement my individual interviews with some open-ended and anonymous surveys? I'll talk with my committee members about that idea to get their feedback."

Memos Specific to Grounded Theory

Conceptual/Theoretical Memos. These memos ask the question "What is going on here?" They describe the development of a category and help to develop ideas on the possible pathways for integrating the theory. Example: "Retirement as a moral conflict" seems connected to the code of "Isolated from professional acknowledgements." Those codes could be clustered together into a broader conceptual category of "Personal and professional identity conflict."

Operational Memos. These memos are very practical. They might have to do with reminding yourself to ask a new question that arose in a previous interview, or about sampling strategies, or who to talk to next. These mostly relate to methodological procedures. Example: "Retirement as a moral conflict": Retirement is a profoundly moral conflict as well as a professional conflict. This code captures any narratives that reflect the clash of moral conflict (e.g., feeling conflicted at leaving a helping profession but wanting to spend more time with family) and professional conflict (e.g., feeling addicted to the accolades that come with professional accomplishments).

How to Ensure Quality When Using Grounded Theory

You can ensure quality in grounded theory studies by *complete transparency* in your data collection process. Through this transparency, a grounded theory researcher is able to demonstrate trustworthiness. Typically, terms like *validity* and *reliability* are not relevant in grounded theory research because those terms are quantitative in origin and are not directly applicable to grounded theory studies. Grounded theory studies, after all, do not hold that there is one unified Truth (with a capital "T") in the world (a relativist position), so it is difficult to claim that one's approach is absolutely valid because truth (with a lowercase "t") is assumed to be rather subjective and largely based on perceptions.

The same problem exists with claims of reliability in grounded theories. Reliability does not apply to grounded theories because reliability is not actually a goal. You should expect the research process to change throughout the study as the topic evolves and the grounded theory becomes more refined.

In lieu of the indicators of validity and reliability then, grounded theories focus on trustworthiness, and there are four main indicators of trustworthiness that a grounded theory researcher should establish: credibility, transferability, dependability, and confirmability. For example, Ellingson (2002) ensured quality by establishing confirmability and credibility in her work. Regarding confirmability, Ellingson kept extensive field notes during her observations (i.e., more than 300 pages). She also demonstrated credibility by staying extensively involved in the field. She conducted observations for 2 years, visiting the hospital weekly for 2 to 5 hours each time. She also amassed "thick" descriptions and triangulated her sources by using transcripts of interviews with patients, companions, and health care providers, along with her observations of them.

A well-constructed grounded theory will meet four central criteria for judging its quality (Strauss & Corbin, 1990):

1. **Understanding.** Because the grounded theory represents reality, it should be comprehensible and make sense both to the participants and to those practicing in the area of study.
2. **Fit.** If a grounded theory is faithful to everyday reality of a substantive area and carefully induced from diverse data, it should fit that substantive area. This means that the grounded theory should not be overly abstract or difficult to apply to real-life scenarios.
3. **Generality.** If the data are comprehensive and the interpretations conceptual and broad, then the grounded theory should be abstract enough and include sufficient variation to make it applicable to a variety of contexts related to that phenomenon. You should be able to apply your theory.

4. **Control.** You should clearly spell out the conditions under which the phenomenon applies, and your hypotheses proposing relationships among concepts should relate to that (and only that) phenomenon.

Benefits of Grounded Theory

The benefits of using grounded theory include the following.

Ecological Validity. Ecological validity is the extent to which research findings accurately represent real-world settings. Grounded theories are typically ecologically valid because they are close to the data from which they were generated. Although the constructs in a grounded theory are appropriately abstract (because their goal is to explain other similar phenomenon), they are context specific, detailed, and tightly connected to the data.

Novelty. Because grounded theories are not tied to any preexisting theory, grounded theories are often fresh and new and hold the potential for innovative discoveries in science.

Parsimony. Parsimony is about using the simplest possible description to explain complex phenomenon. Think Occam's razor: One should not increase, beyond what is necessary, the number of entities required to explain anything. Grounded theories aim to provide parsimonious explanations about complex phenomena by rendering them into abstract constructs and hypothesizing their relationships. They offer helpful and relatively easy-to-remember schemata for us to understand our world a little bit better.

Appropriateness of Grounded Theory for Your Research Project

To determine if grounded theory is a suitable approach for your research, keep in mind these questions:

- Is there truly no theory that adequately explains your phenomenon of interest? Often, there is a theory "out there," but you simply have not found it in the literature. You might need to seek out a librarian who will help you explore parallel fields with applicable theories.
- Is there a theory that partially explains your phenomenon of interest, but it needs to be updated or changed in some way?
- Have other researchers identified various constructs in your field but not tested them in relationship to one another?
- If the answer to any of the above questions is "yes," the next question is *Are you ready to take on a grounded theory study?* Before taking on a grounded theory study, you must understand that they require
  - A substantial commitment of time and energy, without which you may decide against a grounded theory
  - A lot of flexibility and a level of comfort with uncertainty because the focus of one's study is likely to change throughout the process
  - Intensive time in the abductive process. In practical terms, abduction is a continual process of collecting, transcribing, and analyzing data

to create hypotheses. During this process, you will usually need to make changes in your protocol and then go back out into the field to test those hypotheses with more data.

- A thorough literature review and a standardized proposal process, in which you defend your research plans and provide a detailed plan of analysis
- Institutional review board approval, with various contingency plans if your research interests evolve

In terms of specific questions in your research topic, grounded theory is an excellent method to increase understanding of meanings, contexts, or processes and to create an abstract explanation about the relatedness of particular constructs.

Action Plan for Using Grounded Theory

It is important to work with a mentor who understands grounded theory and will provide structured guidance, but who also gives you enough space to figure things out on your own. Glaser (2009) warned novice researchers of relying on mentors who do not understand grounded theory, holding that experienced researchers, in an effort to rescue the novice from confusion, "tend to block the novice researcher . . . by talking of his/her inexperience" and by "forcing their conclusions on the unformed, new novices" (p. 3).

Mentors should advise novice researchers that confusion and being overwhelmed are parts of the grounded theory process, "which are to be tolerated for a short while" (p. 4). It is essential to work with an advisor/chair who understands grounded theory and who will provide adequate support and flexibility. Conducting a grounded theory study is often overwhelming, but with a knowledgeable chair and an eager and open mind, you can employ this method with great success.

You might think that if you are going to do a grounded theory, you will not need to read the literature because you do not want previous findings to bias you. Yet, you will need to demonstrate to your readers (i.e., your research committee members) that you have sufficient justification to employ grounded theory. Otherwise, they will be less likely to approve of your work. Plus, you might be setting yourself up for failure because there could be an existing theory "out there" in the literature that sufficiently explains your phenomenon of interest (and if so, you may choose a quantitative approach that is deductive and proposed to test the validity of that theory).

There is a solution to guard against these pitfalls: Read, read, read. By reading, and critiquing what you read, you will ensure your own understanding of the existing literature and how it might justify a grounded theory. This kind of intensive reading will also provide you with ideas on how to write a proposal for a grounded theory.

Here is a second tip for novice researchers: Write, write, write. Writing a grounded theory is an ambiguous process because it involves more than mere reporting. Writing, like analyses, is emergent in grounded theory. Memo writing is very important in grounded theories because it facilitates reflection, understanding, and analytical insight of emerging data and evolving research study. Memo writing should be a frequent practice while conducting a grounded theory study. These memos should be spontaneous and not overly mechanical, with the purpose to "let your mind rove freely in, around, and from a category" (Charmaz, 2006a, p. 81).

A third tip is to join a network of grounded theory researchers. The Grounded Theory Institute (http://www.groundedtheory.com/) is one example of a network that offers access to an online forum with grounded theory researchers from all over the world. The site also offers a helpful publication and access to various seminars. The site specifically gears itself toward novice grounded theory researchers, many of whom are working on theses or dissertations.

Summary

Grounded theory is a design that researchers use to develop theory inductively from data. Grounded theory differs from other designs in that its end goal is to generate theory rather than to test existing theories or generate only themes. Grounded theory also uses an abductive process of data collection and analysis, in which you continually collect data, generate hypotheses from the data, and then test the hypotheses on more data.

Any discipline can use grounded theory. Topics should aim to increase understanding of meaning, context, or process with questions that are open-ended. Conceptual frameworks should guide—but not dictate—the focus of the grounded theory inquiry and observations.

The role of the researcher is to collect and interpret the data. This process requires adherence to the scientific method, but it also requires abundant creativity. Procedures in grounded theory force the researcher to break through assumptions and create new order out of the old.

All kinds of data can be the building blocks of a grounded theory: field notes, interviews, audio recordings, observations, and information in records and reports. Data collection procedures may change throughout the course of the study, as the grounded theory is refined and sharpened in focus. In addition to data, memos are particularly important to a grounded theory study because they stimulate thinking and provide the researcher with opportunities to focus on emergent categories. Data analysis involves becoming familiar with your data, scanning the data to identify codes, creating categories to organize the codes, developing hypotheses, and then collecting more data and refining hypotheses.

Anyone proposing a grounded theory for his or her thesis or dissertation must conduct a thorough literature review. The key is to use your literature review without letting it stifle your creativity.

You can ensure quality in grounded theory studies by fostering complete transparency in the data collection process. A well-constructed grounded theory will meet four central criteria for judging its quality: fit, understanding, generality, and control. Grounded theories should also focus on trustworthiness and its indicators, which are credibility, transferability, dependability, and confirmability.

KEY SOURCES

Must-Read Book on Grounded Theory

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Thousand Oaks, CA: Sage.

Classic Books on Grounded Theory

Glaser, B. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory*. Mill Valley, CA: Sociology Press.

Glaser, B., & Strauss, A. (1965). *Awareness of dying*. Hawthorne, NY: Aldine.

Glaser, B., & Strauss, A. (1967). *Discovery of grounded theory: Strategies for qualitative research*. Hawthorne, NY: Aldine.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Thousand Oaks, CA: Sage.

Other Books on Grounded Theory

Anfara, V. A., & Mertz, N. T. (2006). *Theoretical frameworks in qualitative research*. Thousand Oaks, CA: Sage.

Geertz, C. (1973). *The interpretation of cultures: Selected essays*. New York, NY: Basic Books.

Glaser, B. (1998). *Doing grounded theory: Issues and discussions*. Mill Valley, CA: Sociology Press.

Maxwell, J. A. (2005). *Qualitative research design: An interactive approach* (2nd ed.). Thousand Oaks, CA: Sage.

Rubin, H. J., & Rubin, I. S. (2005). *Qualitative interviewing: The art of hearing data*. Thousand Oaks, CA: Sage.

Wolcott, H. F. (1995). *The art of fieldwork*. Walnut Creek, CA: AltaMira Press.

Articles on Grounded Theory

Gasson, S. (2004). Rigor in grounded theory research: An interpretive perspective on generating theory from qualitative studies. In M. E. Whitman & A. B. Woszcynski (Eds.), *The handbook of information systems research* (pp. 79–102). Hershey, PA: IGI Global.

Glaser, B. (2009). The novice GT researcher. *The Grounded Theory Review*, *8*(2), 1–21.

Goetz, J. P., & LeCompte, M. D. (1981). Ethnographic research and the problem of data reduction. *Anthropology and Education Quarterly*, *12*, 51–70.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, *18*(1), 59–82. doi: 10.1177/1525822X05279903

Johnson, T. (2008). Doing quantitative grounded theory: A review. *The Grounded Theory Review*, *7*(3), 79–82.

Smaling, A. (2002). The argumentative quality of the qualitative research report. *International Journal of Qualitative Methods*, *1*(3), 62–69. Retrieved from http://ejournals.library.ualberta.ca/index.php/IJQM/article/viewFile/4600/3753

Website on Grounded Theory

Grounded Theory Institute. http://www.groundedtheory.com/

REFERENCES

Adjognon, A. (2014). A grounded theory of political intelligentizing in business administration. *Grounded Theory Review*, *2*(13), 20–29.

Braddy, P. W., Meade, A. W., & Kroustalis, C. M. (2006). Organizational recruitment website effects on viewers' perceptions of organizational culture. *Journal of Business and Psychology*, *20*(4), 525–543.

Brinchmann, B. S., & Sollie, H. (2014). Struggling with and for: A grounded theory of parents managing life with hard-to-treat ADHD teenagers. *Grounded Theory Review*, *1*(13), 3–15.

Carter, R., & Dresner, M. (2001). Purchasing's role in environmental management: Cross-functional development of grounded theory. *Journal of Supply Chain Management*, *37*(3), 12–27. doi:10.1111/j.1745-493X.2001.tb00102.x

Corlett, J. (2000). The perceptions of nurse teachers, student nurses, and preceptors of the theory-practice gap in nurse education. *Nurse Education Today*, *20*(6), 499–515. doi:10.1054/nedt.1999.0414

Charmaz, K. (2006a). *Constructing grounded theory: A practical guide through qualitative analysis*. Thousand Oaks, CA: Sage.

Charmaz, K. (2006b). Measuring pursuits, making self: Meaning construction in chronic illness. *International Journal of Qualitative Studies on Health and Well-Being*, *1*(1), 27–37. doi:10.1080/17482620500534488

Creswell, J. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage.

Ekström, H., Esseveld, J., & Hovelius, B. (2005). Keeping my ways of being: Middle-aged women dealing with the passage through menopause. *The Grounded Theory Review*, *5*(1), 21–53.

Ellingson, L. L. (2002). The roles of companions in geriatric patient-interdisciplinary oncology team interactions. *Journal of Aging Studies*, *16*(4), 361–382. doi:10.1016/S0890-4065(02)0071-3

Fenwick, J., Barclay, L., & Schmied, V. (2001). Struggling to mother: A consequence of inhibitive nursing interactions in the neonatal nursery. *Journal of Perinatal and Neonatal Nursing*, *15*(2), 49–64.

Freeman, R., Ekins, R., & Oliver, M. (2005). Doing best for children: An emerging grounded theory of parents' policing strategies to regulate between meal snacking. *The Grounded Theory Review*, *4*(3), 59–80.

Geertz, C. (1973). *The interpretation of cultures: Selected essays*. New York, NY: Basic Books.

Glaser, B. (1998). *Doing grounded theory: Issues and discussions*. Mill Valley, CA: Sociology Press.

Glaser, B. (2009). The novice GT researcher. *The Grounded Theory Review*, *8*(2), 1–21.

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

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, *18*(1), 59–82. doi: 10.1177/1525822X05279903

Johnston, T. L. (2008). Pushing for privileged passage: A grounded theory of guardians to middle level mathematics students. *The Grounded Theory Review*, *7*(1), 43–60.

Laditka, S. B., Laditka, J. N., Xirasagar, S., Cornman, C. B., Davis, C. B., & Richter, J. V. E. (2008). Providing shelter to nursing home evacuees in disasters: Lessons from Hurricane Katrina. *American Journal of Public Health*, *98*(7), 1288–1293. doi:10.2105/AJPH.2006.107748

Landmark, B. T., Strandmark, M., & Wahl, A. K. (2001). Living with newly diagnosed breast cancer—The meaning of existential issues: A qualitative study of 10 women with newly diagnosed breast cancer, based on grounded theory. *Cancer Nursing*, *24*(3), 220–226.

Nathaniel, A. K. (2004). A grounded theory of moral reckoning in nursing. *The Grounded Theory Review*, *4*(1), 43–58.

Parry, K. W. (1998). Grounded theory and social process: A new direction for leadership research. *The Leadership Quarterly*, *9*(1), 85–105. doi:10.1016/S1048-9843(98)90043-1

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Thousand Oaks, CA: Sage.

White, J., & Weatherall, A. (2000). A grounded theory analysis of older adults and information technology. *Educational Gerontology*, *26*(4), 371–386. doi:10.1080/036012700407857

12

Phenomenology

Paula Dawidowicz, PhD

Introduction

This chapter provides a detailed overview of an approach that is both a method and a methodology and that has a rich philosophical tradition. Fittingly, the chapter begins by discussing the origins and philosophical underpinnings of phenomenology. Examples highlight the use of phenomenology across a variety of disciplines and the research questions that are most appropriate. Issues related to sample size, the role of the researcher (which is rather unique to phenomenological data collection and analysis), and issues related to quality round out the chapter.

Background

Phenomenology is a collection and analysis of people's perceptions related to a specific, definable phenomenon. Other phrases describing *perception* include the lived experience, how people relate to a phenomenon, how people understand a phenomenon, and the meaning people give a phenomenon. A *phenomenon* is an event, an experience, or something that happens to someone. A phenomenon is something that is finite and definable rather than nebulous or unclear. Some examples of a phenomenon include giving birth, losing a first tooth, or failing or passing a test.

As a research design, phenomenology can help you understand the human factors involved in an experience. It answers the question of how people perceive a phenomenon in order to place it within a context. It can also help you clearly understand human perceptions. Whereas other methods only give you a partial view, because meaning is filtered through your own interpretation of information, phenomenology focuses directly on people's experiences with perceptions of a phenomenon. Phenomenology can help you understand why people react a specific way to an event or experience because people respond to situations based on their perceptions. Phenomenology also allows you to examine how transferable responses to experiences are from one participant to someone else. For example, if a person attends a protest rally on baby seals or abortion, the person will perceive that experience in a certain way. Phenomenology can help answer the question "How transferable are those experiences to others with similar backgrounds or in similar circumstances?"

Origins of Phenomenology

Phenomenology originated as a philosophical movement founded by Edmund Husserl. As a philosophy, Husserl's phenomenology focuses on the detailed description of consciousness as experienced from the first-person point of view (Husserl, 1962). Central to Husserl's philosophy is the concept of *epoché*, which means *freedom from suppositions* (Welton, 1999).

Epoché centers around the idea that people cannot feel that they know things without reflecting on them and that only what people think about things gives them meaning. Husserl felt that the perception of the experience itself is the source of knowledge on that experience and that obtaining knowledge is a matter of getting input from people who have experienced a phenomenon directly. Applying epoché to studies requires researchers to recognize their own biases, recognize the impact those biases have on their analysis of data, and purposely set those biases aside.

Husserl's phenomenology began as a pure reporting process without analysis. Husserl wanted to capture how people perceived events or experiences, what they thought of such phenomena, and how they lived through those events or experiences. Over time, and with the help of others like Amedeo Giorgi (1985), Husserl's philosophy evolved into a variety of approaches. Today, these approaches are far from unified. However, there are two main traditions that we commonly see in practice: descriptive (or transcendental) and interpretive phenomenology.

Descriptive (transcendental) phenomenology looks at how to transcend individual experience by reducing the reported experiences from individuals into patterns and themes so you can understand, or transcend, the data to find the commonalities people shared about that phenomenon. Individuals such as Clark Moustakas (1994) and Max van Manen (1997) honed descriptive phenomenology.

Interpretive phenomenology looks at the same kinds of experiences and collects the same data as descriptive phenomenology, but instead of reducing the reported experiences from individuals into patterns and themes, the researcher looks for the psychological or sociological factors that influenced the response. Interpretive phenomenology stems from Martin Heidegger's philosophy (Heidegger & Krell, 2008).

Both schools of phenomenology share the same philosophical underpinnings of humanism and constructivism; both are qualitative approaches to data collection and interpretation. Both schools of phenomenology align with qualitative research because phenomenology is based on the idea that individual perceptions guide actions and responses. Thus, there can be no absolute reality as a result of the analysis, because people understand actions only in terms of the way they perceive them. Data collection consists only of the participants' own words about the phenomenon (see Box 12.1).

Research Topics for Phenomenology

Not all research topics are appropriate for phenomenology. The goal of phenomenology is to understand human interaction with a phenomenon. Therefore, the best topics for phenomenological research are those with

BOX 12.1

Discipline-Specific Applications of Phenomenology

All disciplines can use phenomenology for specific purposes. For example, phenomenology can be used in

- Psychology, to see how parents experienced the loss of a child from sudden infant death syndrome (SIDS) and why they experienced it the way they did (DeCanio, 2000)
- Nursing, to ask how patients experienced nursing care and to look at which parts of that experience resonate for the patient (Huguelet, 2014)
- Epidemiology, to find out how people experience the process of using a specific drug (Ghaffari-Nejad, Ziaadini, Saffari-Zadeha, Kheradmand, & Pouya, 2014)
- Public policy, to determine how people experience the physical process of going to vote in local elections (Renshon, 1975)
- Education, to understand how people experience teaching second language writing (Tsui & Ng, 2010)

questions that consider how and why people do what they do or how they feel or interact with a phenomenon.

Examples of research topics that have been explored using phenomenology include

- The impact of personal expectations on the experience of being married and divorced (Friedman & Friedlander, 2005)
- How parents/guardians and special education students perceive the transition to life after high school (Stein, 2012)
- The experiences of nursing program graduates as they progress from being practical to registered nurses (Rice, 2011)
- How clinicians view their own competency in treating transgender individuals (Johnson & Federman, 2013)

When formulating a research question for a phenomenological study, be sure you can ask a question that can be answered. Suppose a research question asked, “What is the best method of training nurses?” Phenomenology cannot allow you to draw a valid conclusion to such a question because people will be reporting on instructional methods based on their perceptions of which method created their best efficacy. However, efficacy and actual skill development do not always align, so participants’ perceptions cannot provide the answers you need to the level of accuracy another method could, such as one using a measured assessment. A better question for phenomenology would be “How do nurses perceive the quality of their training?”

Also, be sure you interview the right people. Suppose a research question asked, “Why do patients respond better to nurses who explain everything than to nurses who do not?" If you interview nurses for your study, you will not be able to draw conclusions, because the nurses do not have the perspective to be able to explain why patients respond the way they do. You need to interview the patients to uncover why patients respond as they do to different nurses, essentially going to the source rather than going to a group of participants who can only provide their potentially faulty interpretations of patient experiences and perceptions as data to gain your answers.

Can Phenomenology Be Used With Other Methodologies?

Phenomenology can be embedded in the design of another research method, such as a case study or ethnography. A phenomenological case study looks at how different groups of people experience a phenomenon in a specific case bounded by time and location. For example, the case study might examine how people experience the campaigning process for an election in a certain town. To include phenomenology, the researcher would interview campaigners, politicians, and voters to get their perspectives. The researcher would then put the data together to draw conclusions.

A phenomenological ethnography could look at how people experience the orientation process in a certain corporate culture. The researcher could interview recruiters, new hires, people who did not get hired, and managers. They could be asked what cultural expectations there are surrounding the orientation process and how they experienced it.

Although you can embed phenomenology with another qualitative design, you should not combine it with a quantitative design to create a mixed methods study. Doing so affects the integrity of the phenomenology and may even force you to conduct two separate studies instead of one cohesive study.

Reminders About Theoretical and Conceptual Frameworks

In designing a phenomenological study, remember that humanistic concepts inform phenomenology—people's perceptions both reflect and affect how people feel about a phenomenon. Unlike theoretical frameworks, which are designed to test a hypothesis or theory by determining a yes/no response about the truth of a hypothesis, you use a conceptual framework. You develop this framework from a combination of current research and, at times, a theory or two to create a lens or perspective—an angle or approach—to investigate the phenomenon. Some possible frameworks might include critical theory or critical race theory (giving voice to groups that are not normally heard) or a combination of key current research on the phenomenon that indicates why and how you believe the research questions (or interview questions used) are important.

Also, remember that phenomenology requires gathering people's perceptions based on their own words. Researchers should interpret what the participant states in interviews, rather than interpreting others' accounts of the participants' thoughts or superimposing their own perspectives on the participants' statements when interpreting them. This is what is expected in phenomenological research.

Keep in mind that people do not always consciously understand why they perceive a phenomenon the way they do. Sometimes, what they perceive is not actually what is happening. People's perceptions can change, so their conscious perceptions can also change. Also, as people talk about their perceptions, unconscious perceptions emerge. This evolution of one's perceptions is similar to the way that, in ethnography, being observed potentially changes how a participant behaves. As you will see later, validation strategies, such as triangulation, are important ways to address this issue.

Sample Size for Phenomenology

Phenomenology is about getting the depth, not the breadth, of people's perceptions. In determining an appropriate sample size, the goal should be to obtain enough data appropriate to the study, from a sufficient number and variety of individuals. Most often, participant numbers are between 5 and 15. Often, the sample size is between 8 and 12.

Here are some examples of sample size from contemporary studies:

- In a study on how cooperative education could grow talent among South Africans, Groenewald (2004) used a sample size of 10 managers, 5 at educational programs and 5 at collaborating enterprises.
- In a study on the experiences of people with HIV/AIDS in China, Zhou (2010) used a sample size of 21 adults with HIV, 15 men and 6 women, ranging in age from 21 to 46 years.
- In a study on women's experiences of prostitution and substance abuse, Sallmann (2010) used a sample size of 14 women, who identified as White, Black, or biracial and ranged in age from 19 to 48 years.
- In a study on how older women perceive their body/self from the perspective of food and eating, McCormick (2011) used a sample size of 15 women, age 62 to 91 years, who were all members of a large senior center in a rent-subsidized apartment community.

Phenomenology is not meant to result in generalizations. The value of the results of a phenomenological study lies in their transferability (i.e., the ability to apply the learning to another situation). Using larger numbers will not necessarily produce more insight into the experience of a phenomenon. Moreover, analyzing the data for larger numbers of participants is impractical if you consider that an hour-long interview may yield a transcript of 20 to 30 pages. As a result, the appropriate number of participants is smaller than in some other types of qualitative studies.

The Role of the Researcher

In phenomenology, the role of the researcher is to gather, organize, and analyze perceptions from people who have experienced a phenomenon. Avoiding bias during both the data collection and analysis processes is very important. Participants can change their answers based on their perceptions of your bias, and the integrity of their responses can be lost. To avoid researcher bias, do not lead the participants—implying by your questions, facial expressions, or gestures—the perspective you want them to share. Do not push participants to continue answering questions they do not want to answer, share your own stories with the participants, or share participants' stories with other participants. If you have assistants who collect the data, you should carefully train them so that they also do not ask questions in a biased manner.

You should audio record the interviews and use a researcher's journal to record your reflections, ideas, and thoughts about possible connections among data and participants. Also be sure to conduct an ongoing review of responses by the participants, including a review of the researcher's conclusions. Make sure you understand the process of *bracketing*. In qualitative research, the researcher brackets his or her personal experiences, biases, and perceptions prior to conducting research in order to allow the participants to express their own experiences without being influenced by the researcher. Researchers also bracket their experiences during data collection so that the researcher can analyze his or her perceptions separately from those of the participant. In phenomenology, bracketing is particularly important so that the participants' perceptions of a phenomenon remain intact. You also record bracketing in your researcher journal as any potentially biased comments that occur to you during the interview; such comments are recorded so that potential bias during the data analysis process can be considered. For example, if a participant responds with an idea that is the opposite of your beliefs, it is important to note that so that this clash in beliefs will not result in discounting the participant's thoughts before considering their importance and implications to your study.

Data Collection for Phenomenology

Phenomenology requires direct responses from participants, so it has a limited range of data collection sources. Appropriate ways to collect data include interviews, focus groups, journals, open-ended questionnaires, or other similar products that capture individuals' responses to a phenomenon in their own words. Researchers may use the Internet and other technologies to collect data. For example, researchers can conduct interviews or do focus groups via online video-conferencing services such as Skype.

Most often, phenomenological researchers collect data by using interviews, although the length of the interviews may vary depending on the study's purpose. For example, McCormick (2011) interviewed each woman in her study four times for 1 to 1.5 hours per interview. Sallmann (2010) interviewed each of her participants for 45 minutes to 3 hours. Zhou (2010) interviewed each of the adults in her sample for 2 to 3 hours. The focus of phenomenology is to delve deeply into a phenomenon rather than gather shallow amounts of data from larger numbers of people; thus, the time spent with each person reflects a process of gaining that depth. The time required can vary depending on people's readiness to provide that depth and the speed with which they provide it in the interviews.

There are multiple data collection methods available, such as interviews, journals, questionnaires, and focus groups. Whatever combination of collection methods you use, ensure the same methods are used for each participant. If you interview a participant and give the person a questionnaire, you must interview all your other participants using the same questions and administer the same questionnaire to all participants. As an example, Groenewald (2004) interviewed each of his participants and asked them to write essays about their viewpoints; this was done in the same way for all study participants.

The researcher must develop a data collection protocol (a set of questions for prompting responses) in order to ensure complete and thorough data collection. The protocols for different types of data such as interviews, focus groups, and journals will be different, but they are basically a set of core open-ended questions or prompts to which people can respond. For example, Sallmann (2010) used prompts such as "Tell me more about that," "What was that like for you?" and "Please give me an example." Prompts can also be questions specific to the topic in discussion. The frequency of prompts during an interview will naturally vary, depending on how forthcoming and communicative each participant is.

How to Analyze Phenomenological Data

Once you collect data, how do you analyze it? Generally, data analysis involves reviewing the collected data, identifying themes, and synthesizing the results. Techniques will vary based on the type of phenomenological study and its purpose. For example, McCormick (2011) transcribed her four recorded interviews of about 1.5 hours each with 15 women, which yielded about 3,000 pages of transcripts. She assembled the transcripts into a book for each participant. McCormick read the books several times, listened to the recorded interviews without reading the transcripts, and then listened to the interviews while reading the transcripts. She analyzed each book separately. Then, she analyzed the interviews across books. Her goal was to "allow the full array of themes to emerge." Rather than "fragment the transcripts through line-by-line coding," she read the books as whole texts, to allow themes to emerge "continuously, unexpectedly." She used the themes to develop "meaning-clusters" which formed the basis for her description of the phenomenon (McCormick, 2011, pp. 76–77).

Sallmann (2010) applied a hermeneutic analysis. She used a research team to read the interview transcripts for her interviews with her 14 participants, write summaries of them, and begin identifying themes. The team shared the summaries to reach a consensus about the themes and reviewed theoretical literature to clarify the themes. Sallmann then reread the transcripts to examine the themes and produce new themes. She drafted the research manuscript, which she shared with the research team, other reviewers, and some of the study participants. She then integrated their feedback into the final manuscript.

Groenewald (2004) used a process of *explication* rather than *analysis*. Whereas analysis involves breaking data into parts, explication allows examining data in its whole context. Groenewald's explication began with a conscious bracketing of the self—a self-reflection and written record of his biases and perspectives so that he could consciously attempt to exclude them from his examination of the data. He then listened repeatedly to the audio interviews to get a sense of each participant's perspective as a whole, extracted units of meaning from each interview, and clustered them to form themes. He integrated these with the essays each participant provided. He then summarized and validated each interview and essay and wrote a composite summary reflecting the common and individual themes.

In general, descriptive phenomenological studies categorize participants' responses and look at patterns and themes. Interpretive phenomenological studies, especially those conducted by psychologists, examine what people state about their perspectives and then consider what in their background might explain why they have those perspectives.

Data analysis techniques for phenomenology are the same as those used in other qualitative research methods. Because the focus is on sharing participants' perspectives, an *emic* focus is used. In an *emic* focus, you present the perspectives of your participants rather than your own, letting their voices be heard, progressively honing and refining the focus of your conclusions (Schutt, 2011, p. 322). As data collection begins, so does your data analysis. Such ongoing analysis allows you to evolve better and more detailed follow-up probes so you can get data from each participant that will best answer your research questions.

You can code data using multiple strategies, all designed to identify themes and patterns existing within the data. At times, you might find it useful to establish predetermined codes to initially sort your data based on potential participant responses identified during your literature review. At other times, you might use *open coding*, where you identify conceptual categories for the data as you begin your analysis. Open coding often leads to *axial coding*, where data previously clustered through open coding are examined more closely, separated, and coded again based on common characteristics, contexts, or conditions to provide clearer insights into the meanings of the data (Flick, 2013). Finally, you may use *selective coding*, in which you reintegrate the data to identify the underlying themes and the patterns that answer your research questions (Khiat, 2010). Be aware, though, that pulling lines or phrases out to code can result in the miscoding of data if you lose the context or meaning of that line or phrase. For example, two people can say, "Wow, what a long day!" to describe two very different types of days—one long because it included many rough tasks and one long because it included a lot of strenuous entertaining activities. Removing lines from context or searching for words or phrases without reviewing the context of those words or phrases for meaning can result in biased or incorrect results and study conclusions (see Figure 12.1).

Quality in Phenomenological Research

In a phenomenological study, researchers look for trustworthiness. Threats to trustworthiness include too shallow a view of participants' experiences with the phenomenon, bias in interpretation, leading participants' responses, not collecting enough data, or reading into the data rather than letting the data speak for itself.

To ensure quality in a phenomenological study, be sure to record interviews and collect enough data. For example, interview participants several times in case their perspectives change from one session to the next. Let each session begin with the option for them to summarize their previous answers. You should also triangulate sources, which involves using multiple data sources to build a complete picture of participants' perceptions of a phenomenon.

One method of triangulation can be member checks if they are conducted correctly. During member checks, participants review the transcripts of their

What perceptions do you have of your interactions with your mother and why?

I hate my mother. She is a demon, and I would swear I am living in hell on earth!

Don't like | Negative treatment/comments

I love the way my mother always encourages me. She is my inspiration.

Encourages/supports

I think my mother hates me. She always has something bad to say to me.

Don't like | Negative treatment/comments

I find my mother confusing. What is it she wants from me? I can never figure it out.

Confusing | Negative treatment/comments

Sometimes my mother confuses me—I don't know what she wants, and she doesn't get me. But, you know what, she is always on my side and supports whatever I want to do.

Confusing | Encourages/supports

She puts up with me and leaves me alone. I guess she's okay, and I get to do whatever I want.

Okay | Neglectful

I really don't like her. She leaves me alone, but I never feel like she cares what I do.

Don't like | Neglectful

Themes:

Not like—negative

Encouraging/supportive—inspires

Not like—negative

Confusing—negative

Confusing—encouraging/supportive

Okay—neglectful

Not like—neglectful

Note that the responses, if not considered in their entirety, can result in you concluding different, incorrect meanings when you develop themes from your data.

FIGURE 12.1 **Sample coding detail.**

interviews to ensure that their thoughts on the topic remain the same as they have previously stated. As participants reflect on their experiences and perspectives on a phenomenon throughout the data collection process, their perceptions about your interview prompts can change. Member checking is one way you can ensure you have their most accurate thoughts on the phenomenon you are researching. Finally, member checking of your finalized analysis can allow you an opportunity to gain final insights from participants on what you have concluded. The insights they share can clarify the meaning of the data you have collected and the accuracy of your conclusions can bring to light perspectives they share with you that you might not previously have considered. As a result of these discussions, the quality of your data analysis and related study conclusions can be improved.

To ensure quality, do not overgeneralize in your research. Remember, the nature of qualitative research is one of transferability, not generalizability. The general characteristics of the group being examined should be shared, although specific details about individuals within the group should remain hidden, including during member checking of your finalized analysis.

Appropriateness of Phenomenology for Your Research Project

If you are considering doing a phenomenological study, here are some questions to ask yourself:

- Are you examining a phenomenon as opposed to a bounded case?
- Is determining the lived experience of people in relation to a specific phenomenon going to directly answer your question or fill a gap in research or practice?
- Does the basic premise of phenomenology—that perspective directly impacts or is impacted by the nature of the phenomenon—fit your study?
- Will you be able to collect information from the appropriate participants?

If the answer is "Yes" to all of the above, then phenomenology may be an appropriate research method for you.

Summary

Phenomenology is the collection and analysis of people's perceptions of their lived experiences related to a specific, definable phenomenon. Phenomenology originated as a philosophical movement founded by Edmund Husserl and evolved into a variety of approaches. Two main traditions commonly in practice today include descriptive (or transcendental) and interpretive phenomenology.

Appropriate research topics for phenomenological research involve questions that consider how and why people do what they do or how they feel or interact with a phenomenon. Because the goal of phenomenology is getting depth, not breadth, on lived experiences, sample sizes are small, typically ranging from 5 to 15 participants. You as researcher must be careful to avoid bias when collecting data and must follow general ethical considerations that apply to human subject research.

Phenomenological data must capture individuals' responses to a phenomenon in their own words. Appropriate ways to collect data include interviews, focus groups, journals, and open-ended questionnaires. Data analysis involves reviewing the data, identifying themes and patterns, and synthesizing the results through a process that includes open, axial, and selective coding. Techniques will vary based on the type of phenomenological study and its purpose.

To ensure the quality of a phenomenological study, researchers should be sure to record interviews and collect enough data as well as use strategies such as triangulation and member checking.

KEY SOURCES

Giorgi, A. (2009). *The descriptive phenomenological method in psychology: A modified Husserlian approach*. Pittsburgh, PA: Duquesne University Press.

Moustakas, C. (1994). *Phenomenological research methods*. Thousand Oaks, CA: Sage.

Sokolowski, R. (1999). *Introduction to phenomenology*. Cambridge, United Kingdom: Cambridge University Press.

Vagle, M. D. (2014). *Crafting phenomenological research*. Walnut Creek, CA: Left Coast Press.

van Manen, M. (1997). *Researching lived experience: Human science for an action sensitive pedagogy*. London, Canada: Althouse Press.

REFERENCES

DeCanio, C. (2000). *A phenomenological study of perceived presence/spiritual encounters with SIDS parents: A definition based on philosophical perspective* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database.

Flick, U. (2013). *The Sage handbook of qualitative analysis*. Thousand Oaks, CA: Sage.

Friedman, M. L., & Friedlander, M. L. (2005). Toward an understanding of Jewish identity: A phenomenological study. *Journal of Counseling Psychology*, *52*(1), 77–83.

Ghaffari-Nejad, A., Ziaadini, H., Saffari-Zadeha, S., Kheradmand, A., & Pouya, F. (2014). A study of the phenomenology of psychosis induced by methamphetamine: A preliminary research. *Addiction & Health*, *6*(3–4), 105–111.

Giorgi, A. (1985). *Phenomenology and psychological research*. Pittsburg, PA: Duquesne University Press.

Groenewald, T. (2004). A phenomenological research design illustrated. *International Journal of Qualitative Methods*, *3*(1), 1–26.

Heidegger, M., & Krell, D. F. (Eds.). (2008). *Basic writings*. New York, NY: HarperCollins.

Huguelet, P. (2014). The contribution of existential phenomenology in the recovery-oriented care of patients with severe mental disorders. *Journal of Medicine and Philosophy*, *39*(4), 346–367.

Husserl, E. (1962). *Ideas*. New York, NY: MacMillan.

Johnson, L., & Federman, E. J. (2013). Pathways and barriers to care for LGBT veterans in the U.S. Department of Veterans Affairs (VA). *Journal of LGBT Issues in Counseling*, *7*(3), 218–235.

Khiat, H. (2010). A grounded theory approach: Conceptions of understanding in engineering mathematics learning. *The Qualitative Report*, *15*(6), 1459–1488.

McCormick, M. L. (2011). The lived body: The essential dimension in social work practice. *Qualitative Social Work*, *10*(1), 66–85.

Moustakas, C. (1994). *Phenomenological research methods*. Thousand Oaks, CA: Sage.

Renshon, S. A. (1975). Political learning and support for political participation: An examination of some students and their parents. *Nonprofit and Voluntary Sector Quarterly*, *4*(3–4), 200–218.

Rice, E. M. (2011). *Transitioning from practical to registered nurse: A phenomenological inquiry of graduates of a competency based nursing program* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database.

Sallmann, J. (2010). Living with stigma: Women's experiences of prostitution and substance use. *Affilia*, *25*(2), 146–159.

Schutt, R. K. (2011). *Investigating the social world: The process and practice of research*. Thousand Oaks, CA: Sage.

Stein, K. (2012). Experiences of selected emerging adults with emotional or behavioral difficulties in higher education. *Career Development and Transition for Exceptional Individuals*, *35*(3), 168–179.

Tsui, A. B. M., & Ng, M. M. Y. (2010). Cultural contexts and situated possibilities in the teaching of second language writing. *Journal of Teacher Education*, *61*(4), 364–375.

van Manen, M. (1997). *Researching lived experience: Human science for an action sensitive pedagogy*. London, Canada: Althouse Press.

Welton, D. (Ed.). (1999). *The essential Husserl: Basic writings in transcendental phenomenology*. Bloomington, IN: Indiana University Press.

Zhou, Y. R. (2010). The phenomenology of time: Lived experiences of people with HIV/AIDS in China. *Health*, *14*(3), 310–325.

13

Survey Research

Kimberley A. Cox, PhD

Introduction

This chapter presents an overview of survey research, including its origins, purpose, and methods. Common types of survey instruments and their respective advantages and disadvantages are provided, followed by a discussion of the types of information you can collect with surveys. Last, a brief overview of key considerations in designing a survey instrument, including the various types of survey items and response scales and reliability and validity of measurement, is presented.

Background

Survey research is a descriptive, nonexperimental method used to collect information, such as attitudes and behaviors, through interviews and/or questionnaires. Survey research methods can be used alongside other methods in quantitative, qualitative, and mixed methods research. Survey research can be used to explore, describe, and explain a multitude of constructs, characteristics, attitudes, and behaviors. For this reason, it is a popular method used in many disciplines within the social, behavioral, management, and education disciplines.

Survey research is an appropriate method when you have access to a sample of individuals who are available to provide self-report information. It is also an appropriate method when you need to collect information on characteristics of people such as their education or age. Finally, survey research is an appropriate method if your study's research question points to survey research as the best method for obtaining the data needed to answer the question.

Origins of Survey Research

Contemporary survey research has been shaped by social and political events and influenced by various academic disciplines and organizations such as the National Opinion Research Center. As a tool for research, the survey was an outcome of efforts by the business and government sectors in the 20th century. For example, by the 1940s, federal agencies were using polls, such as the Gallup Poll, and surveys, such as the U.S. Census, to gather information about the public and to measure people's opinions about national problems (Converse, 1987). It was not until the 1960s that survey research firmly established itself in academia. Today, numerous disciplines, such as psychology, political science, public policy, public health, marketing, and education, make extensive use of surveys.

The influence of technology on survey research has a long history. In the 1930s, as the U.S. Postal Service expanded and telephones became commonplace, larger samples of the population could be surveyed (Converse, 1987). Later, with the emergence of the Internet in the 1980s and subsequent rapid technological advances in the 21st century, new modes of collecting survey data, such as web-based surveys, have become increasingly popular.

Purposes of Survey Research

Survey research generally serves one or more of three main purposes: exploration, explanation, and description (Martella, Nelson, & Marchand-Martella, 1999). First, a survey can be used to explore a topic that has not been previously examined. For example, Duan, Brown, and Keller (2010) conducted an exploratory study on the experiences of male faculty in academia. The researchers designed a questionnaire that was distributed by mail to collect information on male faculty members' perceptions of their academic career, their values toward family and work, and their experiences in meeting their roles and expectations.

Second, when a survey is used to explain the relationship between two or more variables of interest, its purpose is explanation (Martella et al., 1999). For example, Farc and Sagarin (2009) conducted a study to investigate the relationship between attitude strength and voter behavior in the 2004 U.S. presidential election. The researchers administered questionnaires to college students before and after the election to collect data on their voting behavior and attitudes toward the presidential candidates.

Third, when a survey is used to describe the characteristics or attributes of a population, its purpose is description (Martella et al., 1999). The goal in descriptive survey research is to describe a population rather than understand why a relationship may exist between variables. For example, Barusch and Wilby (2010) conducted a study to describe how older adults experience and cope with depression. The researchers used interviews to collect information on older adults' cognitive impairment, depressive symptomatology, and coping strategies and effectiveness.

Types of Survey Research Methods

A *survey* is a method of collecting data from and about people (Fink, 2009). A *survey instrument* is the tool used to gather data; this term is typically used to differentiate the tool from the survey research it supports. There are two general tools for collecting data through survey research: verbal surveys (or interviews) and written surveys (or questionnaires). A verbal survey or interview refers to a survey instrument that is administered face to face, over the telephone or Internet, or via a combination of these modes. A written survey or questionnaire refers to a survey instrument that contains items that the respondent is expected to read, with little or no direct contact with the researcher, and then report his or her own answers on the questionnaire. There are three main types of questionnaires: (a) self-administered mail questionnaires, where respondents provide their answers to the survey items and return the questionnaire by mail in a provided self-addressed stamped envelope; (b) in-person individual or group-administered questionnaires, which are administered to a single respondent or group of respondents, respectively; and (c) self-administered web-based questionnaires, which are administered through a website or a commercial web-based survey host.

Advantages and Disadvantages of Survey Research Methods

Each type of method has advantages and disadvantages. You must weigh the costs and benefits of each type with consideration of your study's purpose, research question, sampling method, and data analysis plans.

Interviews. Interviews generally yield quality responses from respondents and high response rates because the interviewer is present. The interviewer can clarify survey items and answer questions during the interview. However, interviews can be costly, and they typically require more time, especially if multiple interviewers need to be recruited and trained. Interviews may pose a risk for interviewer bias, and they may offer less privacy than other types of surveys, which may affect respondents' willingness to answer questions truthfully.

Self-Administered Questionnaires. Self-administered questionnaires are typically an inexpensive and quick mode of reaching a large sample. Self-administered questionnaires may offer more privacy than interviews, if they are anonymous. As a result, respondents may be more willing to answer sensitive questions. Because there is no interviewer present, there is minimal concern about interviewer bias or coercion. However, without the researcher present, self-administered questionnaires generally yield lower response rates than interviews. To improve the response rate, self-administered questionnaires should be easily understood and clear because there is no interviewer present to answer questions or provide clarification on survey items. Respondents may intentionally or unintentionally skip survey items, and the researcher does not have any control over these omissions. Questionnaires delivered by postal mail may cost more and require more time to collect data than those administered by telephone or in person.

Self-Administered Web-Based Questionnaires. Self-administered web-based questionnaires may be faster and less costly to administer than other types of surveys. They also tend to provide respondents with a greater sense of privacy, if personal identifying information, such as one's name, is not collected. The data collected through many commercial web-based survey hosts can be exported directly into SPSS or Microsoft Excel programs, thus eliminating time devoted to manual data entry. Although many people have access to the Internet and email, those without access will not be able to participate in a web-based survey. People also possess varying degrees of technical skill that may hinder their ability to complete a web-based survey.

Therefore, the targeted population must have access to and competence in using the Internet and email.

Web-based questionnaires have unique ethical and security concerns compared to other types of surveys, including, but not limited to, issues around the storage of data and the possibility of a breach of respondents' data (Nosek, Banaji, & Greenwald, 2002). You will need to approach these issues and the more general concerns of confidentiality and anonymity common to all research methods in light of these special circumstances.

Types of Information Researchers Collect With Surveys

In survey research, you gather information from participants from a questionnaire or interview. Your role varies depending on how the survey is administered. For example, in a self-administered mail or web-based questionnaire, you have little to no direct contact with respondents. Your only contact with respondents is likely to occur through an introductory email or letter describing the study and instructions for questionnaire completion. In survey research conducted through interviews by telephone, on the Internet (e.g., Skype), or in person, you have direct contact with respondents as you ask questions or present statements and the respondents provide their answers. It is this characteristic of engagement with participants that distinguishes survey research from other kinds of descriptive research, such as observational research, which is characterized by observations of individuals in their natural environment without researcher interference. Common to all modes of survey administration is your responsibility to conduct ethical research.

The purpose of a survey determines the type of information you collect, and the type of information should be aligned with your study's research question and data analysis plan. For example, will groups be compared? Will a relationship between variables be examined? Will the frequency of behaviors be described? Alignment in this regard is critical for collecting the type of information that will produce meaningful data that can be analyzed to answer your research question.

The three main types of information commonly collected with surveys are descriptive, behavioral, and attitudinal. Most surveys include a combination of these types of information. Surveys that collect descriptive information often ask questions to gather information that describes respondents' characteristics, such as education and age. This descriptive information is important in most survey research in that it will permit you to better understand the sample and make inferences about its representativeness to the population. For example, in Farc and Sagarin's (2009) study of university students' attitudes toward presidential candidates, the researchers asked respondents, "Which political party do you usually feel closest to?" to assess political affiliation (Barnes, Jennings, Ingelhart, & Farah [as cited in Farc & Sagarin, 2009]).

Survey researchers are also often interested in gathering information about respondents' attitudes toward various objects, events, things, or people. For example, in Duan and colleagues' (2010) study of male faculty members' perceptions of their academic careers, the researchers asked respondents, "What do you like about your career in academia?"

Survey researchers are also frequently interested in gathering information about respondents' behaviors such as how often one engages in a particular action or activity. Barusch and Wilby (2010), for example, gathered behavioral information on older adults' coping strategies through the use of an open-ended question that asked, "How do you cope with this [depressive symptom]?"

After you decide on a type of survey instrument and determine the information it will collect, your sampling methods should be considered. The term *sampling plan* is used in survey research to refer to the approach used to select the sample from a population (Levy & Lemeshow, 2009). In determining the sampling plan, you must decide whether to use a probability or nonprobability sampling approach. See Chapter 4 for an overview of some sampling strategies relevant to survey research.

BOX 13.1

Tips for Locating Existing Survey Instruments

To locate existing survey instruments, you should review testing reference books and online resources and databases such as the following:

- *FAQ/Finding Information About Psychological Tests* from the American Psychological Association (APA) Science Directorate: http://www.apa.org/science/programs/testing/find-tests.aspx
- *Directory of Unpublished Experimental Measures* published by APA
- *Mental Measurements Yearbook* published by the Buros Center for Testing: http://buros.org/mental-measurements-yearbook
- Online databases such as Health and Psychosocial Instruments (HAPI) and PsycTESTS
- RAND Health Surveys: http://www.rand.org/health/surveys_tools.html
- *Test Collection* from Education Testing Service: http://www.ets.org/test_link/
- *Tests in Print* published by the Buros Center for Testing: http://buros.org/tests-print
- *Test Reviews Online* from the Buros Center for Testing: https://marketplace.unl.edu/buros/

If you locate an existing survey instrument that is suitable for your study, the next step is to obtain permission to use it for your own research. Check the instrument for any copyright restrictions. You may need to contact the publisher or author of the instrument to request permission. Be sure to retain copies of any communications granting permission. Although there may be situations where it is acceptable to use an instrument that is published in a journal without permission, confirmation from the author is typically advised. Many journals hold the copyright for published tests or measures.

Note: The examples provided here are not exhaustive, but rather they are intended to provide suggestions for key sources that may be helpful in locating existing survey instruments. All online sources were current at the time of printing.

Benefits and Limitations of Survey Research

Survey research methods can help you gather information directly from people about how they act and what they think, know, and believe as well as their descriptive characteristics, which can be challenging to measure with other methods. The measurement of behaviors and attitudes, for example, which cannot be directly observed, is possible with survey research methods. Survey research methods are also useful for collecting information from a geographically dispersed sample of individuals.

Although surveys have numerous benefits, there are some limitations to their use. For example, surveys do not provide exact measurements; rather, they supply estimates of the true population under study (Salant & Dillman, 1994). Therefore, you should generally exercise caution in generalizing your findings to a larger population.

The response rate or rate of return on self-administered questionnaires is another potential limitation that can cast doubt on the validity of a study's results. A low response rate could signal that respondents who return the questionnaire are different in important ways from respondents who do not. A low response rate can also limit your ability to generalize findings to the targeted population.

The accuracy of survey data often rests on the ability of respondents to accurately interpret the meaning of survey items and correctly and honestly report their experiences and behaviors. Respondents may also have difficulty recalling information or may misinterpret a survey item. Therefore, surveys are susceptible to bias due to their self-report nature. Bias may also occur as a result of a respondent's intentional or unintentional misreporting of information.

Designing a Survey Instrument

Many phenomena have been studied to such an extent that there exist reliable and valid survey instruments for their measurement. Therefore, it is important that you conduct a thorough review of the literature and search testing reference books, online resources, and databases to determine whether a survey instrument is already available (see Box 13.1). Some circumstances may call for you to modify or use only certain items from an existing survey instrument. In this case, permission is typically needed from the instrument's author. Modifications of survey items may raise potential issues about the items and the instrument's reliability and validity because changes can affect the psychometric properties of an instrument.

You may need to design a survey instrument if one does not already exist to measure the construct or phenomenon of interest. Deciding whether or not to design a survey instrument involves consideration of several important steps, such as

- Determining the factors associated with the construct or phenomenon of interest
- Selecting a mode of administration
- Determining the sample to be studied
- Selecting a sampling plan and method

- Writing questions or statements
- Writing response scales
- Pilot testing
- Adjusting the survey instrument based on pilot testing results

Writing Reliable and Valid Survey Questions and Statements

Good survey questions and statements share some common characteristics: They are easy to understand, free of biased language, and presented one at a time. As a result, they yield responses that are valid and reliable measures of the construct or phenomenon of interest. Survey questions and statements should also be written with knowledge of your plans for data analysis, which begins with identifying the most appropriate scale of measurement—nominal, ordinal, interval, or ratio—for the type of data to be collected according to the statistical analysis that will be performed.

Types of Survey Questions. There are two main types of survey questions: open-ended and closed-ended. Open-ended questions are survey items that ask respondents to provide an answer, such as small amount of text or a number, in their own words; that is, response options are not provided. For example, Duan et al. (2010) asked male faculty, “What do you like about your career in academia?” and “What do you dislike about your career in academia?” Open-ended questions such as these are most appropriate when (a) you are exploring a new topic of study, (b) the list of potential response options is too lengthy for a closed-ended format, and (c) the aim of the study is to obtain diverse responses. Open-ended questions, such as the examples given earlier, require more cognitive thought for respondents to answer (Tourangeau, Rips, & Ransinki, 2000). Content analysis is typically required for analyzing responses from open-ended questions, which can be time-consuming and complex because it involves determining coding categories and then categorizing responses into themes.

Closed-ended questions are survey items presented with a specified set of response options. These response options can be presented in a dichotomous, multiple-choice, ranking, or rating format. Closed-ended questions require that the response options include all of the possible answers that could be expected from respondents. Researchers often include an “Other” category as a response option to capture a response that is not listed, but doing so means that such responses will be analyzed similar to that of an open-ended question.

The overall survey instrument should be visually appealing and professional. It is best to keep color, font size, and font type simple. Response options are typically formatted to either the right of the question or statement or directly below it. Response options for rating scale questions are generally displayed horizontally, whereas response options for multiple-choice and ranking scale questions are generally listed vertically.

Some common errors in writing survey questions and statements include

- *Ambiguity*, which refers to questions and statements that are vaguely written. Ambiguous questions and statements lead to confusion because respondents have to try to make sense of the researcher’s meaning of a survey item.

- *Biased wording*, which refers to words within a question or statement that may trigger an emotional reaction. Biased words may influence a respondent's answer in a way that is not intended by the survey item.
- *Leading questions and statements*, which refer to words that lead a respondent to favor or disfavor a particular perspective. For example, a question beginning "Don't you disagree that . . . ?" may lead a respondent to expect that a particular response is desired.
- *Double-barreled questions and statements*, which refer to questions and statements that ask for responses to two or more things/constructs within the same item. For example, the statement "I think it is important to recycle and conserve water" reflects two different topics—recycling and conserving water. A respondent may wish to answer one way about one of the topics and differently for the other topic.
- *Complicated skip patterns*, which refer to instructions that direct respondents to skip questions or statements that do not apply and proceed to items elsewhere within the survey. If the skip patterns are overly complex, it may be difficult for respondents to follow them.

Some tips to avoid these common errors when writing survey questions and statements include the following:

- Provide clear instructions.
- Write questions and statements that are as simple and as short as possible.
- Write in complete sentences; for example, write "What is your gender?" and not "Gender?"
- Avoid using biased or leading words that may cause an emotional response or lead a respondent to favor or disfavor a particular perspective.
- Avoid using double negatives.
- Avoid using "and" in questions and statements to ensure that only one topic is being measured in each item.
- Avoid complicated skip patterns that can be difficult for respondents to follow.
- Pretest questions and statements for clarity by asking others to read the items aloud and record any issues that arise to inform potential revision.

Another consideration to keep in mind when writing survey questions and statements, especially those that are sensitive in nature, is response bias. One type of response bias is *social desirability*, which refers to the tendency for people to make themselves appear positively or socially desirable by responding to items in a certain way. One way to reduce social desirability is to design the survey such that it ensures anonymity and to communicate to participants the privacy protections you have put in place. Another type of response bias is the tendency for people to only provide extreme answers on a scale. In this case, formatting items so that some are worded for disagreement and some are worded for agreement may help reduce bias.

Types of Scales

Rating scales are used in survey research to quantify a respondent's answer to a survey item. There are two main types of rating scales: Likert and semantic differential. A Likert scale presents a range of response options along a continuum, such as *strongly disagree* to *strongly agree*. Respondents are asked to indicate their agreement or disagreement to a statement by selecting a numerical value. For example:

1 = *Strongly disagree*; 2 = *Disagree*; 3 = *Neither agree nor disagree*; 4 = *Agree*; 5 = *Strongly agree*

A semantic differential scale presents response items with bipolar adjectives following a word, such as in the following item: Good _._._._._ Bad. Respondents place a mark on the continuum between the set of adjectives that best represents the meaning they associate with the word presented. Farc and Sagarin (2009) used a semantic differential scale as one measure of college students' attitudes toward the two candidates in the 2004 presidential election. The participants were presented with the candidates' names followed by response items, such as "Positive–Negative" and "Wise–Foolish."

Pilot Testing a Survey Instrument

A pilot test is the final step in designing a survey instrument (Martella et al., 1999). It allows you to gain preliminary information about how a survey instrument performs under realistic conditions (Martella et al., 1999). You should administer the survey instrument to respondents from the same or similar population of interest in a setting that is as similar as possible to your actual study. The size of the pilot sample may range from as few as 10 respondents to as many as several hundred depending on various factors, such as the purpose of your pilot study (Hertzog, 2008; Johanson & Brooks, 2010), the size of the target population, your access to a representative sample of the target population, the number of items in the survey, and sufficient power for determining statistical significance (Johanson & Brooks, 2010).

You should use the information obtained from a pilot test to evaluate the reliability and validity of your survey instrument (Martella et al., 1999). You can also use the pilot test as an opportunity to ask respondents questions regarding the survey, such as how easily they understood the items, how they arrived at their responses, and their opinions about its user-friendliness, all of which can inform potential revisions to the instrument before your actual study begins.

Reliability and Validity of Measurement

In survey research, *reliability* is the consistency of responses over time. There are two types of reliability that are important: test–retest reliability and split-half reliability. *Test–retest reliability* refers to whether or not the survey instrument yields the same result with repeated administration. To determine test–retest reliability, researchers calculate the correlation of scores (called the reliability coefficient) from participants who completed the survey at two different points in time. *Split-half reliability* refers to the method for evaluating the internal reliability of a survey instrument by correlating one half of the items with the other half. To determine split-half reliability, researchers might score the odd-numbered items and even-numbered items separately, thus yielding two scores for each respondent. A correlation coefficient (called the split-half reliability coefficient) is then computed on these scores to provide the instrument's reliability.

Validity refers to an instrument's ability to measure what it was designed to measure. There are three main types of validity that are important in survey research: face, construct, and criterion validity. *Face validity* refers to whether or not a survey instrument appears to reasonably measure what it intends to measure. *Construct validity* refers to whether or not a survey instrument measures the issue it was intended to measure, typically in relation to a particular theory. *Criterion validity* refers to how well an instrument correlates with another instrument that measures a similar construct, called a criterion measure. Thus, you would validate a new survey instrument by comparing its scores to those from an established instrument known to measure the construct being studied.

Appropriateness of Survey Research for Your Research Project

If you are considering survey research, here are some questions to ask yourself:

- Is my topic appropriate for survey research?
- Is there an existing reliable and valid survey instrument available to measure my phenomenon of interest? If not, then
  - Will I need to design my own survey instrument?
  - Will I be able to gather self-report information from participants?

As with any research study, be sure that your choice of method is aligned with your research purpose, problem, and research question. Consult with your supervisory committee about your research objectives to determine whether survey research is the appropriate method. We also encourage you to contact your institution's institutional review board with any ethics-related questions or concerns about designing and/or administering a survey.

Conclusion

Survey research is a descriptive, nonexperimental research method that entails a set of procedures for exploring, describing, and explaining phenomena, such as attitudes and behaviors, through interviews or questionnaires from a sample of individuals. To determine an approach for selecting a sample of individuals from a larger population, you must decide to use either a probability or nonprobability sampling approach. Although probability sampling methods are preferred, they are not always possible or practical.

When the probability of selecting a participant is unknown, nonprobability sampling methods are used.

Survey instrument is the term that refers to the tool used to gather descriptive, behavioral, and attitudinal information from the sample. A survey instrument can take the form of a questionnaire or interview to be administered in person or by postal mail, email, or the Internet. You should weigh the costs and benefits of each type of survey instrument with consideration of your study's purpose, research question, sampling method, and data analysis plans.

The design of a survey instrument involves a set of procedures that requires careful planning and decision making at numerous stages in the process. Before you decide to design a survey instrument, it is important that you conduct a thorough review of the literature, reference materials, and databases to determine whether an existing survey instrument is available.

In designing a new survey instrument, great care must go into writing survey items and rating scales to yield responses that are valid and reliable measures of your construct or phenomenon of interest. You should conduct a pilot test if you develop a new survey instrument to gain preliminary information about how the instrument performs under realistic conditions.

KEY SOURCES

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method* (3rd ed.). Hoboken, NJ: Wiley.

Fowler, F. J., Jr. (1995). *Improving survey questions: Design and evaluation*. Thousand Oaks, CA: Sage.

Fowler, F. J., Jr. (2009). *Survey research methods* (4th ed.). Thousand Oaks, CA: Sage.

Rossi, P. H., Wright, J. D., & Anderson, A. B. (Eds.). (1985). *Handbook of survey research* (pp. 195–230). New York, NY: Academic Press.

Schuman, H., & Presser, S. (1996). *Questions and answers in attitude surveys: Experiments on question form, wording, and context*. Thousand Oaks, CA: Sage.

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, *54*(2), 93–105.

Sue, V. M., & Ritter, L. A. (2011). *Conducting online surveys* (2nd ed.). Thousand Oaks, CA: Sage.

REFERENCES

Barusch, A. S., & Wilby, F. (2010). Coping with symptoms of depression: A descriptive survey of community-dwelling elders. *Clinical Gerontologist*, *33*, 210–222. doi:10.1080/07317111003773650

Converse, J. M. (1987). *Survey research in the United States: Roots and emergence 1890–1960*. Berkeley, CA: University of California Press.

Duan, C., Brown, C., & Keller, C. (2010). Male counseling psychologists in academia: An exploratory study of their experience in navigating career and family demands. *Journal of Men's Studies*, *18*(3), 249–267. doi:10.3149/jms.1803.249

Farc, M., & Sagarin, B. J. (2009). Using attitude strength to predict registration and voting behavior in the 2004 U.S. presidential elections. *Basic and Applied Social Psychology*, *31*, 160–173. doi:10.1080/01973530902880498

Fink, A. (2009). *How to conduct surveys: A step-by-step guide* (4th ed.). Thousand Oaks, CA: Sage.

Hertzog, M. A. (2008). Considerations in determining sample size for pilot studies. *Research in Nursing & Health*, *31*(2), 180–191.

Johanson, G. A., & Brooks, G. P. (2010). Initial scale development: Sample size for pilot studies. *Educational and Psychological Measurement*, *70*(3), 394–400. doi:10.1177/0013164409355692

Levy, P. S., & Lemeshow, S. (2009). *Sampling of populations: Methods and applications* (4th ed.). New York, NY: Wiley.

Martella, R. C., Nelson, R., & Marchand-Martella, N. E. (1999). *Research methods: Learning to become a critical research consumer*. Needham Heights, MA: Allyn & Bacon.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). E-research: Ethics, security, design, and control in psychological research on the Internet. *Journal of Social Issues*, *58*(1), 161–176.

Salant, P., & Dillman, D. A. (1994). *How to conduct your own survey*. New York, NY: Wiley.

Tourangeau, R., Rips, L. J., & Ransinki, K. (2000). *The psychology of survey responses*. New York, NY: Cambridge University Press.

14

Case Study Research

Kurt Schoch, EdD

Introduction

In this chapter, I provide an introduction to using case study as an approach to exploring and understanding a phenomenon. The chapter begins with a definition of case study research and a description of the origins and philosophical underpinnings of the approach. Several discipline-specific applications of case study methods are presented and the appropriate research questions addressed by case studies are described. This description is followed by methods considerations including case study design, research questions, sample size, data collection, and data analysis. The chapter also helps clarify some misconceptions about case study research.

Background

Case study research involves a detailed and intensive analysis of a particular event, situation, organization, or social unit. Typically, a case has a defined space and time frame: “a phenomenon of some sort in a bounded context” (Miles, Huberman & Saldana, 2014, p. 28). In terms of scope, a case study is an in-depth investigation of a contemporary phenomenon within its real-life context. The case study approach is appropriate especially if you think the context is relevant to the phenomenon—for example, investigating an achievement gap in a high school with a large second-generation immigrant population. Because the boundaries between a phenomenon and its context are not always clear, the case study design relies on multiple data sources for evidence (Yin, 2014, p. 17).

A *case* could be an individual, a role, a small group, an organization, a community, or a nation (Miles et al., 2014, p. 28). Examples of a case include

- A remote town affected by a natural disaster, such as an earthquake (see Parrinello, 2013). In this example, the town or area constitutes the case.
- A refugee group resulting from a political conflict (Pedersen, 2012). In this example, the group is the case.
- A president, a high school principal, or chief senior manager. For example, Scribner and Crow (2012) reported on the case study of a high school principal in a reform setting.

- A specific issue in a company or organization. For example, Shen (2007) described approaches to international industrial relations in multinational corporations.

Two examples of well-known case studies include Sigmund Freud's (1905/1953) study of Dora and Graham Allison's (1971) case study of the Cuban Missile Crisis. Published in 1905, Freud's case study of Dora used recollections, reconstruction, and dream analysis to depict a young woman suffering from hysterical symptoms, including difficulty breathing, nervous coughing, and headaches. Freud thought her case was rather uninteresting in its own right but could help provide further knowledge about Dora's disorder. His point was that analysis of even an ordinary case can have an application in similar situations.

During the Cuban Missile Crisis in 1962, President John F. Kennedy and his leadership team made decisions to avoid nuclear conflict. Allison's (1971) case study of the Cuban Missile Crisis, based on a broad range of data ranging from government documents to interviews with numerous officials, has often served to instruct others about leadership styles and processes, based on how Kennedy handled this difficult situation. This case study also clearly demonstrates how a case study is used for explanatory purposes and not just descriptive or exploratory purposes (Yin, 2014, p. 7).

Finally, it is important to distinguish teaching case studies from those used for research purposes. Teachers use case studies to illustrate a lesson of instruction and spark discussion among students in the classroom. Teachers often use case studies, especially in the fields of law, public policy, and management education. For our purposes within this chapter, a *teaching case study* is different from a *research case study*. Although case studies involving teaching do not necessarily need to present data in a rigorous and scholarly way, this is precisely what research case studies need to do (Yin, 2014, p. 5).

Benefits of Using a Case Study Research Approach

Case study research offers benefits in terms of process as well as outcome. In terms of the process, the case study approach may help prevent the scope of the research project from expanding beyond your original intent because you will focus your research within the confines of space and time on a specific case. A case study also gives you an opportunity to collect different kinds of data about the case and provides you with the chance to get an in-depth look at an organization or individual and the inner workings and interactions of that organization or individual.

In terms of the outcome, the case study design provides a comprehensive understanding of a bounded unit and helps the reader examine that case so he or she can learn from it. It also allows others to apply the principles and lessons learned in a case to other cases or situations and leads to transferability (i.e., the ability to apply the case to another situation), which is different from the generalization that occurs in quantitative studies. For example, if someone wanted to learn what it is like to be a high school principal, a case study could help that person learn about that experience and apply it to another situation or help that person decide whether being a principal is his or her best career path.

Origins of the Case Study Design

Case study has been used as a research design for many years. From the early 1900s to the mid-1930s, the Department of Sociology at the University of Chicago was a leader in case study methods and a source of much literature in the field. The city of Chicago, with its diverse, immigrant population and issues with poverty and unemployment, had abundant material for case studies (Tellis, 1997).

The case study approach became more common in the 1950s and 1960s, and in the past few decades, it has been used in many social science applications. To gain more in-depth, contextual perspectives on specific cases, disciplines such as sociology, psychology, political science, and anthropology have used case studies. More recently, the fields of education and management have used case study designs. Key contributors to the development of case study design and methods include Guba and Lincoln (1985), Stake (2006), and Yin (2014).

Philosophical Underpinnings of Case Study Research

Case study research primarily fits in the postpositivist philosophical orientation. Postpositivism implies the existence of an ultimate reality that we can only approximately—not completely—understand. Case study research also tends to be constructivist, meaning those who engage in it (and those who read it and learn from it) tend to create or construct their own sense of reality and learn from the research process and outcomes. As a learning theory, the emphasis in constructivism is on the learner's ability to make sense and meaning out of his or her experiences; this is the essence of cognitive constructivism. Dewey (1938), Piaget (1974), Kelly (1955), and Gergen (2015) have all contributed to our understanding of constructivism.

Case study research can also contain critical, feminist, and other postmodern concepts, depending on the research purpose. That is, a case study can expose the historical, political, gender, socioeconomic, and other structures that make material differences in peoples' lives and hold them back from their true potential.

The Key to Focusing Case Study Research

Case study research focuses on a specific event, person, place, thing, organization, or unit of some kind (or potentially more than one but typically a small number). Thus, a case study does not involve just conducting a quantitative or qualitative survey of an organization to get an idea of people's perceptions on a topic. The case study must involve multiple sources of data about the entity under investigation to gain a broad, robust understanding. The key to case study research, then, is identifying the case and the boundaries of that case. The key question to ask is, "What is the *case*?" You can help yourself answer that by answering the question, "What is inside and what is outside

BOX 14.1

Discipline-Specific Applications of Case Study Design

Case study research is used widely across disciplines; this type of research is most common in education and other social sciences, as well as law, political science, and health care. For example:

In this discipline . . .	a case study could investigate . . .
Health care	A hospital or specific type of patient
Education	A school or a principal
Business	A business, a chief executive officer (CEO), or a startup organization
Public policy	A natural disaster and policy implications for responding to it
Psychology	An individual with a certain disorder or a mental health center and its responsiveness to community needs

the case?" Yin (2014) refers to this as "bounding the case" (p. 33). If the case is one school, then other schools are not part of the case; if the case is the principal, other teachers or administrators are not part of the case. Therefore, the specific unit of study, or unit of analysis, is a defining characteristic of a case study. Take care when defining the case, because the research questions drive the boundaries and definition of the case (See Box 14.1).

Examples of Research Studies That Use the Case Study Approach

Below are some examples of research studies that use the case study approach. These examples will be used throughout the rest of this chapter to provide examples of key elements of case study research. Note that each has clear boundaries of place and time, which is a defining characteristic of case studies.

- Militello, Schweid, and Sireci (2010) conducted a multiple-case study to explore the use of formative assessment systems (processes that students and teachers use to adjust their learning and teaching strategies). The authors wanted to better understand how educators seek and obtain assessment tools. The study lasted 18 months, used a sample of three school districts (thus, this represents a *multiple-case* study), and focused on the following overarching question: "To what extent does the fit between intended use and system characteristics foster or inhibit the ultimate utility of formative assessment systems for schools?" (p. 34). The results were focused on three key findings, including (a) the reasons a school district might want an assessment system, (b) characteristics of the assessment systems, and (c) how formative assessment systems are used (p. 36).

- Joia (2002) conducted a single explanatory case study, analyzing the use of a web-based virtual learning environment for business students in Brazil. Although this is considered a single-case study, in that the case is the learning environment, three different cities (sites) were used for data collection. Five research questions were used, focusing on the collaborative nature of virtual learning environments, the influence of the physical location of students, use of a moderator, and the students' perceptions of a particular learning environment used in the school. Results were organized around each of the research questions, including the conclusion that there was collaborative learning in place, but the particular platform did not support the accomplishment of collaborative work.
- Lotzkar and Bottorff (2001) observed the development of the nurse–patient relationship between one nurse and one patient over a 3-day period in a hospital's cancer unit. In this single-case study, the researchers' goal was to use interactions between the nurse and the patient in a cancer treatment center to describe how nurse–patient relationships develop (p. 276). Sixty interactions were videotaped over 3 days and analyzed using ethology, a detailed qualitative analysis involving both inductive and deductive methods. Note that the videotaping used for this research initially resulted in some intrusiveness for the participants, until both became accustomed to the research process. It is important to understand in any qualitative research process that the presence of the researcher, and any artifacts of the research process, will disrupt the research environment, making it critical for the researcher to recognize such disruptions and work to minimize them as much as possible. The researchers discovered "behavioral clusters" (p. 275), patterns of behavior that characterized the developing nurse–patient relationship. These clusters lent insight into how effective relationships can contribute to quality patient care.
- Scheib (2003) used a collective case study design to examine role stress among four music teachers at a high school in the Midwestern United States. This study elaborated on previous literature that examined specific stressors in a music teacher's life to learn more about why those stressors exist. Scheib used observations, individual interviews, and document analysis to understand stressors related to areas such as role conflict, role overload, and resource inadequacy. Scheib noted the essential nature of triangulation in balancing all aspects of the data that were collected.
- Campbell and Ahrens (1998) conducted a multiple-case study examining social service programs for rape victims in 22 communities across the United States in the late 1990s. Interviews were conducted with various service providers from programs in these communities; a program in each community was considered a case. The data were used to understand the richness of each program and led to developing an overall model of effectiveness of social service programs (p. 541).

The Process of Doing Case Study Research

All research needs a design, and case study research is no different; design in case study includes the steps that connect the first research thoughts to the final research conclusions. Those first research thoughts begin with a research problem, which involves the identification of a lack of knowledge about some issue or phenomenon. Identification of the problem then leads to crafting the purpose statement and research question(s). Examining the research question(s) then allows you to determine whether a case study is the appropriate research paradigm.

Characteristics for selecting case study research versus other approaches focus on *how* or *why* kinds of research questions directed at exploring and understanding some phenomenon in depth (Yin, 2014). Once the decision is made to conduct a case study, a key decision involves selecting the case itself. Yin (2014) noted the two required elements as "defining the case and bounding the case" (p. 31). The former relates to clearly and concretely indicating the case, which can be a person, place, thing, organization, or phenomenon. The latter relates to scope—what is, and is not, included in the case, whether from time, structure, or other perspectives. Once the case has been identified, you can determine the types of data needed and how those data will be collected.

Case Study Sample Size

In case study research, the case itself is often the sample as well as the number of individuals within the scope of the case whom you choose to interview. Typically, a case study has a sample of one, unless the research project is a multiple-case study. In a multiple-case study, having three to four distinct cases for comparison is probably the most cases that can realistically be handled. When using multiple cases or sampling within a case, it is effective to use a selection method known as *purposeful sampling*. By selecting the cases, and the individuals, documents, and artifacts within the case, purposeful sampling allows you to focus in depth on a phenomenon. It allows you to explore information-rich cases from which you can learn a great deal about issues of central importance to the research (Patton, 2002, p. 46). By selecting unusual cases or finding a good amount of variation to represent diverse cases, purposeful sampling allows the researcher "to select information-rich cases whose study will illuminate the questions under study. Information-rich cases are those from which one can learn a great deal about issue of central importance to the purpose of the research" (Patton, 2002, p. 46).

One way to understand purposeful sampling used for case studies is to contrast it to sampling for quantitative research. Samples obtained for quantitative research studies are often probability samples that are presumed to be representative of the population being studied and are used to generalize to that population; in a way, it does not matter who the individuals in the sample are—only that they are statistically representative of the larger population. In purposeful sampling, however, the goal is to find individuals or cases that provide insights into the specific situation under study, regardless of the general population. Note that you should not select cases or individuals just because they happen to be available. This is a sampling technique known as convenience sampling. Select cases and individuals for the value they can bring to the study, not simply because they are easy to access.

Case Study Sample Selection

Sample sizes in case studies are typically small, which is common in most qualitative research. Sometimes, the selection of samples and cases to use is straightforward and clear, due to the uniqueness of the person or organization or because of special arrangements or access to the case. In some situations, however, there may be many qualified case study candidates, and you may have to use a screening procedure to select the proper ones. Yin (2014) suggested the possibility of asking knowledgeable people about the case candidates or collecting limited documentation on them. What you clearly want to avoid is selecting a case that is representative of something other than what you want to study.

The highlighted research provides some insight into sampling and sample size. Joia's (2002) work is an example of a single-case study analyzing a virtual learning community in Brazil; the case was a web-based learning community. This particular case comprised 43 students enrolled in a graduate course on e-commerce. Scheib (2003), in his collective case study about the role of stress on high school music teachers, first selected the site based on the fact that he had access to it, that it had well-established music programs, and that the music department offered band, choir, and orchestra—a variety of musical programs. After selecting the site, the researcher focused on the four music teachers who taught there; the four teachers comprised the sample.

In other instances, screening for a sample can often be more involved. For example, Militello et al. (2010) used a sample of three schools from three different school districts in their research on formative assessment. They chose the schools by identifying prominent formative assessment companies and the districts that had contracts with these companies. After consulting with the state Department of Education to find out which districts where using assessment in a significant way (p. 34), the researchers narrowed their selection based on the use of formative assessments in middle school mathematics. The researchers then selected three school districts based on phone interviews with assessment personnel from the state Department of Education. The study focused on a single school from each of the three districts.

Sometimes a previous study is used to inform the sample for a follow-up study. For example, Lotzkar and Bottorff (2001), in their nurse–patient relationship case study, included a sample of one nurse–patient dyad. In a previous study, the researchers videotaped eight patients continuously for a period of 72 hours. Out of these eight patients, one nurse–patient dyad was selected because it exhibited the most interactions, allowing the researchers the most opportunity for insight.

Case Study Research Questions

Research questions for a case study can be both quantitative and qualitative but frequently use terms similar to other qualitative research designs. For example, these kinds of questions focus on concepts such as *explain*, *explore*, *describe*, and *understand*. Typically, case study research questions use words such as *how* or *why*. Overall, the case study research questions need to address the substance of what (case) the study is about (Hatch, 2002, p. 10).

To formulate your case study research questions, think of conducting a case study like painting a picture. What does the case look like, whether it is an individual, organization, or situation? What image will the reader have in his or her mind after reading the case study? To paint that picture, what kinds of research questions would you need to ask?

Some case studies have one or two broad research questions. For example, Scheib's (2003) collective case study of school music teachers explored, through the lens of role theory, the open-ended question of why their work life is stressful. In another example, Campbell and Ahrens (1998) published a case study on rape victim services and set out to answer how and why coordinated service programs are effective. The case study on nurse–patient relationships sought to answer the question of exactly how such relationships develop (Lotzkar & Bottorff, 2001). Other case studies state more specific questions. For example, Joia's (2002) case study on a virtual learning community had five specific questions, including "Why is a moderator needed, and how can his importance be measured?" (p. 309).

Some use an overarching question followed by subquestions, such as these from Militello and colleagues' (2010) work:

> To what extent does the fit between intended use and system characteristics foster or inhibit the ultimate utility of formative assessment systems for schools?
>
> 1. Intended Use: What data and action did each district want from the assessment system?
> 2. System Characteristics: What were each of the formative assessment systems designed to do?
> 3. Actual Use: How are school district educators using the assessment systems? (p. 34)

Note that there is an important difference between the research questions for the case study and the interview questions used to elicit information from participants. The research questions are broad, focused on what you as the researcher ultimately want to learn. They are not directed at any one person, document, situation, or occurrence within the case; rather, the questions are directed at the case as a whole, and answers to the research questions will be derived from all the sources of data. You do not ask your participants the research questions; you ask them interview questions that will help you answer your research questions. Interview questions are directed at individuals or groups within the case and often contain the word *you*. For example, the questions might start with, "How do you . . . ?" or "Why do you . . . ?" As a hint, if the word *you* appears in a research question, it is probably an interview question and not broad enough to be a research question.

Patton (2002) identified six different types of interview questions. He noted that "[d]istinguishing types of questions force the interviewer to be clear about what is being asked." The types of interview questions include "experience and behavior questions, opinion and values questions, feeling questions, knowledge questions, sensory questions, and background/demographic questions" (pp. 348–351).

Data Collection Procedures

Types of Data Collection

In case study, as with all research approaches, the research questions drive the data to be collected. From the research questions, the researcher determines the kinds of questions to be asked in interviews, what to observe, what documents to review, and what artifacts to examine. Therefore, multiple sources of data are used in a case study. Sources of evidence may include one or more of the following: (a) documentation; (b) archival records; (c) interviews; (d) direct observation; (e) participant observation; and (f) physical artifacts (Yin, 2014, p. 106).

If you use surveys to collect data, the survey instruments need to be valid and reliable. For interviews, you will need interview protocols, or a list of questions and prompts used to interview participants. Whether your interviews are structured or unstructured, protocols are useful to ensure the consistency of the interviews across the individuals being interviewed. You may also use different interview protocols for different groups within the case. For example, you might use different interview protocols for teachers than for students. Observations also require protocols. Protocols ensure that interviewers understand what they are looking for and how the observations will help answer your research question(s). Similarly, be sure you know how documents and physical artifacts will help answer the research questions; know, in advance, what you are looking for. Yin (2014, pp. 118–129) also advocates the following general principles of data collection in case study research.

1. **Use multiple sources.** Use different types of data and obtain different perspectives by using a variety of people and other sources. For example, Scheib's (2003) case study on music teachers and stress relied on a variety of data, including direct observations of and interviews with the music teachers, school policy publications, documents sent to students and parents, and music concert programs. The case study by Campbell and Ahrens (1998) on rape victim programs used interviews from various sources, including rape victim advocates, crisis center directors, police, prosecutors, and medical staff, as well as the rape survivors themselves. The researchers supplemented these interviews by examining pamphlets and training manuals (i.e., documents) for each community program. In his virtual learning community case study, Joia (2002) collected data from the course's web-based environment, which included a moderated discussion group as well as a variety of features such as email and webmail, a file directory, a bookmark directory, chat sessions, and online polls. The researcher collected quantitative data such as a number of email messages to analyze characteristics such as participation levels. Joia also analyzed qualitative data such as email content to investigate the moderator's role in discussions. Lotzkar and Bottorff (2001) collected data from 60 interactions videotaped by two video cameras over a 3-day period, as well as demographic and clinical data on the patient.

Data can also be collected in multiple phases. For the formative assessment case study, Militello et al. (2010) collected data in three stages during the 2006–2007 school year. In the first stage, the researchers conducted interviews with district administrators, attended district-level meetings, and collected artifacts such as internal memos. The interview protocols at this stage focused on the organizational history and goals of formative assessment. In the second stage, the researchers interviewed test developers, with questions focusing on the purposes of the assessments and whether the assessments fulfill their purposes. At this stage, research data also included technical documentation from assessment companies. In the third stage, the researchers interviewed school-level educators including principals, guidance counselors, and teachers. The interview protocol focused on how educators used the data from formative assessments. The researchers also attended math department and grade-level meetings, using observational notes as data for the study.

2. **Build a case study database.** Keep your notes, documents, narratives, and other materials organized by maintaining a physical or electronic file system. Contents of such a database can include field notes, documents, surveys or other quantitative data, and other narratives. Consider using software such as NVivo Version 10 (QSR International, 2012) to help you stay organized.
3. **Establish a chain of evidence.** Be sure you are able to track the final conclusions or report backward through your notes or database to your approach, protocols, and research questions. From your conclusion, someone should be able to backtrack through the data to see how it is supported. Ask yourself, "Can I do a backward line of sight from my report all the way through to the research questions? Is there alignment between the conclusions, the evidence, and the research questions?"
4. **Exercise care when using electronic sources.** The sheer volume of data and information available electronically can be overwhelming, so it is important to set limits on the amount of time you will spend collecting this kind of data and to be sure you can make a strong case for relevance of this information to your research purpose. Also, the ease of posting and accessing data electronically requires caution regarding the credibility and accuracy of such data. Cross-check and double-check references and sources. In addition, do not forget to seek permission to use electronic sources; just because information is on the Internet does not mean it is in the public domain.

Data Analysis Techniques

Advice from Patton (2002) is well worth heeding: "The best advice I ever received about coding was to read the data I collected over and over and over. The more I interacted with the data the more patterns and categories began to 'jump out' at me" (p. 446). Data, or content, analysis involves several phases: describing, interpreting, drawing conclusions, and determining significance. Specific techniques will vary based on the type of study and its purpose, and individual researchers will often develop their own approach and style. Typically in a case study, data analysis will involve the following steps.

Describing. Basic description of the case involves understanding the "who, what, when, where" of the situation under study. The descriptive phase of the analysis of the data involves several readings and reviews of the data collected, whether from interviews, observations, or other sources. Field notes—notes taken while at the research site or while studying parts of the case—are reviewed extensively to discover patterns or themes. Patterns tend to be descriptive, such as "most students are excited about the first day of class," whereas themes are often more topical, such as "excitement about school."

Emergence of Findings. Through the researcher's interaction with, and often immersion in, the raw data, findings in the form of patterns, themes, or categories tend to emerge. This is the essence of the inductive form of qualitative data analysis, where findings emerge out of the data. The initial stages of this analysis involve open coding, which means that the researcher is open to what the data are saying without bringing in any preexisting codes. Open coding emphasizes recognizing any patterns that emerge from the data, rather than analyzing data based on an existing framework as one might do in deductive analysis with quantitative data. These codes summarize and put labels on the patterns, themes, or categories you see in the data.

Codes become brief shorthand labels to passages of data or other items for easier organization and recognition. Coding is essential when multiple researchers are working on the same project to ensure common understanding and consistency. Basic coding can begin with the initial review of data, but it becomes more systematic upon subsequent data reviews and reflection. Furthermore, codes can be identified in advance of the data analysis, which is known as *a priori coding*, based on what the researcher anticipates seeing in the data, or codes can be allowed to emerge from the data, either from the information itself or from the individuals interviewed. The latter is often called *in vivo coding*. Codes may also arise from a theoretical framework (hence the need for your research to be grounded in the literature). An accompanying step to open coding is axial coding, where relationships or connections are identified among the initial categories and themes identified during open coding.

Although researchers often think of open and axial coding as separate and sequential steps, Corbin and Strauss (2008) noted that "open coding and axial coding go hand in hand" (p. 198). Combining open and axial coding becomes a process of concurrently dissecting the raw data and reassembling it at the same time. While involved in this dissection/reassembly, you may find yourself involved in constant comparative approaches, comparing each piece of data with the ones that came before it; this is a common approach in many types of qualitative research. Selective coding is a final step in the coding process, allowing you to create larger categories that connect the previously identified categories. Creswell (2007) referred to this as a creating a "story line" (p. 67). Here is one final thought on coding: It is helpful to maintain a description or definition of your codes, often in visual form, to allow for easy identification and revision later.

Here is an example of the initial stages of the coding process, provided as a table (Table 14.1) from a dissertation by Cox (2011):

TABLE 14.1
Initial Codes from First Round of Theoretical Sampling

Participant	1	2	3	4	5
Initial Codes	**Values**	Motivations	Tradition	Challenge	Goals
	Service	Teamsmanship	Expectation	Growth	Standards
	Socialization	Adventure	Symbolism of uniform	Adventure	Ethics
	Leadership as shaping	Direction	Values	Overcoming limits	Motivations
	Responsibility as shaping	Resilience	Integrity	Defining moments	Culture shock
	Culture shock	Agility	Proving self	Continuous learning	Growth
	Power of opportunity	Self-reliance	Growth	Ethics, values	Transformation
	Overcoming limits	Learning from leaders	Drive	Transformation	Risk taking
	Self confidence	Integrity	Family	Politics vs values	Learning
	Change world view	Respect	**Responsibility**	Symbolism of rank	Self-worth
			Discipline	Definition of self	Values
			Identify	Calling	**Leadership**
			Change	Tradition	**Civic duty**
			Learning	Right thing	Role models
			Civic behavior		Selflessness
			Transformation		**Belonging**
			Sacrifice		Appreciation for country
					Tolerance
					Tested by experience
					Self confidence
					Judgment
					Mission

Table reprinted from Cox, R. (2011). *The effects of military experience on civic consciousness* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses Database. (UMI No. 3457222).

Cox (2011) describes some of his coding process:

> During initial coding, digital transcripts were hand-coded to establish patterns in the data and to identify the initial codes in [Table 14.1]. . . . After the first five participant interviews and transcriptions, data were hand coded on the transcripts, analyzed, and preliminarily organized in the structure of an NVivo tree and case nodes. (p. 124)

Cox further explained that primary codes appear in bold, and that "secondary codes, unbolded, appear to fit within the scope of primary codes" (p. 123).

Comparing. Final steps in case study data analysis can include making comparisons across the various themes that have emerged from the data, as well as making comparisons across different cases, if the study was a multiple-case study design. It is important to understand that these steps may not necessarily happen linearly, but in an iterative way. You may be comparing themes as you identify them. You may even need to go back and collect more data after completing these steps. This aspect is somewhat unique to qualitative research.

Examples of Data Analysis

Sometimes data analysis includes quantitative methods. In the case study on a virtual learning environment, Joia (2002) described his process of data analysis as data consolidation. He used tables and charts to describe behavior and usage in the virtual community and the geographic distribution of its students in terms of participation. In addition, the researcher analyzed the typology of the web-based community according to a taxonomy established by previous research.

In the nurse–patient relationship case study, Lotzkar and Bottorff (2001) analyzed the videotaped data using quantitative ethological methods in four steps. First, the researchers reviewed the videotaped data to identify and describe behaviors of interest, and then they reviewed the tapes to identify clusters of behaviors indicating the development of the relationship. Third, the researchers described patterns of behavior within each cluster and compared the clusters, and finally, they constructed a detailed behavioral description, including conditions, cause or function, and consequences of the observed behaviors.

Codes are sometimes identified from previous research. As part of his case study on role stress and music teachers, Scheib (2003) used various coding techniques to analyze the interview transcripts, field notes, and documents. Among the codes used by the researcher were the six specific role stressors (e.g., role conflict, role ambiguity, role overload) identified by previous researchers. The researcher used the data to create detailed descriptions of the high school, the teachers, and the context of the case study.

Codes can also come from both previous research and unreported findings. In the Militello et al. (2010) formative assessments study, data analysis consisted of a "coding phase and comparative analysis" (p. 36). The research team created memos from field notes and artifacts. They exported the memos, along with their interview transcriptions, into a computer database, and they described them using an open coding system. They applied codes to represent themes identified in existing literature, as well as codes to represent unreported findings. Using a similar coding system, researchers also analyzed the technical data. The researchers then compared the data to analyze gaps between the intended and actualized use of formative assessments in the case study.

It is not unusual for the process of analysis to need revision along the way. In the Campbell and Ahrens (1998) case study on programs for rape victims, four coders initially reviewed interview transcripts, but some issues arose that required the researchers to independently review the transcripts and develop a reliable system of coding. Once the researchers agreed upon a coding scheme, two to four of the coders rated each interview, filling in three blank tables to describe each community.

The final step is case study reporting. Regardless of the purpose of your case study report—whether you are a student writing a thesis or dissertation or a researcher preparing an executive summary—be sure to

- Provide a thorough description of the case.
- Separate reporting from interpreting (reporting means presenting the facts: what happened, what did you see, etc.; interpreting involves finding meaning in the data).
- Include sections outlining your methods and your literature review, including how that literature led to your research questions.
- Ensure that the reader can easily follow the progression from your original problem, purpose, and methods to your analysis conclusions (there should be a clear sense of alignment among these items). Remember Yin's (2014) admonition to establish a chain of evidence.
- Make it clear what the case study informs and how it lays the groundwork for future studies.
- Write the report in a way so that a person not involved in the case can understand it (and limit the use of technical language so that a broader audience can learn from your experience).

Appropriateness of Case Study for Your Research Project

To determine whether a particular research approach is right for your project, follow these steps:

- Start with the problem (that is, the gap in literature, theory, or practice).
- Determine the purpose of the research. (What do you want to accomplish? Why are you doing the research?)
- Draft the research question that will guide the overall study.
- Based on the research question, determine the design you will use.

Do not decide that you want to do a case study without seriously considering the research questions and whether they are consistent with a case study design. Also, qualitative research is not inherently easier than quantitative research (many would argue it is actually much more challenging). The research design must align with the research problem, purpose, and question.

Consider also the types of questions that you would like to address in your research. Case studies can address questions about who, what, where, how, or why, or any combination of those. Case studies answer questions focused on understanding or explaining.

What Is the Role of the Researcher in a Case Study?

The researcher is the key data collection instrument in a case study, as is true for most qualitative research. Typically, the researcher conducts interviews, administers surveys, reviews and analyzes documents, and observes whatever is being studied. In a case study, the researcher is situated in the activity or organization being studied. By contrast, in quantitative research, the research may be done in a location other than where the object of study is located. For example, the researcher may not need to be present to conduct surveys. In qualitative research, the researcher is less separated from the object of study than in quantitative research.

As a researcher, you must be careful to avoid bias, or the tendency to prejudice or unduly influence the process or results of a research project. Be constantly aware of your own feelings, opinions, and prejudices, and make sure you are open to data and evidence that might not fit your notion or idea of what you might find. Do not enter into case study research, or any research, to demonstrate a previously held position or advocate a particular point of view. You can mitigate potential bias by using techniques shared by all qualitative research, such as journaling, triangulation of data, and member checking.

Can the Case Study Design Be Used With Other Approaches?

A case study can be embedded in a larger research project that uses another approach. For example, an ethnography study could explore the culture of an urban school. As part of that study, the researcher could conduct a case study of the school's principal, focusing on the influence of the principal on the school's culture. In this way, the ethnography uses the case study of the principal to understand the larger picture of the school culture.

A case study can also use the methods of another research approach. For example, if you are doing a case study on an organization, you may employ quantitative surveys to assess the organization climate or conduct phenomenological interviews with management staff to find out about their experiences as managers.

Summary

A case study is an in-depth investigation of a contemporary phenomenon within its real-life context. A case study typically relies on multiple data sources and is bound by both space and time. Any discipline can use case study research, and case studies can be used with other research approaches.

Research questions for a case study should be open-ended, few in number, and straightforward in language. Sample sizes in case studies are typically small, as is common in most qualitative studies. The researcher must be careful to avoid bias when collecting data and follow general ethical considerations that apply to human subject research. The researcher can mitigate potential bias by using techniques such as journaling, triangulation of data, and member checking.

Data in a case study may include interviews, documents, observations, artifacts, surveys, and focus groups. The researcher should collect different types of data and obtain different perspectives from a variety of sources. The researcher should also maintain a file system for the data and be able to backtrack from the conclusions to the evidence to the research questions.

Case study data analysis generally involves several phases: describing, interpreting, drawing conclusions, and determining significance. These steps may not necessarily happen linearly but more in an iterative way. In writing the case study report, the researcher should provide a thorough description of the case, separate reporting from interpreting, and limit the use of technical language.

Threats to trustworthiness can include researcher bias, as well as the possibility that when interviews are conducted or the case is observed, these incidences might not be representative of typical operations. Purposeful sampling, at the beginning of the project, can help ensure that data collected represent typical operations. After collecting data, triangulation, member checks, and other strategies can further ensure trustworthiness of the study.

KEY SOURCES

Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The Qualitative Report, 13*(4), 544–559. Retrieved from http://www.nova.edu/ssss/QR/QR13-4/baxter.pdf

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

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Stake, R. E. (1995). *The art of case study research.* Thousand Oaks, CA: Sage.

Stake, R. E. (2000). Case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed.). Thousand Oaks, CA: Sage.

Stake, R. E. (2006). *Multiple case study analysis*. New York, NY: Guilford Press.

Tellis, W. (1997). Introduction to case study. *The Qualitative Report, 3*(2). Retrieved from http://www.nova.edu/ssss/QR/QR3-2/tellis1.html

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Thousand Oaks, CA: Sage.

REFERENCES

Allison, G. (1971). *Essence of decision: Explaining the Cuban missile crisis*. New York, NY: Little, Brown.

Campbell, R., & Ahrens, C. E. (1998). Innovative community services for rape victims: An application of multiple case study methodology. *American Journal of Community Psychology*, *26*(4), 537–571.

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Thousand Oaks, CA: Sage.

Cox, R. (2011). *The effects of military experience on civic consciousness* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses Database. (UMI No. 3457222)

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage.

Dewey, J. (1938). *Experience and education*. New York, NY: Touchstone.

Freud, S. (1953). Fragment of an analysis of a case of hysteria. In J. Strachey (Ed. and Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 7, pp. 1–122). London, United Kingdom: Hogarth Press. (Original work published 1905)

Gergen, K. (2015). *An invitation to social construction* (3rd ed.). Thousand Oaks, CA: Sage.

Hatch, J. A. (2002). *Doing qualitative research in education settings*. Albany, NY: State University of New York Press.

Joia, L. A. (2002). Analysing a web-based e-commerce learning community: A case study in Brazil. *Internet Research: Electronic Networking Applications and Policy*, *12*(4), 305–317. doi:10.1108/10662240210438434

Kelly, G. A. (1955). *The psychology of personal constructs*. New York, NY: Norton.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.

Lotzkar, M., & Bottorff, J. (2001). An observational study of the development of a nurse-patient relationship. *Clinical Nursing Research*, *10*(3), 275–294. doi: 10.1177/10547730122158932

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

Militello, M., Schweid, J., & Sireci, S. G. (2010). Formative assessment systems: Evaluating the fit between school districts' needs and assessment systems' characteristics. *Educational Assessment, Evaluation and Accountability*, *22*(1), 29–52. doi:10.1007/s11092-010-9090-2

Parrinello, G. (2013). The city-territory: Large-scale planning and development policies in the aftermath of the Belice valley earthquake (Sicily, 1968). *Planning Perspectives*, *28*(4), 571–593. doi:10.1080/02665433.2013.774538

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Pedersen, M. H. (2012). Going on a class journey: The inclusion and exclusion of Iraqi refugees in Denmark. *Journal of Ethnic and Migration Studies*, *38*(7), 1101–1117.

Piaget, J. (1974). *Biology and knowledge*. Chicago, IL: University of Chicago Press.

QSR International. (2012). NVivo (Version 10) [Computer software]. Melbourne, Australia: QSR International.

Scheib, J. W. (2003). Role stress in the professional life of the school music teacher: A collective case study. *Journal of Research in Music Education*, *51*(2), 124–136.

Scribner, S. P., & Crow, G. M. (2012). Employing professional identities: Case study of a high school principal in a reform setting. *Leadership and Policy in Schools*, *11*(3), 243–274.

Shen, J. (2007). Approaches to international industrial relations in Chinese multinational corporations. *Management Revue*, *18*(4), 410–426.

Stake, R. E. (2006). *Multiple case study analysis*. New York, NY: Guilford Press.

Tellis, W. (1997). Introduction to case study. *The Qualitative Report*, *3*(2). Retrieved from http://www.nova.edu/ssss/QR/QR3-2/tellis1.html

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Thousand Oaks, CA: Sage.

15

Action Research

James Jackson, EdD

Introduction

This chapter describes action research, a practical approach to research that involves researchers and key stakeholders as partners. The chapter opens with a description of what action research is and is not and then follows with its historical origins and philosophical underpinnings. To highlight the practical nature of action research, it is then compared to traditional experimental research. The chapter goes on to describe the action research four-phase cycle, including diagnosing, planning action, taking action, and evaluating action. A description of the role of the action researcher and the ways action research is unique follows; in addition, the chapter includes discussion of the role that stakeholders play in executing the research project. What follows is an explanation of various aspects of the methods of action research, including data collection and analysis, sample size, and ways to establish rigor.

Background

Action research is action and research in the same process. It is not *action for research* (doing in order to increase understanding), nor is it research for action (increasing knowledge in order to apply it at a later time). Rather, it is a coming together of two purposes in a single project or process. Action research is not a research method, because many methods of data collection may be used in action research projects. It involves doing research and acting to change situations at the same time (Hughes, 1997).

Action research involves reflection. It is "systematic self-reflective scientific inquiry by practitioners to improve practice" (McKernan, 1996, p. 5). It involves problem solving and change. Herr and Anderson (2005) described action research as

> oriented to some action or cycle of actions that organizational or community members have taken, are taking, or wish to take to address a particular problematic situation. The idea is that changes occur either within the setting and/or within the researchers themselves. (p. 4)

Many definitions of action research emphasize collaboration, describing it as "systematic inquiry that is collective, collaborative, self-reflective, critical, and undertaken by the participants of the inquiry" (McCutcheon & Jurg, 1990, p. 148). Others describe action research as "a form of collective, self-reflective enquiry. . . . Groups of participants can be teachers, students, principals, parents, and other community members—any group with a shared concern. The approach is only action research when it is collaborative" (Kemmis & McTaggart, 1988, p. 5).

Action research is not an "easy way out." It is not a qualitative methodology or a way to describe something unplanned. It is not research that is done in order to apply it to action at some later date, nor is it a single, clearly defined approach. You are not doing action research unless you are taking action in the study, changing a situation, reflecting on the actions taken, and engaging in systematic inquiry.

The use of action research has many possible benefits. For example, action research develops a culture of collaboration and research-based practice, improves the skill level of all participants, provides practical answers to real-world problems, and allows for the application of many tools used in research and evaluation. Action research is intentional and promotes positive social change in local settings. It can be used in a multidisciplinary environment in which participants have varied levels of research or evaluation training, and it is context specific and context dependent, focused on local problems and resources rather than the generalization of results.

Origins of Action Research

Kurt Lewin coined the term *action research* over 60 years ago, when he stated:

> The research needed for social practice can best be characterized as research for social management or social engineering. It is a type of action-research, a comparative research on the conditions and the effects of various forms of social action, and research leading to social action. Research that produces nothing but books will not suffice. (Lewin, 1946, reprinted in Lewin & Lewin, 1948, pp. 202–203)

Jane Addams, the first female Nobel Prize winner and a dedicated social reformer, was an early action researcher, although she is better known for working toward women's suffrage and founding the Chicago settlement house, Hull House, at the turn of the 20th century (Lundblad, 1995; Stanford Encyclopedia of Philosophy, 2014). Her data collection in the field, which predates Lewin, was used to persuade the government to intervene on matters of sanitation, child nutrition, infant care, and a wide variety of social conditions. Furthermore, an examination of her writings reveals terminology, philosophy, and practice that represent a true reflection of action research. Many consider Jane Addams the mother of American social work.

Philosophical Underpinnings of Action Research

In more than any other approach to research, the philosophy and values that form the foundation of action research are essential to the development of processes and procedures—and the interpretation of success.

Action research is pragmatic. Collaboration and empowerment are essential components of any group action research effort. And without these vital principles, true action research does not exist. In action research, the project outcome is only one measure. Investigation into the creation of the outcome is equally, if not more, important.

To understand the difference between the philosophical foundation of action research and other approaches, it is useful to compare the goals and the outcome associated with the construction of two types of houses: one built by a nonprofit such as Habitat for Humanity and the other built by a typical contractor. Both the nonprofit and the contractor could build identical houses, so the outcome would be the same. But the primary goals would be different. For the contractor, the main goal is to make money and earn a profit. For the nonprofit, the main goal is not profit. Instead, the goals of the nonprofit might include providing a needy family with a place to live, instilling personal pride of ownership through sweat equity, empowering communities, developing skills, changing the culture of the neighborhood, and providing an opportunity for volunteerism. Through a more stable environment, these families' health and sense of security grow, leading to long-term changes such as improved education and job opportunities.

Action research is typically a value-laden endeavor. Some of the values that can be expressed through action research include empowerment, responsiveness to a changing environment, collaboration, democracy, and reflection. Sound decision making based on accurate data and timely information from a variety of sources is valued in doing action research, as is a high regard for the environment, culture, and empowerment where all participants benefit. Through action research, social responsibility can be expressed on personal, community, and corporate levels.

Unlike other approaches, action research is not the implementation of a design with a specific outcome. It is a series of steps that include inquiry, cycles of planning, observation, action, and reflection, with each step building on the last. For action researchers, research is an integral part of practice that involves a continuous improvement model. A cycle generally begins at a planning or diagnosing phase, although action research can be implemented at any point in time in a program, personal practice, or stage of development in an inquiry. Action research is intentional—it is a planned intervention that is informed by data and includes a process that is examined and reflected on.

Action Research Compared With Experimental Research

To understand what action research is—and is not—it may be useful to compare it with experimental research. Empirical laboratory and controlled studies try to control most elements and extraneous variables in order to demonstrate cause and effect. Researchers cannot control such conditions in the field when conducting action research. Action research applies the same principles examined by empirical research methods, but applies them in the field (see Box 15.1). It welcomes all variables as sources of information and takes them into account, rather than ignoring or controlling for them.

BOX 15.1

Discipline-Specific Applications of Action Research

Action research is represented in several disciplines, including business, education, sociology, psychology, organizational development, and leadership. It has deep roots in the politics of the oppressed, feminism, antiracism, and community transformation. Individuals and organizations in the fields of medicine, education, business, marketing, government programs, and international development are actively practicing action research today. Some research topics include the following:

- An inquiry into whether a pragmatic-oriented action research methodology may enhance the value of human capital of a business in France (Cappelletti & Baker, 2010)
- An exploration of how a high school teacher used a new way to engage his students in the required reading of classical texts in Greece (Tsafos, 2009)
- An investigation of the relationships between action research and general systems theory using the case study of environmental education activities in a rural community in Brazil (Berlinck & Saito, 2010)
- An analysis of the factors contributing to the longevity and achievements of an action group formed to improve dignity and respect toward patients and other service users at a hospital in the United Kingdom (Crow, Smith, & Keenan, 2010)
- A collaborative study exploring how teachers implemented a Spanish for Heritage Speakers course in Utah, an English-only state (Coles-Ritchie & Lugo, 2010)
- An investigation of aboriginal women, overrepresented in prisons in Australia, requiring "an Indigenous informed conceptual framework utilising a decolonising research methodology inclusive of enduring community and stakeholder dialogue and consultation" (Sherwood & Kendall, 2013, p. 83)
- An application of participatory action research to establish safe drinking water supplies in Bangladesh (Rammelt, 2014)

Researchers conducting empirical studies tend to believe that they leave the environment unchanged while examining and manipulating only the independent variable(s). Action researchers conduct their work with the intent of making positive changes. Empirical research has the experimenter separate from the process to avoid bias and influence. In action research, the researcher is often also a participant whose influence is necessary for the project to succeed.

Both empirical and action research approaches are rigorous and acceptable methods of inquiry with professional, journal, and academic support from many sources.

Design Considerations

This section of the chapter provides information on the types of research topics and questions explored with action research, along with examples illustrating the use of action research with other methodologies. In addition, this section discusses the role of the researcher, as well as sample size and ethical considerations in doing action research.

Roles for the Action Researcher

Topics explored with action research vary, so the roles of the researcher also vary. The list below presents a range of possible roles that the action researcher can take, in terms of being an insider or an outsider. Simply put, an insider is a member of the organization or community under study; an outsider is not.

- Insider self-study: Researcher works alone or with other insiders to study his, her, or their own practice.
- Insider collaborates with insider: Researchers work together to raise consciousness within an organization or community or to form inquiry/study groups with the goal of engaging other members and influencing organizational change.
- Insider in collaboration with outsiders: Researchers who are members of an organization or community invite outsiders, such as consultants, to collaborate on research.
- Reciprocal collaborations: Teams of insiders and outsiders work together in collaborative forms of participatory action research; power relationships are equitable, with insiders and outsiders respecting and learning from each other.
- Outsider(s) in collaboration with insiders: The researcher is an outsider who initiates research projects with members of an organization or community with the goal of community empowerment or organizational learning.
- Outsiders study insiders: Researchers are outsiders who conduct academic research on action research methods or projects; researchers are not participants in the action itself (Herr & Anderson, 2005, pp. 32–33).

For example, Tsafos (2009) had the role of an insider studying his own teaching practice through an action research project. In the project by Crow et al. (2010), Keenan was an insider in collaboration with outsiders. As a nurse responsible for practice development at a local hospital, Keenan collaborated with outsiders Crow and Smith, who were local university researchers. The action research project conducted by Coles-Ritchie and Lugo (2010) included reciprocal collaborations, in which the researchers formed an insider–outsider team to help one of the researchers implement a Spanish for Heritage Speakers course at the high school where she teaches. Cappelletti and Baker (2010) acted as outsiders in collaboration with insiders. As researchers based at universities in Lyon, France, and New York, they collaborated with a French company's management and staff to help the organization learn how to improve its human resources practices. Berlinck and Saito (2010) also served as outsiders in collaboration with insiders. Invited by a planning committee on water resources in Brazil, the team of researchers planned and facilitated discussions involving local administration and community members.

Stakeholders in Action Research Projects

In collaborative action research, stakeholders have an interest in, can affect, or can be impacted by the change that is being addressed. Dick (2002) made the distinction between direct and indirect stakeholders. Stakeholders are described as being directly affected by what is happening or by what is going to happen. For example, in community consultation for the purpose of traffic design, direct stakeholders include residents in the area where the roads and transport are to be changed.

Indirect stakeholders have a stake in the project as well; however, their stakes may not be as obvious or may be harder to identify. In traffic design, indirect stakeholders include the motorists or others who travel through the area where consultation is to take place. Less obviously, the residents of other suburbs are stakeholders, as they will experience increased or reduced traffic because of changes to traffic design in the researched community. It is sometimes difficult to identify all of the indirect stakeholders, and they may identify themselves later by their reactions to the decisions made.

Stakeholders may participate in an action research project at any phase, or they may simply be the recipients of reports and information that result from the project. Stakeholders may provide information, collect data, propose changes, provide input for planning, or engage in any of a variety of tasks and responsibilities that encompass an action research study or project. The collaborative interactions with stakeholders are an essential element of action research.

Research Questions for Action Research

Action research is not appropriate for questions that you want to generalize. In an action research project, the research context determines the research question(s). A general research question might be "How can we make effective change that can be beneficial?" The research question or purpose should somehow reflect this notion of change. Some specific examples of research questions are listed below.

- First, is the quality of the human resource function a determining factor in the sustainable development of the value of human capital? . . . Second, is the value of human capital an essential element in the competitiveness of a company, and how can this be revealed using an AR methodology? (Cappelletti & Baker, 2010, p. 212)
- In what ways do pedagogical decisions in SHS [Spanish for Heritage Speakers] classrooms necessitate a shift in one's discursive alignment? (Coles-Ritchie & Lugo, 2010, p. 200)
- How and in what ways were a teacher-educator and Spanish language teacher able to navigate the secondary school structure to implement an SHS course by applying CTAR [critical teacher action research]? (Coles-Ritchie & Lugo, 2010, p. 200)

In addition, here are some examples of research purposes stated not as questions, but as goals or aims:

- I aimed to develop pedagogic strategies that would nurture student involvement in the learning process. . . . My ultimate goal was to transform students from passive recipients to active readers. (Tsafos, 2009, pp. 198–199)

- The goal of this study is to present a reflection on the relation between action research spiral and the general systems theory (Bertalanffy 1976), specifically with regard to the hierarchy of systems (Berlinck & Saito, 2010, p. 144).

Keep in mind that research questions, goals, or purposes can evolve during the process of action research. For example, Crow and colleagues' (2010) research project began as a short-term endeavor in which the researchers delivered three modules of an education program instructing hospital staff about ways in which to foster dignity and respect in the health care setting. The education program led to the formation of a Dignity and Respect Action Group, whose membership has grown beyond hospital staff to include volunteers, patients, and anyone in the community using the hospital's services.

In the final incarnation of their study, the Crow et al. (2010) acknowledged this evolution of their project: "In retrospect, we can see that our own project began as a 'professionalising' type [aiming to improve practice] and developed further into an 'empowering' one as we sought to empower the 'grass-roots' staff and service users of the hospital" (p. 56). The aim of the resulting study then became to discuss the factors that the researchers believed helped them to "sustain a longer-term project than originally envisaged" (Crow et al., 2010, p. 57).

Finally, recall that action research is more of an approach than a methodology, so it often incorporates other, more clearly defined research methodologies or approaches. For example, Cappelletti and Baker's (2010) action research project examining the value of human capital in a corporate setting used a case study approach. In the action research project of Coles-Ritchie and Lugo (2010), the data presented were part of a larger and longer ethnographic study.

Action Research Cycle as Methodology

From the four-step cycle of Lewin (1946)—plan, act, observe, and reflect—to the many different conceptualizations of the action research cycle that exist today, common elements of an action research cycle include stages of observing, planning, taking action, and evaluating the process that is taking place through reflection. Action research has been described in three-, four-, and seven-step cycles, as well as in even more complicated models that describe cycles within cycles.

Some researchers find the methods for their studies in the steps of an action research cycle. Crow et al. (2010) views action research as a "methodological approach" that has "a cyclical nature involving identification of a problem, action planning, action taking, evaluating and specifying learning and understanding" (p. 56). Tsafos (2009) organized his research around the "standard multi-cycle action research design: plan-act-observe-reflect" (p. 199).

Some researchers object to the notion that action research requires a strict methodology. Berlinck and Saito (2010) viewed action research more as a research "conception" than as a method. To give structure to their project, they used the steps of the action research spiral not as a "set of methodological stages to be followed, but rather as a general directive," with each step corresponding to "different stages of approach." To view action research as a method represents a "reductionism . . . whose ghost must be repelled" (p. 148).

The Role of Theory in Action Research

Theory is important in action research, but there is not a unified opinion among leading writers and practitioners as to the source of the theory or how, when, or for what reason it is applied. Just as there are different approaches (individual vs. collaborative groups), settings (your individual practice vs. inside your workplace with others vs. working with “outsiders”), and methodologies (quantitative vs. qualitative vs. mixed methods vs. evaluation) that share space under the action research umbrella, the situation, orientation of the researchers, knowledge and resources of the local participants, and object of the research can determine theory’s role, form, and importance.

Action research is primarily a framework for exploration and problem solving that theories inform. Researchers often identify theories as the iterations of the action research process take place. Researchers can include traditional theories as participants seek ways of understanding what they are experiencing in their continuous effort to improve. More often than not, theories or frameworks for understanding are unique to the situation and participants.

Like other types of researchers, action researchers usually attempt to outline the theoretical context or framework that informed their project. For example, in their French corporate case study, Cappelletti and Baker (2010) explained why they chose a pragmatic action research framework, as opposed to a critical one (p. 215), as well as how this choice drove their decision to use socioeconomic approach to management (SEAM), which they believe is a French example of a pragmatic approach.

In his project to engage students in classical Greek readings, Tsafos (2009) identified a framework of “praxeology” rather than educational theory. In this way, theory would not drive practice, allowing the process of the action research to remain dynamic (p. 198).

Theory is important to the action research process, but at this time, there is no single, clear, accepted definition of the role of theory in action research, in part because action research is not a monolithic concept or methodology; however, researchers generally accept that theory—whether generated by the participants as an emerging understanding or supplied as a framework by outside collaborators—is an important part of the process.

Ethical Considerations

When doing action research, certain ethical issues, such as the insider/outsider distinction and researcher bias, may pertain in unique or particular ways. Insiders may be in a position of authority or have multiple roles. They may affect issues with participants, interrelated with trust, honesty, and pressure, to give positive results. In action research, the researcher is also a participant, so establishing objectivity and guarding against bias will be different than in other research approaches. It is important to recognize that in participatory action research, in dealing with groups, there are often people with different knowledge levels about confidentiality and other ethical issues.

For example, the action research study on the development of a group fostering dignity and respect at a hospital in the United Kingdom cited several ethical challenges. For example, a small team of people founded the action group and the researchers were aware of the "danger of over-reliance on a few individuals," particularly when it came to who should chair the group meetings. To empower group members, the founding members adopted a "revolving chair" system in which hospital staff members shared the roles of chair and vice-chair (Crow et al., 2010, p. 60).

Another concern was that, over time, hospital staff members who were part of, or perceived to be part of, senior levels of management would replace the people who filled key roles in the group. If this replacement of people in key roles happened, others might have perceived the action group as "just another 'committee' with a 'top-down' agenda" (Crow et al., 2010, p. 60). To prevent this perception, the group made continual efforts to maintain the group as a "'bottom-up' democratic and participative entity" (p. 62).

Thus, the values that provide the foundation for action research are the same values that can help researchers and participants negotiate any ethical concerns that arise.

The Process: How to Do Action Research

This section of the chapter presents information on data collection and analysis in action research. In addition, this section explains how to ensure quality and the importance of rigor in an action research project.

The Action Research Cycle

The process of action research involves cycles. Several different models exist for conceptualizing the cycles. One model that is prescriptive and easy to follow is the four-step cycle presented in *Action Research in Human Services* (Stringer & Dwyer, 2005). The four steps are *diagnosing*, *planning action*, *taking action*, and *evaluating action*.

Diagnosing. The first step in this model of the action research cycle is diagnosing. You begin a project with broad questions or issues that you think are important to you or your organization. Most of these issues will be complex and require a systematic diagnosis. Diagnosing involves outlining broad questions, issues, or problems; understanding the context of the project; identifying the people who can help solve the problem and make changes; establishing collaborative relationships with the stakeholders; evaluating whether the goal is worth the effort; and determining whether resources are available to pursue the change. Diagnosis is a structured and comprehensive process. It is also fluid, holistic, and dynamic. The goal is to create an accurate picture as a starting point—not to make a final decision about what the problem is.

Planning Action. The second step in action research is planning action. It is important not to rush this process. Planning must reflect with whom you work, where you are implementing the project, and what kinds of resources are available. This includes the political environment as well as many other considerations.

Planning action involves setting the stage. This means developing the research questions and analyzing and prioritizing issues. It also requires deciding on data collection techniques, reviewing the feasibility of the project, developing implementation strategies, and preparing a research schedule.

Planning action also involves reflection. Engage in narrative reflection with the research team, stakeholders, outside experts, and others. It is important to contemplate the information and critically analyze it to transform the problem or issue. Reflection is not a one-time step. After the completion of an activity, the research team must review the activity again. Critical reflective inquiry is a key component of action research.

Addressing politics also requires a plan. When you begin an action research project, you are likely to encounter political conflicts. Some people in your organization may be rigid and against any type of change, or there may be members of your organization who try to block, stop, or take credit for the project. You must be prepared to manage the various political forces and ethical situations that may occur.

In addition, you must plan and design the study carefully and ethically. Do not select instruments or methods that could produce misleading conclusions. Strive to use sound, ethical practices in the project, such as appropriately dealing with people and accurately reporting results.

Taking Action. Taking action is, of course, a key component of action research. Taking action involves gathering data, implementation of action, and organizational change. In terms of gathering data, there is no single, clear-cut methodology to use in action research, nor is there one research design that is best for all situations. It is advisable to use more than one strategy and more than one source to answer the research questions. Match the research questions to the data collection methods. Most action researchers will use multiple data collection methods and sources so they can compare the results through triangulation.

Implementation is the biggest challenge. In an action research project, cycles of the process occur simultaneously—participants are constantly observing, reflecting, and taking action. You need to maintain the team's energy and provide support so implementation will be successful. Stakeholders need to assist in implementing the action plans. The success of the action research project is in the hands of the people involved in the project.

Action research can be very effective in implementing organizational change because a group of people is working to resolve an issue or problem; there is not just one person proposing a solution. You do not have top management directing how to solve a problem. Instead, in action research, you have a team of motivated stakeholders who have a clear vision and strategy. They are not afraid of change because they are the ones identifying both the need and the force that will take action to effect change. Numerous theories about change management exist, so it is important to focus on one that the team can share.

Evaluating Action. Most of the evaluations that occur in action research will not have a set criterion to measure against. Stakeholders will evaluate outcomes mainly on what they identify as acceptable standards. Evaluating action involves evaluating the program, the process, and the larger context.

In evaluating the program, the question to ask is "Did the project achieve the desired results?" To determine achievement, review all of the processes and outcomes by assessing and describing their effectiveness. This process may provide reinforcement for the changes implemented, cause you to take the project in a new direction or abandon it, create or spiral into new research cycles, or determine that the project is complete. The goal is to examine the changes to determine whether those changes are the desired outcomes of the stakeholders.

In evaluating the process, keep in mind that action research is not just about accomplishing tasks; it is also about improving systems and people. Feelings of empowerment, improved communications, teamwork, lessened negative political forces, and future productivity are all intended outcomes of action research. This approach can result in changes larger than the immediate objectives, which is what sets action research apart from traditional empirical methods of research. As Kurt Lewin maintained, we are not doing research for the sole purpose of writing papers.

Do not forget the larger context of an evaluation. Action research is not a linear process; it is a spiral of congruent cycles. In action research, you examine where you have been, make decisions about evaluation objectives, develop a plan, review the progress, and make decisions about future research cycles. You should examine the larger context by asking questions such as "How has the organization changed? What greater social good has evolved? What have we learned about ourselves in the process? What other resources will help in future research cycles?" Considering the larger context will help you evaluate the impact the actions in the project have made overall. Examination of the larger context is critical to producing meaningful research.

Remember, "knowledge is never static or complete; it is in a constant process of development as new understandings emerge" (McNiff & Whitehead, 2002, p. 18). The action research cycle will continue to spiral into various concurrent revolutions, producing multiple research cycles. Each cycle will go through the processes of diagnosing, planning, taking action, and evaluating. Action research focuses on processes—not just outcomes. It can range from a very informal to a formal study, and it can vary in duration.

Action research does not have to be complicated to be effective. You can start with a smaller project, and as you grow as a researcher, you can move to more complicated projects. Do not have a goal of making the project an extensive study. A concise study can be as beneficial as—or even more beneficial than—a complex study. Think of action research as a way to be creative in solving problems or issues and in making positive changes. It is a process in which you can pursue change and understanding at the same time.

Data Types for Action Research

Action research is not restricted in the types of data it uses to provide needs assessment, program evaluations, or research other than the ability of the participants to correctly and ethically collect, administer, and interpret the results. Typically, action research uses archival data, surveys, interviews, focus groups, and other common social science instruments that are both standardized and locally prepared; however, physical artifacts and measurements can also be included as well as performance measures related to the topic under study.

Based on the research questions, goals and objectives, and sophistication of the participants, identify what kind of data needs to be collected and what processes to use. Sources of data for action research include all the participants. Keep in mind that there are two streams of data collection: One is the data used to answer the research question, and the other relates to the collaborative process. One type of data that you should always collect should be about whether the collaborative process was a success or failure.

For example, the purpose of the Tsafos (2009) study was to explore ways of engaging students more in their study of classical texts. He collected data from three sources: himself as teacher and researcher, his students, and a critical friend, who was also a teacher of the same grade. His data consisted of a journal, field notes, and audio recordings. From the students, he collected questionnaires and conducted semistructured interviews. The critical friend contributed observations and field notes. The data included opportunities for all three data sources to evaluate whether the new learning process enacted in the project had positive results.

In collecting data, be sure to maintain standards. Doing action research does not mean you can ignore standard levels of practice. Instruments should be reliable and valid. You should maintain integrity in your research. Use only those instruments that you are qualified to administer and interpret. If you lack the expertise, engage an expert or consultant. When preparing local instruments such as surveys, follow recognized guidelines to encourage proper construction, analysis, and safeguards against bias. Make sure to properly protect and reasonably inform participants about the use of surveys and other instruments. Notify participants about whether and how you will protect their identities. Also make sure participants know they can ask questions. Although practitioners often conduct action research in the field, most places of employment, educational institutions, governmental agencies, and other locations for studies have guidelines for conducting research that researchers must follow.

For example, the Cappelletti and Baker (2010) study used the SEAM approach. The researchers outlined a data collection protocol based on the HORIVERT method, which is a tri-axial methodology used by more than a thousand organizations around the world. The HORIVERT method is a way of organizing those involved in the process of change along three axes: policy, tools, and participative change. For each axis, the researchers identified appropriate groups of people within the company under study from whom to collect data. Data collection underwent several phases over a year and a half and consisted of focus groups as well as 175 qualitative, quantitative, and financial interviews. The initial interviews revealed areas of company dysfunction, after which task groups formed to propose solutions. A piloting group composed of the chief executive officer (CEO) and human resources management then supervised which solutions to implement.

Data Analysis Techniques

Data analysis techniques will vary depending on the purpose and specific methodology used for the action research project. You must take care to analyze data correctly and ethically. This level of care is particularly important when the researcher organizes his or her research using the steps of the action research cycle or takes a broad approach, such as a case study or ethnography.

Tsafos (2009), who organized his research around a four-step action research cycle, analyzed his data by delimiting the point of view of each data source; he then conducted reflection meetings with his critical teacher/friend to discuss the action. Tsafos next held a meeting to share the student questionnaires with all staff members and to obtain a wider point of view. Then Tsafos hosted a discussion with both students and teachers. The final analysis drew upon the researcher's journal, the critical friend's final report, and the student interviews.

The data for the Coles-Ritchie and Lugo (2010) project, which were part of a larger ethnographic study, consisted of field notes of classroom observations, school board notes, student records, and email messages. The researchers analyzed the data together to develop emerging themes. Lugo had written reflections and journal entries, and the research team organized these reflections and entries to represent the different phases of planning and implementing the Spanish for Heritage Speakers course. The reflections explain Lugo's journey as a teacher trying to enact change.

Sample Size Considerations

In action research, the sample size depends on the context of the project. Sample size is dependent on the methodologies selected for inclusion in the action research process or cycles. An action research teacher working without a team may consider focusing on improvement with an individual child or the entire classroom. When using a traditional quantitative or qualitative method within action research, follow the rules for sample size for that methodology. For example, the Berlinck and Saito (2010) case study on water resource management had a sample size of one because it focused on a single rural community in central Brazil. Keep in mind that sample size is different than the number of stakeholders. In action research, all participants in the action research project are not necessarily members of a sample.

Rigor in Action Research

Rigor is an attribute of any good research project regardless of the approach used. Rigor means thoroughness and adherence to the rules representing the best practice in the methodology under examination. Some people have criticized action research as lacking sufficient rigor to be a serious academic research approach. This criticism may be correct for some individual action research efforts, but people can apply the same criticism to poorly planned and executed empirical research.

Researchers determine rigor by strictly applying standards that are suitable for a specific methodology; the expectations are not identical for every approach. When assessing the rigor of an action research project, researchers focus attention on the application of the best methodologies for the situation and the questions the research answers. Because action research can range from the exploration or testing of a theory to applied problem solving, rigor defined as precision (as in empirical research) needs supplementation by rigor defined as thoroughness. Thorough research has extensive problem analysis, an extensive literature review, a careful evaluation of assumptions, alternative (inclusive) methods for problem solving, and/or multiple methods of measurement.

Action research is not more or less rigorous than empirical "scientific" research. Action researchers just describe rigor in different terms. In many ways, good action research often combines the rigor associated with empirical quantitative designs with that of qualitative methodologies. For example, *dependability*, a criterion used by qualitative approaches, corresponds to the term *reliability* in quantitative approaches. The corresponding term that an action researcher might use to describe the same criterion is *consistency* (Krefting, 1991). Ernie Stringer (1996) agreed with Lincoln and Guba (1985) that action research is identified as trustworthy by establishing six conditions. The six conditions of trustworthiness are credibility, transferability, dependability, confirmability, degrees of participation, and utility. Procedures to attain trustworthiness are prolonged engagement, persistent observation, triangulation, participant debriefing, negative case analysis, referential adequacy, member checks, transferability, dependability, confirmability, and participation.

Appropriateness of Action Research for Your Research Project

In determining whether action research is appropriate for your research project, consider the essential characteristics of action research. Ask yourself the following questions:

- What are my core values?
- Is the philosophy and approach of action research consistent with how I view myself?
- Do I have the time to involve stakeholders?
- Do I have the resources to accomplish the scope of the project?
- What are the ramifications of entering in a certain place in the action research cycle?

Consider how you feel about the validity of an action research project: Think about what it would take to convince you of the trustworthiness or validity of a study or dissertation. It appears that the less you can control (account for through empirical methods) in the environment, the more you have to describe in order for the research community to accept your results as valid. It also appears that many or all of the validity concerns in empirical studies are identical to those of action research studies. Do you only trust laboratory-controlled studies or real-life explanations when you make decisions?

Consider how you feel about the acceptability of action research. Acceptability is a different concept than validity and is more closely related to traditions, political forces, and contexts that tend to favor one type of "legitimate" inquiry over another. Differentiate educational inquiry acceptability from in-the-field implementation acceptability. Reflect on your own thoughts about using an action research approach in a dissertation and about using an action research approach in your own work setting.

Summary

Action research involves action and research, reflection, problem solving, and change. It is a way of engaging in systematic inquiry while taking action to change a situation. Action research is not a single, clearly defined approach. It is a pragmatic approach in which collaboration and empowerment are essential components. Investigation of the process of the action is as important as the outcome. Individuals and organizations in a wide range of fields, from education to business to international development, actively practice action research.

Doing action research involves cycles of planning, observing, acting, and reflecting. Many different models of the action research cycle exist, and a study often consists of cycles within cycles. Because action research is more of an approach than a method, it almost always incorporates other, more clearly defined research methods. Topics explored with action research vary, so the roles of the researcher also vary. Researcher roles can range from insiders studying their own practice, to insiders working in collaboration with outsiders, to outsiders conducting academic research on insiders engaged in an action research project.

Theory is important in action research, but there is not a unified opinion among leading writers and practitioners as to theory's role, form, and importance; however, it is generally accepted that theory, whether generated by the participants as an emerging understanding or supplied as a framework by outside collaborators, is an important part of the process. Some ethical issues to consider when doing action research include the insider/outsider distinction, pressure to give positive results, and researcher bias, especially because the action researcher is also often a participant.

Data collection and analysis techniques will vary depending on the purpose and specific methodology used for the action research project. You must take care to analyze data correctly and ethically. Doing action research does not mean you can ignore standard levels of practice—good action research must be rigorous, thorough, and trustworthy.

KEY SOURCES

Coghlan, D., & Brannick, T. (2005). *Doing action research in your own organization*. Thousand Oaks, CA: Sage.

Lewin, K., & Lewin, G. W. (Eds.). (1948). *Resolving social conflicts: Selected papers on group dynamics*. New York, NY: Harper & Row.

McKernan, J. (1996). *Curriculum action research: A handbook of methods and resources for the reflective practitioner* (2nd ed.). Oxford, United Kingdom: Routledge.

McNiff, J., & Whitehead, J. (2006). *All you need to know about action research*. London, United Kingdom: Sage.

Sagor, R. (2005). *The action research guidebook: A four-step process for educators and school teams*. Thousand Oaks, CA: Corwin Press.

Stringer, E., & Dwyer, R. (2005). *Action research in human services*. Upper Saddle River, NJ: Pearson Education.

REFERENCES

Berlinck, C. N., & Saito, C. H. (2010). Action research for emancipation informed by Habermas and Hierarchy of Systems: Case study on environmental education and management of water resources in Brazil. *Systemic Practice & Action Research*, *23*, 143–156.

Cappelletti, L. G., & Baker, C. R. (2010). Developing human capital through a pragmatic oriented action research project: A French case study. *Action Research*, *8*, 211–232.

Coles-Ritchie, M., & Lugo, J. (2010). Implementing a Spanish for heritage speakers course in an English-only state: A collaborative critical teacher action research study. *Educational Action Research*, *18*(2), 197–212.

Crow, J., Smith, L., & Keenan, I. (2010). Sustainability in an action research project: 5 Years of a Dignity and Respect Action Group in a hospital setting. *Journal of Research in Nursing*, *15*, 55–68.

Dick, B. (2002). *Stakeholders and participation*. Retrieved from http://www.scu.edu.au/schools/gcm/ar/areol/areol-session04.html

Herr, K., & Anderson, G. L. (2005). *The action research dissertation: A guide for students and faculty*. Thousand Oaks, CA: Sage.

Hughes, I. (1997). *Action research electronic reader*. Retrieved from http://www.scu.edu.au/schools/gcm/ar/arr/arow/rintro.html

Kemmis, S., & McTaggart, R. (Eds.). (1988). *The action research planner*. Geelong, Australia: Deakin University Press.

Krefting, L. (1991). Rigor in qualitative research: The assessment of trustworthiness. *The American Journal of Occupational Therapy*, *45*(3), 214–222.

Lewin, K. (1946). Action research and minority problems. *Journal of Social Issues*, *2*(4), 34–46.

Lewin, K., & Lewin, G. W. (Eds.). (1948). *Resolving social conflicts: Selected papers on group dynamics*. New York, NY: Harper & Row.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

Lundblad, K. S. (1995). Jane Addams and social reform: A role model for the 1990s. *Social Work*, *40*, 661–669.

McCutcheon, G., & Jurg, B. (1990). Alternative perspectives on action research. *Theory into Practice*, *24*(3), 144–151.

McKernan, J. (1996). *Curriculum action research: A handbook of methods and resources for the reflective practitioner* (2nd ed.). Oxon, United Kingdom: Routledge.

McNiff, J., & Whitehead, J. (2002). *Action research: Principles and practice*. London, United Kingdom: Routledge.

Rammelt, C. F. (2014). Participatory action research in marginalized communities: Safe drinking water in rural Bangladesh. *Systemic Practice and Action Research*, *27*, 195–210.

Sherwood, J., & Kendall, S. (2013). Reframing spaces by building relationships: Community collaborative participatory action research with Aboriginal mothers in prison. *Contemporary Nurse*, *46*(1), 83–94.

Stanford Encyclopedia of Philosophy. (2014). *Jane Addams*. Retrieved from http://plato.stanford.edu/entries/addams-jane/

Stringer, E. [Ernest], & Dwyer, R. (2005). *Action research in human services*. Upper Saddle River, NJ: Pearson Education.

Stringer, E. [Ernie]. (1996). *Action research: A handbook for practitioners*. Thousand Oaks, CA: Sage.

Tsafos, V. (2009). Teacher–student negotiation in an action research project. *Educational Action Research*, *17*(2), 197–211.

16

Program Evaluation

Louis Milanesi, PhD

Introduction

In this chapter, I cover the foundational role of stakeholders in program evaluation, followed by an exploration of the logic model of evaluation. As part of the logic model discussion, key differences between process and formative evaluation, as well as outcome and summative evaluation, are described. I then highlight examples of program evaluation across the disciplines. In an analysis of methodological aspects, I explain sample size, data collection and analysis, and issues of quality and also provide resources for researchers conducting program evaluations.

Background

Evaluation is the systematic assessment of the operation and/or outcome of a program. Much like a scientific research study, program evaluation seeks to answer a question or set of questions. The scope and complexity of programs vary widely—from large-scale, national or international projects to smaller scale efforts of local governments or individual organizations. Thus, the scope and complexity of program evaluations also vary widely. Additionally, program evaluations are sometimes encountered as elements nested within broader studies related to policy development or policy evaluation.

The scope and complexity of program evaluations also have implications for use as a doctoral study within a professional doctorate program or as a thesis or dissertation in a Doctor of Philosophy (PhD) program. In fact, applications of program evaluation designs are commonly found in thesis and dissertation research in professional doctorate programs, such as the Doctor of Education (EdD), Doctor of Business Administration (DBA), and Doctor of Nursing Practice programs. Some program evaluations may contribute significant knowledge to the home disciplines and rise to thesis- or dissertation-level rigor. However, not all program evaluations are an appropriate model for a full thesis or dissertation study. If you are considering a program evaluation for your research, consult with your research committee to help make this determination.

To help you fully understand program evaluation, I will address the following components: stakeholders; logic model development (the program theory of change); determination of goals, objectives, or outcomes; needs assessment; process and formative evaluations; and outcome and summative evaluations.

I will also introduce and emphasize five core standards common to evaluation practice: utility, feasibility, propriety, accuracy, and accountability (Yarbrough, Shulha, Hopson, & Caruthers, 2011).

Stakeholders

Collaboration among stakeholders is essential to the success of any program, and the program evaluator should be considered a key contributor in that collaboration. Understanding how other stakeholders view the evaluator's role is an important step in engaging with a project. The perception of degree of the evaluator's involvement in the project is often determined at the time that he or she is brought onto the team.

The expectations of the team also can be reflected in the resources allocated to evaluation activities. The optimal situation occurs when the evaluator is recruited in the early planning stages and is seen as a long-term member of the program team. The least desirable scenario occurs when the evaluator is recruited at the end of the program cycle simply to satisfy a mandated requirement. In the latter case, opportunities for the evaluator to contribute to the program's successful deployment, as well as to collect data that would contribute to a robust evaluation, may be lost. The evaluator may decide not to perform an evaluation if evaluation standards cannot be achieved.

The evaluator's ability to develop and maintain collegial relationships with all stakeholders is critical. The terms *stakeholders* and *partners* are often used synonymously; in this chapter, however, a useful distinction is made between stakeholders who have key roles and investments in a program and partnerships between individuals and organizations that exist outside the scope of the program. Stakeholders include individuals and organizations that are involved in shaping and delivering a program as well as individuals who are involved in sustaining a program and those impacted by the program. Some or all of these stakeholders may be connected with one another through cooperative partnerships that link their organizations. For example, one organization may be a service provider for another, one might be the parent organization of another, or two or more may be connected through contracts.

Such partnerships can serve to bind stakeholders in the execution of the program activities; however, they may also bring background to the program project that can interfere with its execution. Cases when the connections between stakeholders involve only their work on the program being evaluated should prompt specific concerns for an evaluator because stakeholder commitment to the project's activities and timeline may become inconsistent.

Logic Model Development

Evaluators can make significant contributions to the success of a program when they are engaged in the formative stages of its development. One such early contribution involves helping stakeholders formulate a detailed logical model of how the program operates. Logic models are helpful tools in program evaluation and performance measurement. They provide a causal map of how a program, policy, or initiative should work. In many ways, the logic model can be thought of as a program schematic that identifies the program/project resources, activities, inputs and outputs, and outcomes.

A logic model is a visual aid of your program that you can show a potential funder, auditor, or organizational executive. The visual aspect can be appealing because it removes technical language and shows how the components of the program should flow. Furthermore, logic models present how the program will lead to an outcome and why the activities within a program will lead to an outcome. For example, the logic model identifies the required resources (inputs) needed to accomplish each step a program uses to effect change. It also defines what immediate gains (outputs) are expected at each step.

This sequence of incremental change eventually achieves the desired end results of the program (outcomes) and contributes to the program's long-term goals. This approach is sometimes called the program theory of change. More information on the program theory of change is presented in another section of this chapter. Working across different stakeholder groups in the logic model process allows the evaluator to assess the degree to which key groups value the spectrum of outcomes the project should realize. Individuals and groups may come to the project with different goals or agendas; therefore, this process can facilitate communication once all parties' objectives are understood.

Evaluators can ensure that adequate planning has gone into a program as they work with program leadership to examine, understand, or develop the logic model. This process often identifies gaps or misalignments within the planning that the evaluator can use to guide change in the initial plan and prevent the program from foundering.

Moreover, well-constructed logic models will assist the evaluator in constructing the assessment model for the program by pointing to the *what*, *where*, and *when* dimensions of data collection and analyses. For example, were the expected input resources available to support a particular stage? Did that aspect of the program result in the expected output accomplishments? A robust logic model will help facilitate those pointed questions. The W.K. Kellogg Foundation (2004) *Logic Model Development Guide* provides an excellent online resource for anyone interested in reading more about the use and development of logic models. Also, viewing the image results from a web search for *logic model* will provide numerous pictorial illustrations of simple through complex model designs.

Late Engagement: Determining Goals, Objectives, and Outcomes

A logic model provides a roadmap for the individual components of a program evaluation model, but only if a model has been defined. Unfortunately, a program evaluator may be recruited to assist a program that is already under way without the benefit of this detailed operational model. Here, the evaluator may be seen as an external agent who has narrow engagement within the project.

In this case, the evaluator will need to work quickly to gain a clear sense of the needs of different stakeholder groups and the way in which they relate to one another. Then, stakeholders are engaged in a collaborative process in defining goals, objectives, and outcomes. Stakeholders tend to discuss the program initially in terms of some higher level, long-term goals to which it will contribute. Because these higher level goals are usually not amenable to evaluation until the program has been fully deployed, the evaluator should assist stakeholders in defining immediate, intermediate, and long-term objectives and the outcomes that would represent successful achievement of each stage of the program's development.

After defining objectives and outcomes, the focus and scope of the evaluation effort can be negotiated. The evaluator can draw on the commonly accepted standards related to the potential for utility, feasibility, propriety, and accuracy to formulate alternative strategies and designs. Part of this negotiation will include determining what resources are dedicated to the evaluation effort. Resources related to funding, as well as access to data, information, and people, directly impact the feasibility of the proposed evaluation. The importance of resources cannot be overstated, because many evaluation models that appear strong in a written proposal do not hold up when assumptions about available resources are unmet. Evaluators who are engaged late in a program's implementation are often faced with deciding if a meaningful evaluation is possible.

The potential for not meeting certain standards may contribute to evaluators deciding not to engage with a program. Some of these include the following:

- The evaluator may not feel adequately qualified within the program field of endeavor (see utility standard).
- The evaluator is concerned that there is a high risk of misuse or misinterpretation of the results (see utility standard).
- The evaluator believes that the opportunities for evaluation will not equitably represent the cultural or political interests of key stakeholders (see feasibility standard).
- The evaluator perceives a threat to the rights or dignity of program participants or other stakeholders by performing requested activities (see propriety standard).
- The evaluator perceives a limited opportunity to acquire dependable and consistent information to answer the key questions posed within the proposed evaluation model (see accuracy standard).

Needs Assessment

Needs assessments can initially be used to determine the extent of need for new programs in a community, an organization, or even a specific unit in an organization. Thus, a needs assessment directs attention to key opportunities for service and practice improvements. An evaluation question for a needs assessment might be the following: What is the problem and its nature (i.e., its magnitude, antecedent factors)? Some examples might include a community college that wants to conduct a needs assessment to determine current gaps in what programs should be offered, a nonprofit that wants to conduct a needs assessment to determine its focus and allocation of resources, or a human resource department that wants to survey employees within a work unit to assess the need for new training programs. Methodologies for conducting needs assessments can include focus groups, personal interviews, review of administrative data, archival data analysis, or surveys. McKillip (1998) provides an easily accessible and concise guide to conducting needs assessments.

Process and Formative Evaluations

As with many aspects of methods, theoretical arguments arise, and the field of evaluation is no exception. The terms *formative evaluation* and *process evaluation* are often used synonymously to classify evaluation activity. Although there is a distinction between the two approaches, it is important to note they are complementary rather than contradictory.

Process evaluation focuses on documenting and analyzing the way a program works to isolate and understand influences that affect its operation. That is, a process evaluation aims to understand how the program achieves the results it does. A major focus of process evaluation is relating the experiences of the program to others. Therefore, process evaluations are commonly employed to benchmark the approaches used in exemplary programs so that they can be replicated within similar organizations and contexts.

A formative evaluation differs in that it is ongoing and seeks to assist decision makers in refining and improving a program. The collection, synthesis, and feedback of information are continuous. Thus, formative evaluations would be employed to achieve continuous improvement of newly developed programs, especially those that have not yet achieved a desired level of efficiency or consistency. Dehar, Caswell, and Duignan (1993) provided one of the most lucid explanations of the differences between the two types of evaluation typologies.

Two evaluation questions for a formative/process evaluation might be the following:

1. How did the program achieve the observed outcome?
2. Was there a better way to arrive at the outcome?

Evaluators can use both qualitative and/or quantitative techniques to answer these questions.

Outcome and Summative Evaluations

Outcome evaluation and *summative evaluation* are terms that are often used interchangeably in discussing program evaluation, and this interchangeability may contribute to some confusion. A convenient way to draw distinctions between the two is to examine differences between the scope and sophistication of the research involved. Outcome evaluations tend to be focused, whereas summative evaluations tend to be more comprehensive.

Outcome evaluations tend to focus on specific objectives defined for the program. Program objectives tend to be more short term in nature when compared to the long-term goals of the program. Thus, the achievement of program objectives can be immediately measured and evaluated at specific points as the program elements are deployed. On the other hand, overall program goals may not be realized until years in the future.

Trochim (2006) views summative evaluation broadly and subdivides it into several subtypologies of activities. Some examples of summative evaluations include impact analyses and cost-effectiveness or cost–benefit analyses.

Comprehensive summative evaluations can have significant impact on the future of the program and the stakeholders involved. Therefore, impact analyses, cost–benefit analyses, or other politically sensitive evaluations are major endeavors and are usually best performed by experienced evaluators. For example, new evaluators tend to have a professional tunnel vision and sometimes focus on a narrow set of outcomes defined within the intended scope of the program. Frequently, good programs can also generate unanticipated yet significant positive outcomes that extend beyond their original scope. Thus, a program may be undervalued if a novice evaluator fails to capture these unanticipated added benefits.

Another common form of summative analysis is cost–benefit analysis, which attempts to determine the resources required to achieve a specific amount of change. Whereas impact analysis tends to employ a macro-level focus to measure the broad, holistic benefit of a program, cost–benefit analysis often employs a more micro-level focus on costs associated with achieving specific objectives within the program. Such analyses are then used to explore where more cost-effective alternatives may be needed. These studies can result in leaner programs that are more resilient to changes.

Impact evaluations assess the intended or unintended effects of the program. Frequently, an impact evaluation takes place sometime after the program's implementation and addresses the magnitude of any change that has taken place. Impact evaluations usually try to assess the degree of change directly attributable to the program in relation to other considerations. For example, an impact evaluation could assess the degree of change, in relation to an overall problem, that would not have been realized in the absence of the program. Therefore, impact analyses are commonly associated with policy development and especially with policy evaluation.

Evaluation Standards

The Joint Committee on Standards for Educational Evaluation, originally formed in 1975, has been involved in developing and revising core standards for evaluation practice (Yarbrough et al., 2011). Although initially focused on program evaluation within educational settings, the standards are largely generalizable to a variety of evaluation contexts. Recent publications have aligned the core standards to application for personnel evaluation and student evaluation.

The core themes of the standards include the following:

- Utility standards are intended to increase the extent to which program stakeholders find evaluation processes and products valuable in meeting their needs.
- Feasibility standards are intended to increase evaluation effectiveness and efficiency.
- Propriety standards support what is proper, fair, legal, right, and just in evaluations.
- Accuracy standards are intended to increase the dependability and truthfulness of evaluation representations, propositions, and findings, especially those that support interpretations and judgments about quality.
- Accountability standards encourage adequate documentation of evaluations and a meta-evaluative perspective focused on improvement and accountability for evaluation processes and products.

Comparing Program Evaluation With Other Approaches

Because program evaluation largely involves employing research methods to solve applied, real-world problems, it is similar to other research designs and approaches such as case studies, cross-sectional survey research, quasi-experimental between-group designs, longitudinal single-group designs, and longitudinal quasi-experimental mixed designs.

Although program evaluation structures often contain one or more of these familiar research designs, they usually differ from academic research regarding intended external validity. Whereas academic research tends to be broadly generalizable by design, program evaluations usually tend toward maximizing the utility of their results to align to the specific social environmental demands within which the program needs to operate. Thus, the extent to which findings from a specific program evaluation are generalizable to other programs depends on how closely those programs and their environmental contexts resemble the program being studied.

Benefits of Program Evaluation

The practice of program evaluation provides numerous benefits for the organization. It can serve as a foundation for program planning, or it can be a way to identify gaps in current practice and needs for new programs or additional services within current programs. Program evaluation can help determine the feasibility of proposed programs or projects or improve the functioning of current programs to increase effectiveness and efficiencies. Finally, program evaluation can provide evidence of accountability.

Applied research skills also benefit the evaluation practitioner. Program evaluation represents a field of practice in which practitioners can translate much of their education in research methods into valuable professional skills in applied practice settings. For example, the evaluation practitioner should be able to quickly formulate the most appropriate research questions the evaluation will address and immediately recognize whether quantitative or qualitative data are required to answer each question. Perhaps the program has archived data related to the program delivery; evaluation practitioners should be able to determine the quality of these data as well as whether and how they can be incorporated into the overall evaluation design.

Program Evaluation in Various Disciplines

The practice of program evaluation lends itself well to serving a wide spectrum of disciplines and professions, including education; business; health care; public policy; and psychological, social, and human services.

Education. Evaluation in education has a long and rich history dating to the 19th century. More recently, the relevance of program evaluation has been underscored due to increased emphasis on research-based decision making to improve curriculum and educational processes, federal grant requirements, expectations of accreditation agencies, and efforts to align professional development activities to current organizational challenges.

Business. Businesses employ a variety of research techniques, including program evaluation. The data-driven decision-making processes employed in many quality assurance models share similar foundations with program evaluation. The performance of both internal and external programs is commonly practiced in business and has parallels with evaluating the performance of groups and individuals.

Health Care. In the field of health care, the application of program evaluation to support better health spans organizations of all sizes in both the public and private sectors. Large government institutions like the Centers for Disease Control and Prevention promote the use of evaluation practices and studies to identify best practices in preventing disease; large private-sector health care providers conduct ongoing evaluations of practices, treatment outcomes, and patient compliance; and smaller local clinics evaluate allocation of scarce resources for patients' most immediate needs.

Psychological, Social, and Human Services. Successful social service programs across all disciplines are usually those that effectively employ program evaluation techniques and strategies. Commonly financed by increasingly scarce public funding, successful programs capitalize on their ability to establish that significant needs for services exist, align their program planning to meet those needs, and document the most effective and efficient means of delivering services.

Public Policy. Public programs are often closely tied to policy decisions. Summative evaluation strategies, such as impact analyses and/or cost-effectiveness analyses, are used to assess policy decisions.

Program Evaluation Alignment and Design

The research literature has numerous examples of research topics best explored with program evaluation. Some of these are described below.

- Need for service programs: needs assessments (Altschuld & Kumar, 2009; Barnett, 2012; Soriano, 2013; World Health Organization, 2000)
- Efficiencies of service program processes: formative/process evaluations, cost–benefit analyses (Coulon et al., 2012; Saunders, Evans, & Joshi, 2005; Steckler & Linnan, 2002; Wholey, Hatry, & Newcomer, 2004)
- Effectiveness of service program outcomes: outcome evaluations, impact analyses (Mohr, 1995; Rossi, Lipsey, & Freeman, 2003; Schalock, 2001; Williams-Reade, Gordon, & Wray, 2014)
- Efficiencies of educational intervention processes: formative/process evaluations, cost–benefit analyses (Davis, Orpinas, & Horne, 2012; Jimison et al., 1999; Oakley, Strange, Bonell, Allen, & Stephenson, 2006)
- Effectiveness of educational interventions: outcome evaluations, impact analyses (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Paul, Yeats, Clarke, Matthews, & Skrabala, 2015)
- Evaluation of instructional curriculum: formative/process evaluations, outcome evaluations, impact analyses, cost–benefit analyses (Volkwein, 2011)

Program evaluation can be amenable to doctoral research whose focus is on applied research. However, it is important to note that program evaluation may not readily apply to the academic model of a specific discipline. A thesis- or dissertation-level research question may not always be embedded within a program evaluation.

Sometimes, the applicability of program evaluation to dissertation and thesis research projects may be a question of scope. An applied program evaluation that encompasses state, regional, or national data can add vital information to the research literature in the field. Students who have questions about whether program evaluation is an appropriate strategy should consult members of their research committee for guidance.

To judge problems that can be addressed in a doctoral-level project, determine whether or not the research/practice problem is

- Justified. Is there evidence that this problem is significant to the professional field?
- Grounded in literature. Can the problem be framed in a way that will enable the researcher to either build on or counter the previously published findings on the topic?
- Original. For research doctorates (PhD), does the problem reflect a meaningful gap in the research literature? For the professional doctorates (EdD and DBA), does the problem describe a meaningful gap in practice?
- Amenable to scientific study. Can a scholarly, systematic method of inquiry be applied to address the problem?

Discuss these aspects of the research problem with all members of a research committee to ensure a consensus that it is reasonable to argue that a specific program evaluation can meet the expectations of a doctoral project (see Box 16.1).

BOX 16.1

Questions for the Evaluator

The most basic program evaluation question is as follows: Does the program work? There are, however, many other questions an evaluator might ask.

- Is there a need for a new program?
- What would a new program require to yield success?
- Can an existing program be improved?
- Is the operation of the program worthwhile?
- Does the program meet the needs of its participants?
- What is the impact of the program?
- Does the program produce any unintended consequences?

Methodological Approaches Used in Program Evaluation

A full spectrum of methodologies is commonly involved in program evaluation, including qualitative, quantitative, and mixed methods. Qualitative methods can be used throughout the evaluation of a program. These techniques are especially crucial in the early stages of a program's development. Interviews, open-ended surveys, and/or focus groups are effective tools in starting the program. These approaches afford the opportunity to fill critical gaps in knowledge or understanding that may exist related to planning or guiding the program during early stages of deployment.

Quantitative approaches are useful to answer how much and to what extent. They can provide powerful and convincing documentation of the problem a program seeks to address. Segmentation of such data can also guide program planning to target specific populations at greater risk. Quantitative analyses are the gold standard when it comes to examining the productivity and success of a program through outcome, summative, or cost–benefit evaluations. Mixed methods research is a common approach in program evaluation. Mixed methods designs leverage the advantages of both quantitative and qualitative approaches to address complex and often evolving research questions the evaluator must answer over the longitudinal cycles encountered in program evaluation. The quantitative and qualitative components of mixed methods designs can be applied either sequentially or concurrently to align with the research questions and how they relate to the developmental stages of the program. Sequential mixed methods designs are employed when data analyzed in one component are used to inform and shape the data collection in the subsequent component. Concurrent mixed methods models collect data at the same time and triangulate the results from one component to expand or better explain the results of the other component.

Evaluation Theories and Philosophies of Practice

Program evaluation practice began and evolved for many years as an application of scientific methods to address day-to-day problems faced by individuals and organizations. Therefore, unlike many paradigms that coalesce around a central unifying theory (Kuhn, 1962), program evaluation evolved as method of practice within different professional fields. In the United States, evaluation models in education and organizational management can be traced to the 19th and early 20th centuries. Major investments in government-sponsored programs brought increased attention to the practice of program evaluation after World War II, and general theoretical orientations of evaluation practice began to emerge toward the end of that century.

Additionally, broader theoretical orientations have evolved within the field of program evaluation. Program evaluators tend to embrace general theoretical orientations in their approaches to evaluation, as outlined by Alkin (2004) and others. These orientations can be understood as higher level theories regarding different purposes and rationales for program evaluation. Examples would include utilization focus, methods, social valuing, accountability and control, and social inquiry. These varied theoretical orientations serve to define specializations of professional practice among professional evaluators. Some evaluators opt to identify with one or more higher level theoretical perspectives, whereas others remain more eclectic in their approach to defining the field and practice. Regardless of the lens used in their general orientation, most evaluators return to focus on the more specific conceptual framework of the program logic model to organize an individual evaluation design.

These distinctions between larger theoretical perspectives on program evaluation often reflect the individual practitioner's orientation on the underlying purpose of program evaluation, or the *why* of evaluation practice. However, how an individual evaluation is conducted links more specifically to the conceptual framework defined by its logic model. Because logic models are also sometimes referred to as the *program theory of change*, those new to the field are sometimes confused when discussing evaluation theory. For practical purposes, the logic model is the specified conceptual model that is critical for defining the purpose of the evaluation, the questions it needs to investigate, and the best methods to employ.

The Role of the Researcher

Evaluators may have varying influence on the organization depending on whether they are internal or external to the program under review. External evaluators are individuals who are independent of the organization that houses or sponsors the program being evaluated. Therefore, external evaluators are viewed as more likely to be objective in their analysis and reporting due to their independence from the program. Major grants often require that a program evaluation be conducted of any funded projects; this requirement usually specifies that an external evaluator should conduct the evaluation. An external evaluator comes in at specific milestones of the project, which may be either time based or event based. For example, many grants specify that an assessment of progress from the external evaluator be included in the annual report submitted to the funding agency, and many require a summative evaluation once all program activities are functioning.

Internal evaluators are members of organizations that house or sponsor programs that are being evaluated. An internal evaluator is familiar with a program, making that position more agile and allowing for immediate course corrections. The role of the internal evaluator is especially important during the launching of larger and more complex programs. An internal evaluator can help in quickly identifying any impact of unintended events or gaps in initial planning that are undermining the program's progress or effectiveness. Early interventions based on this information can help conserve valuable resources and increase the program's likelihood for success.

Both external and internal evaluators need to ensure that they guard against potential bias in the design, data collection, analyses, and reporting of their evaluations. The accuracy and accountability dimensions of the Joint Committee on Standards for Educational Evaluation (2011) *Program Evaluation Standards* speak to specific considerations in avoiding bias, and agencies that sponsor or require evaluations may provide additional guidance.

For example, external evaluators engaged in high-stakes evaluations often are required to file extensive affidavits that document their independence from the project under review. Internal evaluators need to not only attend to the guidance provided within the standards but also question whether their role in the organization creates a conflict of interest, whether their personal investment in the project will contribute to bias, or whether their friendships with those conducting the program will make it difficult to provide objective feedback. Thus, it is best that evaluators, whether external or internal, decline assignments where significant doubts exist regarding their ability to render an accurate, objective, and unbiased appraisal of a program.

All university research studies, including program evaluations, are submitted to an institutional review board (IRB) for a review of ethical standards in research. Program evaluations frequently require that additional permissions be obtained from various stakeholders involved in the evaluation. Additional IRB considerations may pertain to studies conducted in which students are serving as internal evaluators. Consult the IRB applications for your university for more clarification. When using program evaluation, you should always keep ethical considerations in mind.

Identifying and Engaging Stakeholders

The evaluator's initial contact with key stakeholders will usually occur during a collaborative meeting to define the role of the evaluator as well as the nature, scope of effort, and reporting involved in the evaluation. This will be the first opportunity to begin developing a list of key contacts needed throughout the evaluation; therefore, it is important to have sufficient representation of all key stakeholders.

It is also wise to come to this meeting prepared to ask structured questions regarding who to contact and how to begin entry into the groups and organizations within the project's stakeholders. As you identify the processes and outcomes that the evaluation may include, you should simultaneously ask about the key contacts with whom you need to interact. For example, you may learn about a potential source of archival data. You next need to ask the following questions: Who would provide me with that data? Who would give me permission to access the data? Are there any individuals higher in the organization I should be aware of and who would endorse my using the data? Perhaps you find that you need to collect data, which may lead to the following questions: Who oversees the activities where the data collection will take place? Who will actually collect the data? If program employees are tasked with data collection, which supervisors will need to be contacted?

Some evaluators develop their own worksheets that flow much like a standardized interview to help in documenting important contacts including key informants, key collaborators, and gatekeepers. Best practices for engaging stakeholders include the following:

- Strive to be regarded as a resource rather than a judge. Collaborate with stakeholders to define what successful outcomes would look like and present the evaluation as a central means to achieving the objectives those outcomes represent.

- Avoid communicating weaknesses, which tend to imply judgment or static conditions, in favor of focusing on opportunities for improvement that key to greater success.
- Be approachable as an empathetic listener; avoid both antipathy, which hinders engagement and information, and sympathy, which hinders objectivity.
- Consider how to best enter each group. Reflect on who introduces you and how he or she does it and consider how this shapes the initial impressions stakeholders form regarding your purpose and trustworthiness.
- Learn the formal organization structure via organizational charts and strive to be aware of informal social structures and opinion leaders.
- Explore how communication about the program is conducted within and across stakeholder groups.
- Be cautious of oversharing information across stakeholder groups.
- Avoid the perception of favoring individuals or groups within the community of stakeholders.
- Remain objective and consistent.

Data for Program Evaluation Studies

Data for program evaluations may be aligned to resolve a single research question or several different questions. Single research questions can be simple and straightforward (e.g., specific outcome analyses) or complex and multifaceted (e.g., cost–benefit analyses, impact analyses). Alternatively, an internal evaluator is usually engaged with a project over an extended period and will face a series of questions that are aligned to a program's need for information. Thus, evaluation commonly requires collecting different types of data to triangulate in addressing more complex questions or to inform and guide a program over time. Qualitative data will be needed to address the *what* and *how* aspects of questions, whereas quantitative data will best serve the *how much* and *to what extent* aspects.

Data Collection Procedures

Whereas other types of research afford researchers the luxury of extending sampling and data collection until an adequate sample is achieved, data analyses in program evaluation are often used for decision making, planning, and deployment of the program. Because any delays to data collection and analyses can jeopardize the entire program, develop data collection plans, procedures, and permissions well in advance to ensure that the evaluation component is effective and does not place the project at risk. Evaluators need to ensure that they do no harm to the program. This does not mean that an evaluator should not call attention to poorly performing programs; it does mean that poorly planned evaluations can interfere with program deployment or processes and contribute to a poorly performing program.

Each of the data collection procedures in Table 16.1 are commonly employed in program evaluation, and each has benefits that must be weighed against the logistic requirements they present.

TABLE 16.1
Data Collection Procedures

Approach	Advantages	Potential Drawbacks or Challenges
Archival data	Often a source of rich data; already exists and does not require collection; low logistical demand	Gaining access/permission to use; limitations from how collected or coded; may be poor quality or nonrepresentational; may be too old
Surveys	Easy and common method for collecting information from large numbers of people; average participant is familiar with surveys; moderate logistical demand relative to data collected; can be cost efficient	Logistic considerations related to identifying/accessing a sample relative to survey method used (Internet, mail, phone, hand-distributed hardcopy); language barriers; low return rates
Interviews	Allows more in-depth questioning/probing; can verify that participants understand questions; allows interviewer to verify that respondent fits the sampling criteria; potential to gain greater insight into nuances of research question	Logistic considerations related to identifying/accessing a sample relative to survey method used (phone, face to face); potential for interviewer bias; physical presence of interviewer may influence response
Field observations (including participant observation)	Direct observation of activity versus self-report; ability to recognize unanticipated influences stemming from social-ecological context surrounding observed activity; helps in understanding interrelationships between program stakeholders	Logistic considerations related to identifying/accessing appropriate observation site; potential for observer bias in coding data; physical presence of observer may influence behavior of those observed; safety of observer; requires significant time commitment for data collection
Document review	Qualitative alternative to quantitative archival data; good source of retrospective information; already exists and does not require collection; low logistical demand	Gaining access/permission to use; may be biased; may be poor quality or nonrepresentational; may be too old
Focus groups	Participant interaction can generate rich information; potential to gain greater insight into nuances of research question; can help in understanding interrelationships between stakeholders	Logistic considerations related to identifying/recruiting appropriate group; potential for overly dominant group members; biased information; presence of group leader may influence responses of group

Careful planning and execution of appropriate sampling strategies are essential to successful program evaluation. Determining the appropriate sample size is a primary consideration. Because the specific elements of an evaluation design include a wide application of data analysis methods, sampling approaches need to be aligned to each research question and analytic strategy employed within each component of the evaluation design. Some factors to consider would include the size of the target population the program is intended to serve, the scope of the program's application within that population, adequate statistical power to support the quantitative analyses proposed, and a means of ensuring adequate saturation in qualitative components.

Although sample size is always an important concern, sampling strategies and techniques are also critical in acquiring valid and meaningful data. Strategies would include aligning appropriate sampling methods for each component of the evaluation model. For example, a random sampling approach is often desired in controlling for alternative explanations of inferential quantitative analyses; however, purposeful-selection techniques may be better suited to answer particular research questions or in situations where subpopulations would be statistically underpowered. Purposeful selection of a smaller number of participants is common practice in qualitative components of an evaluation. Often, purposeful selection for interviews, narrative methods, or focus groups entails identifying and recruiting key informants who possess valid knowledge, information, and/or expertise required to answer a research question.

Other data collection methods also present unique requirements to consider in devising an effective sampling strategy. For example, sampling protocols for observational data collection can require significant planning and previewing. Even the simple step of defining the sampling window needs to consider how the target behavior varies across time and within social and physical environments to ensure observers are in the right place, at the right time of day, and over an appropriate period to provide a valid accounting of the behavior. Similarly, document analysis will require substantial advanced planning to map a sampling plan regarding how archived documents are located, how they are categorized for relevance to the research question, and how many are selected for coding and analysis.

Always keep in mind the logistics required to support the procedures you select. Ask yourself the following questions: What will it take in resources to conduct this study? What are the optimal collection methods that will allow me to collect the data I need? What is the time frame in which I need to complete data collection? Also, use these questions to determine when to decline a request for a last-minute evaluation.

Data Analysis Techniques

A wide variety of analytic techniques are employed in program evaluation. Combinations of the following techniques would be applied depending on whether a program is applying simple exploratory techniques in early needs assessments and formative evaluations or more sophisticated multivariate analyses to perform summative, impact, or cost–benefit evaluations.

Program evaluation may involve one or more of the following data analysis techniques:

- Systematic coding and analysis of qualitative data for either emergent or a priori themes
- Descriptive statistics, usually segmented by subpopulations or conditions
- A wide variety of inferential quantitative analyses

It is important to point out that data analysis strategies should be considered within the logistic context of the availability of data or the ability to collect new data in a timely manner. An optimal technique looks good on paper; however, the optimal feasible technique will produce a successful evaluation.

For program evaluators who know how to use existing data effectively, archival data—information already stored by organizations—is often a valuable resource. This practice has become an increasingly popular means of supporting information-based decision making through a variety of techniques that are sometimes called data mining. Archival data can provide an efficient means of answering time-sensitive questions quickly, or alternatively, it can provide comprehensive data to provide reliable answers to other questions.

Evaluation planning with stakeholders should also address data collection in advance and tie data collection into the normal flow of routine activities whenever possible. Four simple questions to consider regarding archival data are the following:

1. Do the data exist?
2. Are the data good? The quality of the data is very important. What types of quantitative or qualitative data are available? What is the level of measurement for quantitative data? Are the data of adequate quality (free of obvious bias in the content or significant gaps in data collected)?
3. Are the data relevant? What types of statistical analysis can be used in the data? What types of questions can you ask of these data? Do the existing data fairly and equitably represent all stakeholders' needs?
4. May I use it? Which stakeholders need to grant permission to use the data? Will these stakeholders allow unbiased reporting of results?

How to Ensure Quality

As with other forms of research, program evaluators must provide evidence of the trustworthiness, validity, and/or reliability of their measures and establish that their research designs answer the questions posed by the evaluation, while also attempting to consider and assess alternative explanations of the observer effects. Since 1975, the Joint Committee on Standards for Educational Evaluation has served to provide a coalition of professional associations dedicated to the quality of evaluation practice. Its *Program Evaluation Standards* (Yarbrough et al., 2011) presents detailed definitions that provide excellent references from which to benchmark the quality of the evaluation process and products. Evaluation processes should meet standards for utility, feasibility, propriety, and accountability, whereas the products produced in the evaluation should be held to standards of utility, propriety, accuracy, and accountability. Additional information on these standards can be found on the Joint Committee on Standards for Educational Evaluation website at http://www.jcsee.org/.

Appropriateness of Program Evaluation for Your Research Project

In determining whether a program evaluation is right for your research project, ask yourself these questions:

1. What types of questions in my topic could program evaluation address?
2. Have I carefully reviewed the evaluation standards?
3. What is the feasibility of the planned/desired evaluation? Is a specific evaluation strategy/design premature?
4. Can an appropriate evaluation design be performed within the limits defined by available resources?
5. What are potential utility/risks to the different stakeholder groups?
6. What level of support for the evaluation exists across the various stakeholders?
7. What key gatekeepers will need to support the evaluation to ensure access to information, people, and resources?

Action Plan for Using Program Evaluation

The following information is suggested as a sample action plan to supplement the reflection questions posed earlier when considering program evaluation for your research.

- Discuss program evaluation with your advisor or research committee early in your planning. At a general level, find out if program evaluation studies meet the expectations for the degree granted in your program. Would the specific evaluation plan for your study meet the expectations for the degree granted in your program? For example, research in professional doctorate programs is generally oriented to improving practice within the field; however, research conducted within PhD programs is expected to contribute to scientific knowledge in the discipline. A program evaluation of a specific program is usually conducted to improve practice. This most common application more directly aligns to the research expectations for professional doctorates. Alternatively, an evaluation of a program that is based on a clearly defined logic model/theory of change that underlies the structure of the program activities could be framed as a test of that model/theory. This approach to conducting the evaluation would be designed to generate the more generalizable knowledge about the underlying model/theory expected within PhD programs. In either case, you will need to determine whether the expertise to supervise a program evaluation study exists within your current committee structure.

- Engage methodologists familiar with this type of research practice early in formulating the program evaluation plan. Consider having the program evaluation expert serve as the chair for your committee because it blends content and research expertise. Discuss with a methodologist the scope of the evaluation.
- Identify stakeholders and assess feasibility. The feasibility of an evaluation often depends on access to key information, people, and/or resources. Therefore, access/feasibility should be explored early in the planning process to identify key gatekeepers who control access.
- Search the library for articles and doctoral theses and dissertations using program evaluation using the following key search terms: *program evaluation*, *outcomes assessment/evaluation*, *process evaluation*, *summative evaluation*, *impact analysis*, and *cost-effectiveness evaluation*. Search the literature to get a sense of whether and how other researchers in your field use program evaluation and the justification for using these methods.
- Read books and articles about program evaluation methods to learn about the many and varied methods of data collection and data analysis. (See the Key Sources section at the end of this chapter.) Review and become acquainted with the evaluation standards.

The following organizations may prove useful if you plan on expanding your professional skills to include program evaluation:

- The **Association of Institutional Research** is a longstanding organization that specializes in institutional research specifically as it applies to research and assessment in higher education.
- The **American Educational Research Association** is generally regarded as the premiere organization serving the practice of educational research practice.
- The **W.K. Kellogg Foundation** is a major philanthropic organization that hosts an excellent Knowledge Center that contains useful guides for logic model development and an evaluation handbook that is useful in program development and grant writing.
- The **American Evaluation Association** is the membership organization for professional program evaluators and offers professional development training.

Summary

In this chapter, I presented a brief introduction to the broad spectrum of activities that are associated with the practice of program evaluation, including how fundamental research skills are applied to examine a diverse range of questions that can be best organized using a well-defined program logic model as a conceptual frame of reference. I also introduced a core set of program evaluation designs, explained the terms commonly used by evaluation practitioners, and explained how increasingly sophisticated evaluation designs are aligned to address a progressive sequence of research questions posed across stages in the logic model sequence. I then illustrated how research skills related to sampling and analyses are applied by commonly employed evaluation designs.

I explored the logic model and how it illustrates that program evaluation applications can span the full continuum of research inquiry. Beginning with initial needs assessments, where little is known, and leading through more complex summative analyses that seek to validate the causal impact of a program, the framework of the logic model situates specific research elements on the continuum of research inquiry. I also focused on how logic models provide the most relevant conceptual framework to guide specific evaluations. This focus is used to provide the most directly applicable information to guide program development and deployment. Broader theoretical orientations within the field of program evaluation, such as utility, methods, valuing, accountability and control, and social inquiry, can be explored further through the resources provided in this section.

I examined whether program evaluation is an appropriate approach to research questions. Program evaluation is primarily an applied research paradigm that is amenable to a variety of applications that center on needs for programming; program planning, development, and employment; and program effectiveness and efficiency. Program evaluation usually deals with addressing gaps in practice; however, it also can be used to address gaps in research literature when engaging in relevant research questions and applying rigorous research design.

The fundamental knowledge, skills, and abilities you acquire related to data collection and analysis can be put to direct and practical application in program evaluation. Even though some evaluators specialize in either quantitative or qualitative approaches, the diversity of research questions encountered in program evaluation underscores the advantage of having adequate skills in both approaches to employ coordinated mixed methods designs.

Program evaluation provides timely, accurate, and focused information to effect social change through effective and efficient programs, practices, policies, and interventions. It can apply to an array of research designs and methods; it is highly adaptable to a broad range of applied research questions, where its impact tends to be clear and immediate. Early involvement of the evaluator can increase the likelihood of success and the efficient use of program resources.

Because program evaluation is usually conducted within ongoing processes in an organization, specific evaluations are frequently challenged by constraints common to field research in general and more specific constraints within the organization. Field research, like program evaluation, often engages with operations and processes that are already under way and evolving. Therefore, key opportunities for data collection, analysis, and input may be lost if the evaluator is engaged too late in development. Field studies often present challenges related to access to people, places, and information, as well as recruiting and coordinating engagement and support across multiple individuals and groups. Finally, program evaluation requires that adequate resources are available to ensure a successful evaluation project.

KEY SOURCES

Patton, M. Q. (2008). *Utilization focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.

Soriano, F. I. (2013). *Conducting needs assessment: A multidisciplinary approach* (2nd ed.). Thousand Oaks, CA: Sage.

Steckler, A. B., & Linnan, L. (Eds.). (2002). *Process evaluation for public health interventions and research*. San Francisco, CA: Jossey-Bass.

Wholey, J. S., Hatry, H. P., & Newcomer, K. E. (2004). *Handbook of practical program evaluation* (2nd ed.). San Francisco, CA: Jossey-Bass.

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Sage.

REFERENCES

Alkin, M. C. (2004). *Evaluation roots*. Thousand Oaks, CA: Sage.

Altschuld, J., & Kumar, D. (2009). *Needs assessment: An overview*. Thousand Oaks, CA: Sage. Retrieved from http://www.sage-ereference.com/view/needs-assessment-an-overview/SAGE.xml?rskey=5Br3kd&row=10

Barnett, K. (2012). *Best practices for community health needs assessment and implementation strategy development*: *A review of scientific methods, current practices, and future potential report*. Report of Proceedings from a Public Forum and Interviews of Experts, Atlanta, GA. Retrieved from http://www.phi.org/uploads/application/files/dz9vh55o3bb2x56lcrzyel83fwfu3mvu24oqqvn5z6qaeiw2u4.pdf

Coulon, S. M., Wilson, D. K., Griffin, S., St. George, S. M., Alia, K. A., Trumpeter, N. N., . . . Gadson, B. (2012). Formative process evaluation for implementing a social marketing intervention to increase walking among African Americans in the Positive Action for Today's Health trial. *American Journal of Public Health*, *102*(12), 2315–2321. doi:10.2105/AJPH.2012.300758

Davis, M., Orpinas, P., & Horne, A. M. (2012). Design and evaluation of prevention programs. In E. Vera (Ed.), *The Oxford handbook of prevention in counseling psychology* (pp. 91–108). Oxford, United Kingdom: Oxford University Press. doi:10.1093/oxfordhb/9780195396423.013.0007

Dehar, M. A., Casswell, S., & Duignan, P. (1993). Formative and process evaluation of health promotion and disease prevention programs. *Evaluation Review*, *17*(2), 204–220.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405–432. doi:10.1111/j.1467-8624.2010.01564.x

Jimison, H., Adler, L., Coye, M., Mulley, A., Jr., & Eng, T. R. (1999). Health care providers and purchasers and evaluation of interactive health communication applications. *American Journal of Preventive Medicine*, *16*(1), 16–22.

Joint Committee on Standards for Educational Evaluation. (2015). *Program evaluation standards*. Retrieved from http://www.jcsee.org/program-evaluation-standards

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

McKillip, J. (1998). Needs analysis: Process and techniques. In L. Bickman & D. J. Rog (Eds.), *Handbook of applied social research methods* (pp. 261–284). Thousand Oaks, CA: Sage.

Mohr, L. B. (1995). *Impact analysis for program evaluation*. Thousand Oaks, CA: Sage.

Oakley, A., Strange, V., Bonell, C., Allen, E., & Stephenson, J. (2006). Health services research: Process evaluation in randomised controlled trials of complex interventions. *British Medical Journal*, *332*(7538), 413. doi:10.1136/bmj.332.7538.413

Paul, C., Yeats, J., Clarke, C. P., Matthews, M., & Skrabala, L. (2015). *Assessing and evaluating Department of Defense efforts to inform, influence, and persuade*. Santa Monica, CA: RAND.

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2003). *Evaluation: A systematic approach*. Thousand Oaks, CA: Sage.

Saunders, R. P., Evans, M. H., & Joshi, P. (2005). Developing a process-evaluation plan for assessing health promotion program implementation: A how-to guide. *Health Promotion Practice*, *6*(2), 134–147. doi:10.1177/1524839904273387

Schalock, R. L. (2001). *Outcome-based evaluation* (2nd ed.). New York, NY: Kluwer Academic/Plenum.

Soriano, F. I. (2013). *Conducting needs assessment: A multidisciplinary approach* (2nd ed.). Thousand Oaks, CA: Sage.

Steckler, A. B., & Linnan, L. (Eds.). (2002). *Process evaluation for public health interventions and research*. San Francisco, CA: Jossey-Bass.

Trochim, W. M. (2006). *The research methods knowledge base* (2nd ed.). Cincinnati, OH: Atomic Dog. Retrieved from http://www.socialresearchmethods.net/kb/

Volkwein, J. F. (Ed.). (2011). *Assessing student outcomes-why, who, what, how? New directions for institutional research, assessment supplement 2009*. Chichester, United Kingdom: Wiley.

Wholey, J. S., Hatry, H. P., & Newcomer, K. E. (2004). *Handbook of practical program evaluation* (2nd ed.). San Francisco, CA: Jossey-Bass.

Williams-Reade, J., Gordon, B. A., & Wray, W. (2014). A primer in program evaluation for MedFTs. In J. Hodgson, A. Lamson, T. Mendenhall, & D. R. Crane (Eds.), *Medical family therapy* (pp. 283–299). New York, NY: Springer Publishing. doi:10.1007/978-3-319-03482-9

W.K. Kellogg Foundation. (2004). *Using logic models to bring together planning, evaluation, and action: Logic model development guide*. Retrieved from http://www.smartgivers.org/uploads/logicmodelguidepdf.pdf

World Health Organization. (2000). *Workbook 3: Needs assessment*. Retrieved from http://www.unodc.org/docs/treatment//needs_assessment.pdf

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Sage.

17

Interviewing Essentials for New Researchers

Linda M. Crawford, PhD, and Laura Knight Lynn, PhD

Introduction

It is likely that you have turned to this chapter because you have decided that interviewing will be a best source for collecting your data. Interviewing can provide rich information from study participants, including the text of responses, vocal tone and inflection, and body language. However, along with these benefits, there are also risks in interviewing of which new researchers are typically not aware. Many researchers assume you can just get up and collect the interview data, but, if they are not careful, much can happen in the data collection, analysis, and interpretation processes that can impair the integrity of the data. In this chapter, we define various types of interviews, provide information regarding choice of either individual or focus group interviews, describe the interview process, examine some interviewing pitfalls and solutions, and describe the important role of culture and power in interviewing.

Types of Interviews

Interviews fall into two classifications. One classification relates to the *form* of the interview as structured, semistructured, or unstructured. A second classification addresses the *members* of the interview, including individual or group interviews.

Interview Form

There are three basic forms for interviews: structured, semistructured, and unstructured. Each form is progressively less rigid in the questioning format and process. For a structured interview, the researcher writes out a script of questions to be posed and asks the questions of each participant exactly as stated and in the exact sequence. The researcher records the responses but does not probe for clarification or additional information. Structured interviews are often conducted as quantitative, Likert-type scale surveys. One example is the researcher who stops participants in a mall and asks for answers to a few questions. Another example is the structured phone interview. In qualitative research, structured interviews are used when the intent is to precisely compare participant responses. For example, for a study of the problem of low minority enrollment in talented and gifted educational programs, the researcher might ask each participant about satisfaction with the program admittance procedures and with the program itself without probing for details.

For semistructured interviews, the researcher constructs interview questions related to the research question(s) and also anticipates probes that might be used to explore participant responses (Rubin & Rubin, 2012). In a semistructured interview, the interview questions are posed to each participant, and probing questions are used to follow-up and gather deeper information from the participant. The research plan suggests an order to the interview questions, but probing questions might alter that order. For example, for a study of the problem of low minority study enrollment in talented and gifted educational programs, the following might hold:

- Research question: How do minority parents perceive the relevance of talented and gifted programs to their children?
- One interview question: From what you know of the talented and gifted programming that has been provided for your child, how well do you think the program meets your child's needs?
- Probing question: Tell me more about how your child has talked about his or her experience in the talented and gifted program.

Probing questions would be used in relation to the initial response from the participant and may not be used with all participants. Although probing questions can be anticipated, they also might be developed during the course of an interview, in which case the researcher must record the newly created probes. Semistructured interviews are a usual form for qualitative studies, particularly for novice researchers. Unstructured interviews are more open than semistructured interviews.

In unstructured, or open, interviews, the interviewer constructs questions "on the fly" in response to participant contributions that are related to the research topic. For example, in relation to the previous example of minority participation in talented and gifted program, an open interview might begin with a very broad question to a parent, such as "How do you feel the talented and gifted program fits your child's needs?" The next interview question would be constructed in relation to what the participant said. With unstructured interviews, the line of questioning may vary among participants depending on their responses.

Semistructured interviews are typical in qualitative research and are recommended for novice researchers. Structured interviews may not offer the researcher sufficient opportunity to probe participant responses. Without extensive experience as a qualitative researcher, unstructured interviews may not offer sufficient focus to answer the research question.

Interview Members

There are two membership groups for interviews: individual interviews and group interviews, known as focus groups. Individual interviews are conducted in a one-on-one situation with the researcher, or other interviewer who asks questions, and the interviewee who responds to them. For a focus group interview, the researcher assembles about 6 to 10 persons relevant to the study topic and moderates a discussion among them that is guided by interview questions (Brinkman & Kvale, 2015).

Selecting Between Individual or Focus Group Interviews

The choice between conducting individual or focus group interviews requires you to consider the following:

1. Study design (whether the design itself requires individual or group interviews)
2. Availability and time constraints (whether it is easier to schedule individuals or groups)
3. Location (whether there is a place to conduct individual interviews or focus groups in which audio and/or video recording is feasible)
4. Influence of group talk over individual talk in relation to the research questions (whether the focus group discussion enhances or unduly influences individual responses)

Study Design

Phenomenological studies require individual interviews because they intend to derive composite themes from individual experiences. Because the unit of analysis for narrative studies is at the individual level, narrative studies also require individual rather than focus group interviews. Individual interviews are usually used in grounded theory studies, but focus group interviews could also be used, and both individual and focus group interviews are commonly used for case and ethnographic studies. In the context of grounded theory, case study, and ethnographic research, focus groups may draw out a variety of viewpoints stimulated by the interaction among group participants (Brinkman & Kvale, 2015).

Availability and Time Constraints

Study design is the major consideration in deciding between individual and group interviews but the practical considerations of availability and time constraints are important as well. If participant availability is limited, and if participant availability occurs at a singular time, then a group interview might be preferred, particularly if the researcher also has time constraints in accessing participants. A third consideration in selecting between individual and focus group interviews is access to a suitable location.

Location

Individual and focus group interviews need to be conducted in a location that provides accessibility, visual privacy, freedom from distractions, and video/audio recording capability. The location of the interview needs to be convenient to and permissible for the participants, both in time and location. For example, if the interview is conducted on the job site and during work time, the researcher needs to obtain permission to use the job site and time for the study. In addition, the space must be large enough to seat all participants comfortably. The interview location should afford the participant(s) privacy. In other words, the interview should not be conducted in an open area, such as a cafeteria, but rather in a closed door room. Using a closed door room for interviews also minimizes distractions. The location of the interview needs to allow for clear audio and video (if required) recording. Accurate recording of the interview is essential for data analysis, and the researcher should test all recording equipment prior to data collection and also have backup equipment available. The space in which the interview is conducted should allow for audio recording without feedback or distortions and for video recording of the entire group. Video recording is especially important for focus group interviews in order to identify who is talking and to whom.

Group Talk Influence

When determining if individual interviews or focus group interviews are better for a study, you must consider the influence of group talk over individual responses. In a focus group, the responses of one person might influence the responses of others. This influence can be beneficial in that it can stimulate ideas in individuals, but it can also be detrimental in that it can bring the conversation in directions not associated with the research questions, prevent individuals from speaking in opposition to others, or result in groupthink (Brinkman & Kvale, 2015).

The Interview Process

Interviewing is not a simple a task and requires attention to detail. There are several stages that must be considered and described in the research proposal in support of effective interviewing, including preparing for the interview, conducting the interview, and transcribing interviews. Having a clear set of interview questions is essential. This section will focus on logistics of the interview process in terms of preparing for, conducting, and transcribing interviews.

Constructing an Interview Protocol

An interview protocol is a tool used to ensure consistency in the interview process. It is a form that guides the interviewer, including a location to record name of interviewee, date and time of the interview, location of the interview, name of the interviewer, and any other relevant details. It includes introductory and closing remarks as well as the questions to be asked. In summary, an interview protocol should include the following:

1. A place to record details of the interviewee, interviewer, time, place, and recording mechanism
2. The introductory explanation of the interview and its purpose and how it will be conducted (tape recording, detailed notes, etc.)
3. A list of questions to be asked, with possible probes used to obtain more information. Probes would include statements such as “Please explain,” “Can you say more about that?” and “Please describe what kind.”
4. A closing statement that includes detailing what will happen next with the data, thanking the participant for being involved in the interview, and asking permission for further contact to gain clarification as needed

The layout of the interview protocol may include a section for the interviewer to record subjective observations of the participant or reactions to the participant's responses. As advised by Josselson (2013), the purpose of the interview protocol is to ensure consistency across interviews; it is not intended for verbatim notation of the participant's responses, although the interviewer may make some brief notations that may be reviewed during analysis. Figure 17.1 shows a sample interview protocol.

Clear protocols are important in all cases but especially when there is more than one interviewer. In that case, the additional interviewer should use the same protocol and be trained in interview implementation for consistency.

Preparing for Interviews

When preparing for interviews, you must consider the site, the participants, and yourself. In relation to the site, be sure that all relevant parties at the site are informed of the time, dates, and locations of the interviews. For example, if the interviews are to take place in a conference room during the work day, be sure to schedule the conference room for the required times. Think about even small details, such as putting a trashcan outside the door so that a custodian does not interrupt the interview in order to collect it. The site also includes any recording equipment. Test the equipment to ensure it records voices clearly, and if using video, ensure that the video frame includes all participants. Also, have a backup system for both audio and video recording in the event of an equipment failure.

Confirm scheduled dates and times with participants. Ensure that participants have received informed consent materials and that you have previously obtained any required signatures or have a clear plan to obtain any required signatures at the time of the interview. If, during the interview, it seems the session might go over the set time, ask for permission to continue the interview but also have a plan for adjusting the interview to fit into the time frame without going over the schedule.

In terms of preparing yourself for interviews, practice interviewing with a friend or colleague, and if possible, record the practice sessions on videotape. Obtain feedback from your partner in regard to your posture, your manner of asking questions, and any subtle facial expressions or voice tone that might influence participant responses. The goal here is to learn if, for example, you are asking questions too quickly or using body language to communicate a judgment. The goal can also be to learn if you are missing opportunities to probe more deeply into responses. Videotaping practice sessions can give you the opportunity to self-critique beyond feedback from your partner.

Conducting Interviews

You must decide who is going to conduct the interview. We consider it preferable for the researcher to conduct the interview in order to bring the interviewer closer to the participants and the data. However, if you have a close personal or professional relationship with the interviewee, then it may be advisable to have someone else conduct the interview. Certainly, if you have any supervisory or other power over the interviewee, you should seek out someone else to conduct the interview. If someone other than yourself conducts the interview, then you must determine the qualifications of the

Interview Protocol

Date of interview:

Location of interview:

Start time: End time:

Name of interviewee:

Name of interviewer:

Recording mechanism:

Introduction:
Thank you for taking your time to meet with me today. As you know, this interview will contribute information for a research study intended to [state purpose of the study]. You have signed an informed consent form, but, as a reminder, you may decline to answer any question you do not wish to answer or withdraw from the interview at any time. This interview will take approximately [X] minutes. Do you have any questions before we begin?

Interview Questions	Interviewee Responses	Interviewer Observations/Reactions
1. List interview questions, one per row.	Do not notate verbatim response but may notate some thoughts that jump out at you.	Notate observations, such as *interviewee has drawn back in her chair and crossed her arms,* and personal reactions, such as *I'm wondering if my question has offended or challenged her.*
2.		
3.		
Etc.		

Potential probes:

- Please tell me more about . . .
- How did you know . . .
- What kind . . .
- What was the best approach among those you named?
- What would you have liked done differently?

Conclusion:
Thank you for your time today. I very much appreciate your contributing to this study. [State any next steps, such as member checking.]

FIGURE 17.1 **Sample interview protocol.**

interviewers and present them in the study document. For example, you might observe interviewers conducting pilot interviews to confirm that they conduct the interview as intended by the study design and methods.

A question sometimes arises as to whether interviews must be conducted face-to-face or whether they can be conducted using audio/video conferencing tools, such as Skype, or by telephone. We consider it preferable to conduct interviews face-to-face because that personal presence enhances the establishment of rapport, allows the researcher greater access to subtle body language cues from the interviewee, and permits the researcher greater control over the interview environment. However, technological tools expand the geographic area from which participants may be drawn and thus can be useful for studies for which the researcher cannot physically meet with the participants. Given a choice of technological tools, we advise using a tool with a video element. With such conferencing tools come problems with stability of the tool, Internet bandwidth to accommodate the interaction, and avoidance of "breaking up" when interacting and recording the interview. Telephone interviews can be more stable but, of course, lack the video capability.

When conducting the interview, the researcher must avoid body posture, body language, voice tone, and linguistic constructions that communicate judgment, either positive or negative, or lead the participant. For example, a physical response that communicates "I like what you said" would inject researcher bias into the interview. Body posture and language must remain neutral. Likewise, asking a follow-up probe such as "Don't you think people would be better off if . . . " communicates an opinion on the part of the researcher, whereas a more open question such as "What do you think would be better for . . . " allows the respondent to more freely express opinion.

You may find yourself overwhelmed with input when conducting an interview. You are hearing words from the participant(s), seeing body language, and reacting yourself, both emotionally and cognitively. Josselson (2013) suggests not taking notes during interviews. The interviews are recorded so you have access to them, and taking notes can distract both you and the interviewees.

Transcribing Interviews

Transcription is the process of transferring audio-recorded statements from the recording to a printed text. This is a time-intensive process that involves concentration and attention to detail in order to accurately represent the participants' words and meaning. There are three options for transcribing interviews: (a) the researcher can transcribe them, (b) the researcher can hire someone to transcribe them, or (c) the researcher can use software to transcribe the interviews. There are pros and cons to each choice.

The benefit of you as the researcher transcribing the interviews is that you hear the participants' voices and become more intimately acquainted with their utterances. However, accurate transcription is quite time consuming and requires listening and relistening to recordings in order to capture the utterances accurately, including filler words, such as *um*, as well as a good-quality recording from which to transcribe; a tool that allows you to start, stop, and repeat sections; and a high level of typing skill (Brinkman &

Kvale, 2015). A 30-minute interview can take approximately 180 minutes to transcribe.

There is the choice of using professional transcribers. You as the researcher can save time by hiring a transcriptionist, but hiring professionals is expensive and takes you a step away from the data. Even if you hire a transcriber, you still should listen to the interviews while reading the transcripts in order to verify transcript accuracy. Furthermore, you should give the transcriber explicit instructions, such as how to indicate pauses in responses, how to indicate inflections, and how to punctuate to indicate meaning. If more than one transcriber is employed, issues of consistency among transcriptions can arise because human beings hear and record audio input with some variation. Using an experienced transcriber, though, can be a benefit in that the transcriber may be versed in conventions that communicate the meaning of utterances. You, then, must become versed in how to read those conventions.

Software, such as Dragon Naturally Speaking, as well as other programs, can be used both to record and to transcribe interviews. Using software for transcription can be efficient, but the software usually has to be trained to understand specific voices. Training is typically needed for each interviewee, which can result in inaccurate transcription. Using such software can also be cumbersome when interviewing multiple people because the program may be unable to distinguish voices in a group interview. Given those issues, we do not recommend using voice recognition/transcription software. We recommend audio recording of interviews and manual transcription. Manual interview transcription is a time-intensive process for individual interviews and even more so for group interviews. It is imperative to have start, stop, and repeat capability when listening to and transcribing interviews. NVivo software provides that capability, as do other software packages and transcription tools.

Group interviews present a unique transcription challenge in that people sometimes talk over each other, and it may be difficult to discern who is talking from just an audiotape. A video recording of group interviews will assist in overcoming these challenges; however, use of video recording presents a challenge in and of itself in that the transcriber must attend to both audio and video input. Transcription conventions, such as using a bracket to indicate overlapped utterances, can clarify the flow of conversation within a group interview (Brinkman & Kvale, 2015).

Interviewing Pitfalls and Solutions

The cause of risks with interviews as the data collection approach is simple:

1. The primary interview tool is YOU, and
2. You are HUMAN.

What does this mean? It means you are subject to the influence of your collection of human experiences, your memories, your impressions, and your feelings. These serve you well in daily life, but during interviews, our recollections, memories, and impressions can “muddy the waters.” The aspects of yourself that help you navigate from point A to point B when moving through a crowded conference or shopping mall are the same aspects that can hinder you as a qualitative researcher and interviewer. How so? Your experiences, feelings, and filters color

- What you see in interviews
- What you hear in interviews
- What you find important in interview utterances

It is important to understand that what is not part of your mental and emotional background is often filtered out. Also, you are primed to your hunches and notions, which facilitate attaching your biases to the data.

With this subjectivity in mind, Table 17.1 describes some common pitfalls in interviewing, why the pitfall is problematic, and some ways to avoid the pitfalls.

In summary:

1. Be sure, be calm, and go slow.
2. Have a clear protocol with a list of possible probes.
3. Have a thorough plan for recording and documenting the interview.
4. Reflect on your own worldview and assumptions.

The fourth point, reflection on personal worldview and biases, is important enough to discuss a bit further. You bring to the interview the person you are, with all your past experiences and assumptions. Although you try to create an open mind within yourself in order to receive the participant's experiences without bias, your own worldview and experiences can color how you hear the words of the interviewees. That is reality (Josselson, 2013). The charge to you as a researcher is to recognize those personal influences and to *bracket* them. *Bracketing* is "a method used in qualitative research to mitigate the potentially deleterious effects of preconceptions that may taint the research process" (Tufford & Newman, 2010, p. 80). There are several ways to bracket, including memoing, reflexive journaling, and external interviews (Tufford & Newman, 2010). *Memos* are notes written primarily to the self during the course of data collection and analysis that record descriptions of some aspect of the study environment or phenomenon as the researcher is seeing it, meanings the researcher attributes to analytical categories, and/or explanations of patterns emerging among categories (Schwandt, 2015).

Reflexive journaling is an ongoing narrative of the researcher's rationale for the study, assumptions, values, and relationships to the participants in terms of culture and power. Reflexive journaling begins before even writing the research questions. Whereas memoing is the researcher's understanding of the data, reflexive journaling reveals the researcher's presuppositions in relation to the study and his or her relationships with the participants (Tufford & Newman, 2010).

A third way to bracket one's researcher bias is to engage in *external interviews*, that is, interviews with those not participating in the study but who may allow the researcher to discern biases in questions, probes, or judgmental reactions to participant responses. Such interviews allow the researcher to recognize, and correct, unintentional physical or verbal responses to participants that communicate a judgment, closing off the flow of information (Tufford & Newman, 2010).

TABLE 17.1
Interviewing Pitfalls and Potential Solutions

Interviewing Pitfalls	Why Problematic	Ways to Avoid These Pitfalls
Rushing the informed consent/ explanation of steps process	Interviewee feels confused and pressured to complete interview.	Carefully plan the informed consent process. Be sure to present information in a calm and relaxed way. Be sure it is clear to interviewees that it is their choice to participate and they can withdraw at any time.
Rushing the interview	Limited data lacking rich information to answer research questions.	Continue the calm, relaxed, and pleasant approach. Because many people tend to talk quickly, overtly trying to slow down a bit can be helpful.
Too much facial expression/ reactions	May close responses or influence nature of responses.	Find a comfortable, neutral facial expression and maintain it. If you tend to be naturally facially expressive, practice managing that in practice interviews. You want to be very careful not to express surprise, agreement, pleasure, or offense in reaction to the interviewee.
Frequent nodding	May influence nature of responses as it indicates agreement.	Limit nodding.
Lack of probes and follow-up questions	Limits richness of data.	Plan possible probes and use them to urge a participant to expand or more directly answer the question.
Only hearing and documenting part of responses	Impacts integrity and validity of data.	Audiotape or document verbatim. Listen to transcripts following interview and/or review notes.
Unprepared for emotional responses	Unethical to have participant feeling vulnerable without follow-up or support.	Have references available for counseling for emotionally laden interviews.
Too connected to interviewee	Interviewee could be eager to please, and this can impact the integrity and truth in data.	It is best not to interview people you are connected to in some way.
No rapport with interviewee; distant or robotic	Can limit responses.	Make sure you are natural with normal eye contact and responses so interviewee is comfortable.
Ordering of questions not well planned	If more intrusive questions are first, can limit responses.	Start with more innocuous questions such as demographic, factual type questions and build to the deeper questions.

continued

TABLE 17.1 *(continued)*
Interviewing Pitfalls and Potential Solutions

Interviewing Pitfalls	Why Problematic	Ways to Avoid These Pitfalls
Questions include presuppositions	Can be leading and influence the integrity of the data.	Use a direct preface. For example, instead of "How has this changed your life?", first ask, "Has this had an impact on your life?" Then ask, "How?"
Inappropriate use of "why" questions	Can be leading because there may not be a known "why"; "why" questions can create defensiveness.	Make sure that the interviewee is likely to have an explanation of why. For example, with children, why questions are often confusing. Reword the question to avoid the use of "why."
Incompletely transcribing interviews	Makes data invalid and biased.	Use software to slow down recording and completely transcribe. Consider hiring a transcriptionist.
Not audio recording interviews or documenting verbatim	If verbatim information is not available, can cause data to be limited and invalid.	Audio record interviews and be prepared to document verbatim on site if interviewee refuses to be audiotaped.
Issues of power not considered	If interviewee perceives the interviewer as a person of "power," can influence or limit what is reported.	When meeting with the site, be sure to describe your role and background. Inquire as to whether the site contact foresees any issues. This information helps in protocol preparation and planning the data collection process.
Issues of culture not considered	Every environment has a culture, and specific people have cultures. If this is not considered, researcher can intrude or offend unintentionally and limit the nature of data that is reported.	Interviewer needs to gather information from the site on cultural considerations for both the environment and the people interviewed. Special consideration for ethnic norms should be considered (e.g., eye contact, appropriateness of shaking hands).

In addition to the pitfalls described earlier, issues of culture and power are often overlooked in interviewing. Because interviewing involves a one-on-one (or one-to-group) relationship, these issues can impact data collection.

Culture and Power in the Interview Process

Culture is everywhere. A business team has a culture, a university department has a culture, a K-12 school has a culture, and a hospital floor has a culture. Before entering the environment in which you will conduct interviews, you should find a way to understand its culture. Talk to your site contact; this contact is an insider who can provide valuable insight into particular nuances of the culture. He or she can give you tips on limiting intrusiveness by your presence and useful approaches for style in engaging in the interview. Many important questions can be answered, such as the following: What should you be sensitive to? Are employees time pressured? Are they fearful of exposure?

There is also the ethnic/religious culture element. Even if you are collecting data in the United States, some of your interviewees could be uncomfortable being interviewed by a woman, engaging in eye contact, or shaking hands. Some questions that seem okay for American culture could be offensive to those from a different culture. Try to get an understanding of your site population and the possibility of diverse national origins. That way, you can prepare well in advance of the interview. Be careful to understand but not stereotype. When unsure, simply ask the interviewee for permission or about his or her comfort level.

In addition to culture, power or perceived power can make a difference in data collection. A participant can perceive you to be a person of power without your knowing it. Do you hold an advanced position in your field? Do you have a college degree? Do you have a master's degree? Are you a doctoral student? Are you known for something in your field? Some participants may look up to you for these reasons and want to please you. You should consider your setting, make sure you encourage candid responses, and explain the purpose of your research and the usefulness of honest, candid data. In some cases, perceived power differentials cannot be overcome. You need to be perceptive of clues that an interviewee is modifying responses to please you. For example, notice if the interviewee looks to you for affirmation of responses or expresses undue admiration of your position. In such situations, you might want to have a person other than yourself conduct the interview, or it might be that the participant's data become marginal to the study.

Summary

Interviewing can be a dynamic way to collect data and can yield rich, in-depth qualitative data. What we have tried to include here are some reminders and tools for you to consider before you make your way out into the world as a qualitative interviewer. Becoming a qualitative interviewer will require practice and also studying some of the key sources offered below. Remember that just as you would perfect a survey or select the best quantitative assessment for your construct, when collecting your interview data, you are the research tool that needs to be perfected and sharpened. You need to make sure you are prepared and practiced in order to limit ways you could unintentionally invalidate your data. Also, novice researchers often make the mistake of thinking that doing interviews is easy. It is not. Anyone conducting interview research should gain some training in interviewing skills, preferably in a class on interviewing. At the very least, read one of the many resources available on interviewing skills, and practice before conducting your first interview. Colleagues will be happy to provide feedback on the data you receive as well as your style and presence during the interview.

KEY SOURCES

Brinkman, S., & Kvale, S. (2015). *Interviews: Learning the craft of qualitative research interviewing* (3rd ed.). Thousand Oaks, CA: Sage.

Gubrium, J. F., Holstein, J. A., Marvasti, A. B., & McKinney, K. D. (2012). *The SAGE handbook of interview research: The complexity of the craft* (2nd ed.). Thousand Oaks, CA: Sage.

Josselson, R. (2013). *Interviewing for qualitative inquiry: A relational approach*. New York, NY: Guilford Press.

Rubin, H. J., & Rubin, I. S. (2012). *Qualitative interviewing: The art of hearing data*. Thousand Oaks, CA: Sage.

Seidman, I. (2013). *Interviewing as qualitative research: A guide for researchers in education and the social sciences* (4th ed.). New York, NY: Teachers College Press.

REFERENCES

Brinkman, S., & Kvale, S. (2015). *Interviews: Learning the craft of qualitative research interviewing* (3rd ed.). Thousand Oaks, CA: Sage.

Josselson, R. (2013). *Interviewing for qualitative inquiry: A relational approach*. New York, NY: Guilford.

Rubin, H. J., & Rubin, I. S. (2012). *Qualitative interviewing: The art of hearing data*. Thousand Oaks, CA: Sage.

Schwandt, T. A. (2015). *The SAGE dictionary of qualitative inquiry* (4th ed.). Thousand Oaks, CA: Sage.

Tufford, L., & Newman, P. (2010). Bracketing in qualitative research. *Qualitative Social Work*, *11*(1), 80–96. doi:10.1177/1473325010368316

About the Editors

Gary J. Burkholder, PhD, is Vice President of Research and Accountability at Laureate Education, Inc., as well as Senior Research Scholar and a faculty member in the psychology and public health programs at Walden University. He previously served as Senior Research Associate and Statistician at the Institute for Community Research in Hartford, Connecticut, and Director of Survey Center and Data Management at the Institute for Community Health Promotion at Brown University. Dr. Burkholder has served as a mentor for nearly 100 doctoral and master's thesis students, and he has been the recipient of the Walden University Bernard L. Turner Award for excellence in mentoring doctoral dissertations. He publishes and presents in the areas of research methods, social psychology, and public health.

Kimberley A. Cox, PhD, holds a degree in social psychology. For nearly a decade, she has served as a faculty member in the School of Psychology at Walden University, where she teaches graduate courses and mentors dissertation students. She was recently recognized with the Walden University Bernard L. Turner Award for excellence in her role as a dissertation mentor. Dr. Cox's research interests focus on the application of social psychological theory and methods to social issues. She also has interest in online teaching and learning, and she is a past recipient of Walden University's Faculty Research Initiative Grant and Presidential Research Fellowship in Distance Learning.

Linda M. Crawford, PhD, is Director of Quality Assurance at Laureate Education, Inc., and a faculty member at both Walden University and University of Phoenix. She has held both teaching and administrative positions in the fields of K–12, undergraduate, and graduate education. She has twice received the Walden University Bernard L. Turner Award for excellence in mentoring doctoral dissertations, and she has also been the recipient of the Walden College of Education's Extraordinary Faculty Award. Involved in online education for nearly two decades, Dr. Crawford has developed and taught courses for multiple universities in the areas of research, measurement and evaluation, curriculum theory and design, instructional practice, and educational law. Her research centers on mentoring graduate students and building a sense of community among students in the online environment.

Index

D

E

F

G

H

I

J

Q

R

T

U

V

W

Y

Z