EDUCATIONAL RESEARCH

THIRTEENTH EDITION

Competencies for Analysis and Applications

Geoffrey E. Mills | Adam W. Jordan

Educational Research

Competencies for Analysis and Applications

Thirteenth Edition

Geoffrey E. Mills
Southern Oregon University

Adam W. Jordan
College of Charleston

Content Management: Rebecca Fox-Gieg
Content Production: Janelle Rogers
Product Management: Drew Bennett
Product Marketing: Krista Clark
Rights and Permissions: Jenell Forschler

Cover Image by LWA/Dann Tardif/DigitalVision/Getty images

Library of Congress Cataloging-in-Publication Data
Names: Mills, Geoffrey E., author. | Gay, L. R., author.
Title: Educational research: competencies for analysis and applications/Geoffrey E. Mills, Southern Oregon University, Adam W. Jordan, College of Charleston.
Description: Thirteenth Edition. | Boston: Pearson, [2022] | Twelfth edition: 2019. | Includes bibliographical references and index.
Identifiers: LCCN 2022008576 | ISBN 9780137535101 (Paperback)
Subjects: LCSH: Education—Research. | Education—Research—Methodology.
Classification: LCC LB1028 .M495 2022 | DDC 370.72—dc23/eng/20220419 LC record available at https://lccn.loc.gov/2022008576

ScoutAutomatedPrintCode

ISBN 10: 0-13-753510-4
ISBN 13: 978-0-13-753510-1

Pearson's Commitment to Diversity, Equity, and Inclusion

Pearson is dedicated to creating bias-free content that reflects the diversity, depth, and breadth of all learners' lived experiences.

We embrace the many dimensions of diversity, including but not limited to race, ethnicity, gender, sex, sexual orientation, socioeconomic status, ability, age, and religious or political beliefs.

Education is a powerful force for equity and change in our world. It has the potential to deliver opportunities that improve lives and enable economic mobility. As we work with authors to create content for every product and service, we acknowledge our responsibility to demonstrate inclusivity and incorporate diverse scholarship so that everyone can achieve their potential through learning. As the world's leading learning company, we have a duty to help drive change and live up to our purpose to help more people create a better life for themselves and to create a better world.

Our ambition is to purposefully contribute to a world where:

- Everyone has an equitable and lifelong opportunity to succeed through learning.
- Our educational content accurately reflects the histories and lived experiences of the learners we serve.
- Our educational products and services are inclusive and represent the rich diversity of learners.
- Our educational content prompts deeper discussions with students and motivates them to expand their own learning (and worldview).

Accessibility

We are also committed to providing products that are fully accessible to all learners. As per Pearson's guidelines for accessible educational Web media, we test and retest the capabilities of our products against the highest standards for every release, following the WCAG guidelines in developing new products for copyright year 2022 and beyond.

You can learn more about Pearson's commitment to accessibility at **https://www.pearson.com/us/accessibility.html**

Contact Us

While we work hard to present unbiased, fully accessible content, we want to hear from you about any concerns or needs with this Pearson product so that we can investigate and address them.

Please contact us with concerns about any potential bias at **https://www.pearson.com/report-bias.html**

For accessibility-related issues, such as using assistive technology with Pearson products, alternative text requests, or accessibility documentation, email the Pearson Disability Support team at **disability.support@pearson.com**

Pearson

This page intentionally left blank

Preface

About the Authors

Geoffrey E. Mills

Born in Australia, Geoff moved to the USA in 1986 to undertake doctoral studies at the University of Oregon. After completing his PhD in 1988, Geoff accepted his first teaching position at Southern Oregon State College (now Southern Oregon University). After 12 years of teaching, Geoff moved to university administration and served as Dean of the School of Education at Southern Oregon University for 12 years. Geoff returned to teaching and spent five years overseeing administrative licensure programs. Geoff is now an Emeritus Professor of Education at Southern Oregon University.

Geoff has traveled extensively and given invited research presentations in Australia, New Zealand, Greenland, United Kingdom, Canada, and many states in the US (including consultancies in the Marshall Islands, Guam, Saipan, Palau, American Samoa, and the US Virgin Islands). In addition to *Educational Research: Competencies for Analysis and Applications*, Geoff is the author of *Action Research: A Guide for the Teacher Researcher*, 6e (2018).

Adam W. Jordan

A native of beautiful Madison County, Georgia, Adam is a former middle and high school teacher with a focus on alternative education. He received his PhD in Education in 2013 from the University of North Carolina at Chapel Hill. Currently, he is an associate professor of Special Education at the College of Charleston in Charleston, South Carolina, where his research is focused on well-being in school spaces for teachers and students. In addition to his academic work, Adam co-directs a nonprofit called the *All Y'all Social Justice Collective*, where he works to support teachers and public education advocates in the Southern United States and Appalachia.

About This Book

Welcome to *Educational Research: Competencies for Analysis and Applications*. This text is designed primarily for use in the introductory course in educational research that is a basic requirement for many graduate programs. Because the topic coverage of the text is relatively comprehensive, it may be easily adapted for use in either a senior-level undergraduate course or a more advanced graduate-level course. Perhaps one of the first things that you will notice about this book is that it is in its 13th edition! It has truly stood the test of time since its first edition in 1976 by L.R. Gay and immediately became the benchmark for all educational research textbooks. During nearly a half century of iterations, the book has continued to evolve and reflect the growth of educational research designs and applications with diverse, equitable, and inclusive examples throughout the text.

The philosophy that guided the development of the current and previous editions of this text was the conviction that an introductory research course should be oriented more toward skill and application than toward theory. Thus, the purpose of this text is for students to become familiar with research mainly at a "how-to" skill and application level. The authors do not mystify students with theoretical and statistical jargon. They strive to provide a down-to-earth approach that helps students acquire the skills and knowledge required of a competent consumer and producer of educational research. The emphasis is not just on what the student knows but also on what the student can do with what they know. It is recognized that being a "good" researcher involves more than the acquisition of skills and knowledge; in any field, important

research is usually produced by those who through experience have acquired insights, intuitions, and strategies related to the research process. Research of any worth, however, is rarely conducted in the absence of basic research skills and knowledge. A fundamental assumption of this text is that the competencies required of a competent consumer of research overlap considerably with those required of a competent producer of research. A person is in a much better position to evaluate the work of others after they have performed the major tasks involved in the research process. The overall strategy of the text is to promote students' attainment of a degree of expertise in research through the acquisition of knowledge and by involvement in the research process.

Organization

In the 13th edition, Part I, Foundational Concepts and Processes, includes discussion of the scientific and disciplined inquiry approach and its application in education. The main steps in the research process and the purpose and methods of the various research designs are discussed. In Part I, each student selects and delineates a research problem of interest that has relevance to their professional area. Throughout the rest of the text, the student then simulates the procedures that would be followed in conducting a study designed to investigate the research problem; each chapter develops a specific skill or set of skills required for the execution of such a research design. Specifically, the student learns about the application of the scientific method in education (Chapter 1) and the ethical considerations that affect the conduct of any educational research (Chapter 2), identifies a research problem and formulates hypotheses (Chapter 3), conducts a review of the related literature (Chapter 4), develops a research plan (Chapter 5), selects and defines samples (Chapter 6), and evaluates and selects measuring instruments (Chapter 7). Throughout these chapters are parallel discussions of quantitative and qualitative research constructs. This organization, with increased emphasis on ethical considerations in the conduct of educational research and the skills needed to conduct a comprehensive review of related literature, allows the student to see the similarities and differences in research designs and to understand more fully how the nature of the research question influences the selection of a research design. Part II, Research Designs, includes description and discussion of different quantitative research designs, qualitative research designs, mixed methods research designs, and action research designs. Part III, Working with Quantitative and Qualitative Data, includes two chapters devoted to the statistical approaches and the analysis and interpretation of quantitative data, and two chapters describing the collection, analysis, and interpretation of qualitative data, as well as a new chapter focused on secondary data research. Part IV, Reporting and Critiquing Research, focuses on helping the student prepare a research report, either for the completion of a degree requirement or for publication in a refereed journal, and an opportunity for the student to apply the skills and knowledge acquired in Parts I through III to critique a research report.

Strategy

This text represents more than just a textbook to be incorporated into a course; it is a total instructional system that includes stated learning outcomes, instruction, and procedures for evaluating each outcome. The instructional strategy of the system emphasizes the demonstration of skills and individualization within this structure. Each chapter begins with a list of learning outcomes that describes the knowledge and skills that the student should gain from the chapter. In many instances, learning outcomes may be assessed either as written exercises submitted by students or by tests, whichever the instructor prefers. In most chapters, a task to be performed is described next. Tasks require students to demonstrate that they can perform particular research skills.

Because all students work with individual research problems, they can demonstrate the competency required by a task as it applies to their own problem. With the exception of Chapter 1, an individual chapter is directed toward the attainment of only one task (occasionally, students have a choice between a quantitative and qualitative task).

Text discussion is intended to be as simple and straightforward as possible. Whenever feasible, procedures are presented as a series of steps, and concepts are explained in terms of illustrative examples. In a number of cases, relatively complex topics or topics beyond the scope of the text are presented at a very elementary level, and students are directed to other sources for additional, in-depth discussion. There is also a degree of intentional repetition; a number of concepts are discussed in different contexts and from different perspectives. Also, at the risk of eliciting more than a few groans, an attempt has been made to sprinkle the text with touches of humor—a hallmark of this text spanning four decades—perhaps best captured by the pictures and quotes that open each chapter. Each chapter includes a detailed, often lengthy summary with headings and subheadings directly parallel to those in the chapter. The summaries are designed to facilitate both the review and location of related text discussion. Finally, each chapter (or part) concludes with suggested criteria for evaluating the associated task, along with an example of the task produced by a former introductory educational research student. Full-length articles, reprinted from the educational research literature, appear at the ends of all chapters presenting research designs and serve as illustrations of "real-life" research using that design. For the 13th edition all of these articles have been annotated with descriptive and evaluative annotations.

New to This Edition

Like the 12th edition, the 13th edition reflects a combination of both unsolicited and solicited input. Positive feedback suggested aspects of the text that should not be changed—the writing style and the focus on ethical practice, for example. Those aspects remain. Part I, Foundational Concepts and Processes, retains the same seven chapters from the 12th edition. Part II, Research Designs, includes all of the research design chapters that were covered in the 12th edition. Part III, Working with Quantitative and Qualitative Data, brings together discussions of descriptive statistics, inferential statistics, and qualitative data collection and analysis as well as a new chapter on secondary data research. Part IV, Reporting and Critiquing Research, effectively remains the same.

Content changes reflect the inclusion of new topics and the expansion or clarification of existing topics. There are many improvements in this edition, and we describe the more significant highlights here:

- **Chapter 20 is a new chapter titled Secondary Data Research.** This chapter reflects the changing educational research landscape that has been significantly impacted by the global COVID-19 pandemic. The chapter guides the researcher through the steps involved in analyzing previously existing data in new ways to answer new research questions and find new trends in data.
- **Chapter 4 has undergone significant revision because of the way technology has affected the literature review process.** The use of online and digital technologies is growing in popularity and effectiveness for researchers in the field of education and other disciplines. Changes include an expanded Digital Research Tools feature that covers annotation, brainstorming, citation management, organization, and writing management.
- **Chapter 12 on Single-Subject Experimental Research design has been significantly updated to reflect changes in the field.** The example article in this text has been updated to represent current trends in single-subject research.

- **Chapters 18 and 19 on descriptive and inferential statistics have been significantly updated,** and are focused on using R, Excel, and SPSS. R is a little different than Excel and SPSS in that it is open-source (read: free) but in order to operate R you need some very basic coding skills. This may frighten some readers from the start, but don't worry. We are going to send the reader off with a nice start to R programming and offer suggestions for future reading that will enhance their skillset.

Key Content Updates by Chapter

Chapter 1: Updated example abstracts to reflect current research trends, discussed the important role research plays in the real lives of students, particularly in a world influenced by a global pandemic.

Chapter 2: The updated chapter reflects the general principles from the Ethical Principles of Psychologists and Code of Conduct adopted by the American Psychological Association on January 1, 2017. This chapter also reflects the challenging ethical issues faced by researchers negotiating access during a pandemic and the heightened importance of informed consent and protection from harm for research subjects.

Chapter 3: Updated "Digital Research Tools" contributes to the researcher's thinking about selecting and defining a research problem and how these tools can connect researchers around the world.

Chapter 4: Updated use of online and digital technologies that contribute to the effectiveness of researchers in the field of education and other disciplines. Changes include an expanded Digital Research Tools feature that covers annotation, brainstorming, citation management, organization, and writing management.

Chapter 5: This chapter continues to provide a comprehensive discussion and templates for quantitative and qualitative researchers to develop, revise, and improve on their research plans.

Chapter 6: This chapter provides comprehensive, step-by-step directions on how to select samples in qualitative and quantitative research designs.

Chapter 7: This chapter provides a comprehensive discussion of the difference between constructs and variables and guides the reader through the process of selecting and developing measuring instruments.

Chapter 8: Updated coverage of web-based survey tools contributes to the reader's understanding of easily accessible, and user friendly, survey instruments.

Chapter 9: This chapter provides a comprehensive discussion of correlational research as well as relationship and prediction studies that use a correlational research design.

Chapter 10: This chapter is the seminal chapter on causal-comparative research and has stood the test of time! The chapter distinguishes between causal-comparative research and other quantitative research designs and provides an excellent example of causal-comparative research in practice.

Chapter 11: Another seminal chapter that has evolved over the past 45+ years but is largely considered to be the benchmark by which other authors measure their ability to comprehensively define experimental research designs and the threats to internal and external validity.

Chapter 12: This chapter, focused on single-subject research, includes further clarified examples and a newly annotated research article, which represents recent trends in single-subject research.

Chapter 13: Updated use of online digital research tools including Zoom, Otter.AI, and Dragon Anywhere to assist narrative researchers with data collection.

Chapter 14: This chapter provides a comprehensive discussion of the different types of ethnographic research and the use of ethnographic research techniques to produce an ethnographic study.

Chapter 15: This chapter provides a comprehensive description of the case study research process as well as how to conduct and analyze multiple case studies.

Chapter 16: This chapter describes in detail the three core mixed methods research designs: explanatory sequential, exploratory sequential, and convergent parallel and guides the reader through the process of determining the design of best fit for the research questions being considered.

Chapter 17: This chapter provides a comprehensive description of the action research process and is based on the dialectic action research spiral that is widely used by teacher researchers.

Chapter 18: With a focus on descriptive statistics, this chapter has been further clarified with updated R programming.

Chapter 19: Like Chapter 18, Chapter 19 has been brought up to date with recent R programming. The ease of use of this chapter, a strength noted by reviewers, has been maintained.

Chapter 20: This chapter represents a significant change to this edition as a new chapter on secondary data research. As the lives of educational researchers changed in the global COVID-19 pandemic, we realized that the introductory student could benefit greatly from a brief, but clear guide to the use of secondary data. We hope you enjoy this new addition to the text.

Chapter 21 (formerly Chapter 20): This chapter provides a comprehensive description of qualitative data collection sources and techniques as well as strategies to address the trustworthiness of qualitative research.

Chapter 22 (formerly Chapter 21): This chapter provides a comprehensive description of the definitions and purposes of qualitative data analysis and interpretation before, during, and after data collection and describes the steps to be followed to ensure the credibility of your qualitative research study.

Chapter 23 (formerly Chapter 22): This chapter provides guidelines for writing a research report and helps with the identification of target professional journals for publication.

Chapter 24 (formerly Chapter 23): Finally, this chapter provides guidelines for evaluating the major sections and subsections of a research report in order to self-evaluate a written research report.

In addition, we have added new tables and figures throughout the text. Every chapter has been edited and updated. References have been updated as well.

Pedagogical Features

The 13th edition includes several features that facilitate learning, including:

Research articles presented in their entirety: Each chapter provides one complete research article. Each article includes annotations from the authors that illustrate key points that will aid comprehension of important features of high-quality research. Articles have been selected that highlight important, contemporary concerns within the field, with a particular emphasis on research that focuses on equitable teaching principles.

Tasks guide students through the research process: Chapters include a running list of Tasks that help students break the research process into smaller objectives. By taking this step-by-step approach, students have ample opportunity to apply what they are learning in the text to their larger research interests. Suggested responses to select Tasks in Appendix B allow students the opportunity to evaluate their comprehension.

Pearson eText, Learning Management System (LMS)-Compatible Assessment Bank, and Other Instructor Resources Pearson eText

The Pearson eText is a simple-to-use, mobile-optimized, personalized reading experience. It allows you to easily highlight, take notes, and review key vocabulary all in one place—even when offline. Seamlessly integrated videos and other rich media will engage you and give you access to the help you need, when you need it. To gain access or to sign in to your Pearson eText, visit: https://www.pearson.com/pearson-etext
Podcasts: New to this edition is a series of podcasts that include the authors having conversations with various methodologists. These podcasts allow the reader to gain insight into the lives and careers of educational researchers. It is our hope that these podcasts connect the reader to the content in a way that bridges the gap between the theoretical and the practical.

Podcasts: As mentioned earlier, new to this edition is an introduction to Geoff and Adam's podcast "Talking Educational Research with Geoff and Adam." These podcasts are designed to connect the reader with the "real life" stories of methodologists in the field. Enjoy an introduction podcast as well as method-specific episodes in qualitative research, quantitative research, action research, and single-case design.

Artifacts: As a supplement to the digital text, R files are provided as accompaniments to the R examples found in text. These files allow the reader to download "ready to go" files that can be used alongside the examples as an enrichment exercise.

LMS-Compatible Assessment Bank

With this new edition, all assessment types—quizzes and application exercises—are included in LMS-compatible banks for the following learning management systems: Blackboard (9780137535323) Canvas (9780137535330), D2L (9780137535354) and Moodle (9780137535392). These packaged files allow maximum flexibility to instructors when it comes to importing, assigning, and grading. Assessment types include:

- **Learning Outcome Quizzes:** Each chapter learning outcome is the focus of a Learning Outcome Quiz that is available for instructors to assign through their Learning Management System. Learning outcomes identify chapter content that is most important for learners and serve as the organizational framework for each chapter. Each Learning Outcome is accompanied by a 4–5 question multiple-choice quiz with feedback, allowing the reader to further test their knowledge of the section content.
- **Application Exercises:** At least two application exercises are provided in each chapter. These exercises allow the reader to apply the knowledge gained in the chapter to either a case study or other extended response type item. A model response written by an expert is provided to help guide your learning.
- **Chapter Tests:** Suggested test items are provided for each chapter in multiple-choice format.

Instructor's Manual (9780137535415)

The Instructor's Manual is provided as a Word document and includes resources to assist professors in planning their course.

PowerPoint® Slides (9780137535460)

PowerPoint® slides are provided for each chapter and highlight key concepts and summarize the content of the text to make it more meaningful for students.

Note: All instructor resources—LMS-compatible assessment bank, instructor's manual, and PowerPoint slides—are available for download at www.pearsonhighered.com. Use one of the following methods:

From the main page, use the search function to look up the lead author, or the title. Select the desired search result, then access the "Resources" tab to view and download all available resources.

From the main page, use the search function to look up the ISBN (provided above) of the specific instructor resource you would like to download. When the product page loads, access the "Downloadable Resources" tab.

This page intentionally left blank

Brief Contents

This page intentionally left blank

Contents

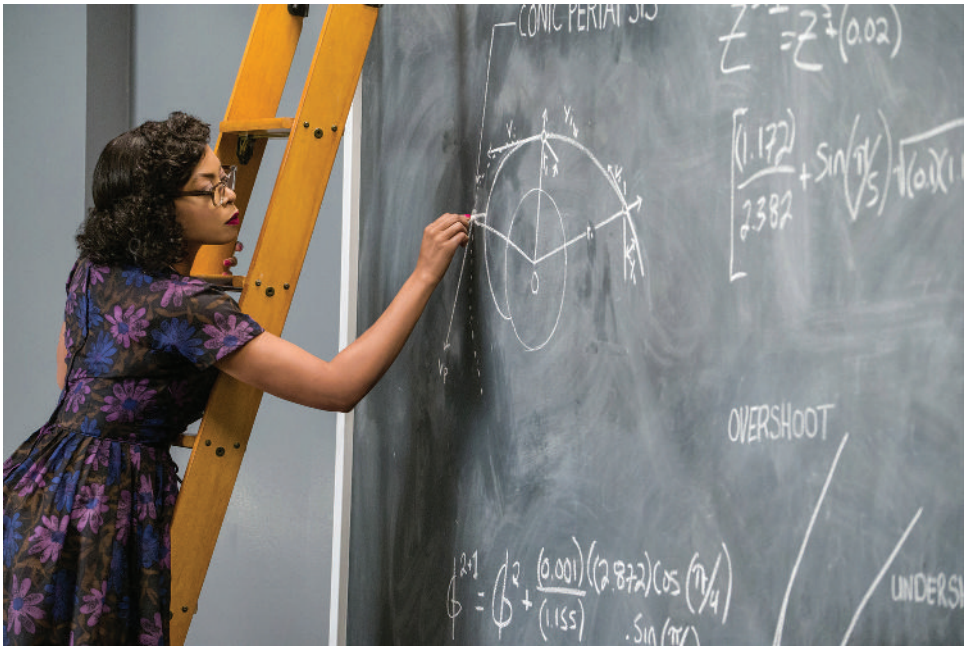
This page intentionally left blank

Research Articles

This page intentionally left blank

Chapter 1

Introduction to Educational Research



Hopper Stone/TM & 20th Century Fox Film Corp/Everett Collection

Hidden Figures, 2016

"Despite a popular stereotype that depicts researchers as spectacled, stoop-shouldered, elderly gentlemen… who endlessly add chemicals to test tubes, every day thousands of men and women of all ages and postures conduct educational research in a wide variety of settings." (p. 3)

Learning Outcomes

After reading Chapter 1, you should be able to do the following:

1.1 Briefly describe the reasoning involved in the scientific method.

1.2 Explain why researchers would use quantitative, qualitative, mixed methods, or action research designs to address a specific research problem.

1.3 Briefly define and state the major characteristics of these research designs: survey, correlational, causal–comparative, experimental, single-subject, narrative, ethnographic, case study, mixed methods, and action research.

1.4 Explain the purposes of basic research, applied research, evaluation research, research and development (R&D), and action research.

Completing Chapter 1 should enable you to perform the following tasks:

TASKS 1A, 1B

Identify and briefly state the following for both research studies at the end of this chapter:

1. The research design
2. The rationale for the choice of the research design
3. The major characteristics of the research design, including research procedures, method of analysis, and major conclusions (See Performance Criteria, p. 18.)

TASK 1C

Classify given research studies based on their characteristics and purposes. (See Performance Criteria, p. 18.)

Welcome!

If you are taking a research course because it is required in your program of studies, raise your right hand. If you are taking a research course because it seems like it will be a really fun elective, raise your left hand. We thought you may not be here of your own free will. Although you may be required to take this course, you are not the innocent victim of one or more sadists. Your professors have several legitimate reasons for believing this research course is an essential component of your education.

First, educational research findings contribute significantly to both educational theory and educational practice. As a professional, you need to know how to find, understand, and evaluate these findings. And when you encounter research findings in professional publications or in the media, you have a responsibility, as a professional, to distinguish between legitimate and ill-founded research claims. Second, although many of you will be primarily critical consumers of research, some of you will decide to become educational researchers. A career in research opens the door to a variety of employment opportunities in universities, research centers, and business and industry.

At this time in the history of the world, there is perhaps an unprecedented need for critical thinking and valid research. The world has faced enormous challenges related to the Novel Covid-19 pandemic that has killed millions of people. K–12 schools and universities around the world are facing fiscal, cultural, and educational peril. The ways we have traditionally thought about teaching and learning for all student populations are being challenged, as are the ways we have traditionally thought about research designs. Words like *mask, social distance, hand sanitizer, touchless, isolation*, and *high-risk group* are part of our everyday conversations. And increasingly, political and medical pundits bombard the airways with references to "random trials" and "clinical" research focused on the development of therapeutics and vaccines. Perhaps now more than ever educational researchers need to better understand what works, and what doesn't work, in 21st century classrooms that no longer can rely solely on face-to-face interactions with students, and home environments where working parents are now also expected to teach home schoolers! In short, the need for thoughtful, socially responsive educational researchers has never been greater. Engaging in meaningful educational research endeavors is destined to make a powerful difference for teachers, students, and parents during these difficult times.

Despite a popular stereotype that depicts researchers as spectacled, stoop-shouldered, elderly gentlemen (a stereotype I am rapidly approaching!) who endlessly add chemicals to test tubes, every day thousands of men and women of all ages and postures conduct educational research in a wide variety of settings. Every year many millions of dollars are spent in the quest for knowledge related to teaching and learning. For example, in 2020 the U.S. Department of Education budget was $64 billion, including an allocation of $187.5 million for research, development and dissemination (https://www2.ed.gov/about/overview/budget/budget20/summary/20summary.pdf). Educational research has contributed many findings concerning principles of behavior, learning, and retention of knowledge—many of which can also be applied to curriculum, instruction, instructional materials, and assessment techniques. Both the quantity and the quality of research are increasing, partly because researchers are better trained. Educational research classes have become core components of preservice teacher education programs, as well as the cornerstone of advanced degree programs.

We recognize that educational research is a relatively unfamiliar discipline for many of you. Our first goals, then, are to help you acquire a general understanding of research processes and to help you develop the perspective of a researcher. We begin by examining the scientific method.

The Scientific Method

Learning Outcome 1.1 **Briefly describe the reasoning involved in the scientific method.**

What is knowledge? And how do we come to "know" something? Experience is certainly one of the fundamental ways we come to know about and understand our world. For example, a child who touches something hot learns that high heat hurts. We know other things because a trusted authority, such as a parent or a teacher, told us about them. Most likely, much of your knowledge of current world events comes secondhand, from things you

have read or heard from a source you trust. Or, increasingly, it might have come from unvetted sources of social media like Facebook and Instagram.

Another way we come to know something is through thinking, through reasoning. Reasoning refers to the process of using logical thought to reach a conclusion. We can reason *inductively* or *deductively*. **Inductive reasoning** involves developing generalizations based on observation of a limited number of related events or experiences. Consider the following example of inductive reasoning:

Observation: An instructor examines five research textbooks. Each contains a chapter about sampling.

Generalization: The instructor concludes that all research textbooks contain a chapter about sampling.

Deductive reasoning involves essentially the reverse process—arriving at specific conclusions based on general principles, observations, or experiences (i.e., generalizations)—as shown in the next example.

Observations: All research textbooks contain a chapter on sampling. The book you are reading is a research text.

Generalization: This book must contain a chapter on sampling. (Does it?)

Although people commonly use experience, authority, inductive reasoning, and deductive reasoning to learn new things and draw new conclusions from that knowledge, each of these approaches to understanding has limitations when used in isolation. Some problems associated with experience and authority as sources of knowledge are graphically illustrated in a story told about Aristotle. According to the story, one day Aristotle caught a fly and carefully counted and recounted the legs. He then announced that flies have five legs. No one questioned the word of Aristotle. For years his finding was accepted uncritically. Unfortunately, the fly that Aristotle caught just happened to be missing a leg! Whether or not you believe the story, it illustrates the limitations of relying on personal experience and authority as sources of knowledge.

The story also points out a potential problem with inductive reasoning: Generalizing from a small sample, especially one that is atypical, can lead to errors. Deductive reasoning, too, is limited by the evidence in the original observations. If every research text really does have a chapter on sampling, and if this book really is a research text, then it follows that this book must have a chapter on sampling. However, if one or more of the premises is false (perhaps some research texts do not have a chapter on sampling), your conclusion may also be wrong.

When we rely exclusively on these common approaches to knowing, the resulting knowledge is susceptible to error and may be of limited value to understanding the world beyond our immediate experience. However, experience, authority, and inductive and deductive reasoning are very effective when used together as integral components of the scientific method. The **scientific method** is an orderly process entailing a number of steps: recognition and definition of a problem, formulation of hypotheses, collection of data, analysis of data, and statement of conclusions regarding confirmation or disconfirmation of the hypotheses (i.e., a researcher forms a **hypothesis**—an explanation for the occurrence of certain behaviors, phenomena, or events as a way of predicting the results of a research study and then collects data to test that prediction). These steps can be applied informally to solve everyday problems such as the most efficient route to take from home to work or school, the best time to go to the bank, or the best kind of computer to purchase. The more formal application of the scientific method is standard in research; it is more efficient and more reliable than relying solely on experience, authority, inductive reasoning, and deductive reasoning as sources of knowledge.

Limitations of the Scientific Method

The steps in the scientific method guide researchers in planning, conducting, and interpreting research studies. However, it is important to recognize some limitations of the method. First, the scientific method cannot answer all questions. For example, applying the scientific method will not resolve the question "Should we legalize euthanasia?" The answers to questions like this one are influenced by personal philosophy, values, and ethics.

Second, application of the scientific method can never capture the full richness of the individuals and the environments under study. Although some applications of the method lead to deeper understanding of the research context than others, no application—and in fact no research approach—provides full comprehension of a site and its inhabitants. No matter how many variables one studies or how long one is immersed in a research context, other variables and aspects of context will remain unexamined. Thus, the scientific method and, indeed, all types of inquiry give us a simplified version of reality.

Third, our measuring instruments always have some degree of error. The variables we study are often proxies for the real behavior we seek to examine. For example, even if we use a very precisely constructed multiple-choice test to assess a person's values, we will likely gather information that gives us a picture of that person's beliefs about their values. However, we aren't likely to have an adequate picture of how that person acts, which may be the better reflection of the person's real values.

More broadly, all educational inquiry, not just the scientific method, is carried out with the cooperation of participants who agree to provide researchers with

data. Because educational researchers deal with human beings, they must consider a number of ethical concerns and responsibilities to the participants. For example, they must shelter participants from real or potential harm. They must inform participants about the nature of the planned research and address the expectations of the participants. These factors can limit and skew results. All these limitations will be addressed in later sections of this book.

Application of the Scientific Method in Education

Research is the formal, systematic application of the scientific method to the study of problems; **educational research** is the formal, systematic application of the scientific method to the study of educational problems. The goal of educational research is essentially the same as the goal of all science: to describe, explain, predict, or control phenomena—in this case, educational phenomena. As we mentioned previously, it can be quite difficult to describe, explain, predict, and control situations involving human beings, who are by far the most complex of all organisms. So many factors, known and unknown, operate in any educational environment that it can be extremely difficult to identify specific causes of behaviors or to generalize or replicate findings. The kinds of rigid controls that can be established and maintained in a biochemistry laboratory, for instance, are impossible in an educational setting. Even describing behaviors, based on observing people, has limits. Observers may be subjective in recording behaviors, and people who are observed may behave atypically just because they are being watched. Chemical reactions, on the other hand, are certainly not aware of being observed! Nevertheless, behavioral research should not be viewed as less scientific than natural science research conducted in a lab.

Despite the difficulty and complexity of applying the scientific method in educational settings, the steps of the scientific method used by educational researchers are the same as those used by researchers in other more easily controlled settings:

1. *Selection and definition of a problem.* A problem is a question of interest that can be tested or answered through the collection and analysis of data. Upon identifying a research question, researchers typically review previously published research on the same topic and use that information to hypothesize about the results. In other words, they make an educated guess about the answer to the question.
2. *Execution of research procedures.* The procedures reflect all the activities involved in collecting data related to the problem (e.g., how data are collected and from whom). To a great extent, the specific procedures are dictated by the research question and the variables involved in the study.
3. *Analysis of data.* Data are analyzed in a way that permits the researcher to test the research hypothesis or answer the research question. Analysis usually involves application of one or more statistical techniques. For some studies, data analysis involves verbal synthesis of narrative data; these studies typically involve new insights about the phenomena in question, generate hypotheses for future research, or both.
4. *Drawing and stating conclusions.* The conclusions, which should advance our general knowledge of the topic in question, are based on the results of data analysis. They should be stated in terms of the original hypothesis or research question. Conclusions should indicate, for example, whether the research hypothesis was supported. For studies involving verbal synthesis, conclusions are much more tentative.

Different Approaches to Educational Research

Learning Outcome 1.2 **Explain why researchers would use quantitative, qualitative, mixed methods, or action research designs to address a specific research problem.**

All educational inquiry ultimately involves a decision to study or describe something—to ask some question and seek an answer. All educational inquiry necessitates that data of some kind be collected, that the data be analyzed in some way, and that the researcher come to some conclusion or interpretation. In other words, all educational inquiry shares the same four basic actions we find in the scientific method. However, it is not accurate to say that all educational research is an application of the scientific method. Important differences exist between the types of problems researchers investigate and the questions they ask, the types of data they collect, the form of data analysis, and the conclusions that the researcher can draw meaningfully and with validity.

The Continuum of Research Philosophies

Historically, educational researchers used approaches that involved the use of the scientific method. However, over the last four decades, researchers have adopted diverse philosophies toward their research. Now, there are certain philosophical assumptions that underpin an educational researcher's decision to conduct research. These philosophical assumptions address issues related to the nature of reality (ontology), how researchers know what they know

(epistemology), and the methods used to study a particular phenomenon (methodology), with an emphasis on quantitative or qualitative methods. As Creswell[1] notes, historically, researchers compared the philosophical assumptions that underpinned qualitative and quantitative research approaches in order to establish the legitimacy of qualitative research, but given the evolution of qualitative and quantitative research over the past four decades, there is no longer any need to justify one set of philosophical assumptions over another set of assumptions.

Quantitative Research

Educational researchers have also followed well-defined, widely accepted procedures for stating research problems, carrying out the research process, analyzing the resulting data, and verifying the quality of the study and its conclusions. Often, these research procedures are based on what has come to be known as a *quantitative approach* to conducting and obtaining educational understandings. The quantitative framework in educational research involves the application of the scientific method to try to answer questions about education. At the end of this chapter you will find an example of quantitative research published in *Child Development* (a refereed journal): "Can Instructional and Emotional Support in the First-Grade Classroom Make a Difference for Children at Risk of School Failure?" (Hamre & Pianta, 2005). As the title suggests, this research investigates the ways in which children's risk of school failure may be moderated by instructional and emotional support from teachers.

Quantitative research is the collection and analysis of numerical data to describe, explain, predict, or control phenomena of interest. Part II of the text will address in detail specific quantitative research designs that satisfy the assumptions underpinning a quantitative approach to research. A quantitative research design entails more than just the use of numerical data. At the outset of a study, quantitative researchers state the hypotheses to be examined and specify the research procedures that will be used to carry out the study. They also maintain control over contextual factors that may interfere with the data collection and identify a sample of participants large enough to provide statistically meaningful data. Many quantitative researchers have little personal interaction with the participants they study because they frequently collect data using paper-and-pencil, noninteractive instruments. The analysis of numerical data can be complex, but addressed systematically, it can be manageable and Part III of the text will provide a detailed description for how to work with quantitative data.

Underlying quantitative research methods is the philosophical belief or assumption that we inhabit a relatively stable, uniform, and coherent world that we can measure, understand, and generalize about. This view, adopted from the natural sciences, implies that the world and the laws that govern it are somewhat predictable and can be understood by scientific research and examination. In this quantitative perspective, claims about the world are not considered meaningful unless they can be verified through direct observation.

Qualitative Research

Qualitative research is the collection, analysis, and interpretation of comprehensive narrative and visual (i.e., nonnumerical) data to gain insights into a particular phenomenon of interest. Part II of the text will address in detail specific qualitative research designs that satisfy the underpinning assumptions of a qualitative approach to research. Qualitative research designs are based on different beliefs and purposes than quantitative research designs. For example, qualitative researchers do not necessarily accept the view of a stable, coherent, uniform world. They argue that all meaning is situated in a particular perspective or context, and because different people and groups often have different perspectives and contexts, the world has many different meanings, none of which is necessarily more valid or true than another.

Qualitative research designs tend to evolve as understanding of the research context and participants deepens (think back to the discussion of inductive reasoning). As a result, qualitative researchers often avoid stating hypotheses before data are collected, and they may examine a particular phenomenon without a guiding statement about what may or may not be true about that phenomenon or its context. However, qualitative researchers do not enter a research setting without any idea of what they intend to study. Rather, they commence their research with "foreshadowed problems."[2] This difference is important—quantitative research usually tests a specific hypothesis; qualitative research often does not.

Additionally, in qualitative research, context is not controlled or manipulated by the researcher. The effort to understand the participants' perspectives requires researchers using qualitative methods to interact extensively and intimately with participants during the study, using time-intensive data collection methods such as interviews and observations. As a result, the number of participants tends to be small, and qualitative researchers analyze the data inductively by categorizing and organizing it into patterns that produce a descriptive, narrative synthesis.

[1] *Qualitative Inquiry & Research Design: Choosing Among Five Approaches* (4th ed.) by J. W. Creswell and C. N. Poth, 2018, Thousand Oaks, CA: Sage.

[2] *Argonauts of the Western Pacific* (p. 9), by B. Malinowski, 1922, London: Routledge.

Qualitative research differs from quantitative research in two additional ways: (1) Qualitative research often involves the simultaneous collection of a wealth of narrative and visual data over an extended period of time, and (2) as much as is possible, data collection occurs in a naturalistic setting. In quantitative studies, in contrast, research is most often conducted in researcher-controlled environments under researcher-controlled conditions, and the activities of data collection, analysis, and writing are separate, discrete activities. Because qualitative researchers strive to study people and events in their naturalistic settings, qualitative research is sometimes referred to as *naturalistic research, naturalistic inquiry*, or *field-oriented research*.

These two characteristics of qualitative research, the simultaneous study of many aspects of a phenomenon and the attempt to study things as they exist naturally, help in part to explain the growing enthusiasm for qualitative research in education, especially in applied teacher practitioner–oriented research. Some researchers and educators feel that certain kinds of educational problems and questions do not lend themselves well to quantitative methods, which use principally numerical analysis and try to control variables in very complex environments. As qualitative researchers point out, to have relevance to real-world settings, findings should be derived from research conducted in real-world settings.

At the end of this chapter, you will find an example of qualitative research published in *Action in Teacher Education* (a refereed journal): "Developing Teacher Epistemological Sophistication about Multicultural Curriculum: A Case Study" (Sleeter, 2009). This research investigates how teachers' thinking about curriculum develops during a teacher preparation program and how the lessons from the case study might inform teacher education pedagogy. And, of course, the use of the word *epistemological* in the title introduces you to the language of educational research!

Mixed Methods Research

Mixed methods research combines quantitative and qualitative designs by including both quantitative and qualitative data in a single study. The purpose of mixed methods research is to build on the synergy and strength that exists between quantitative and qualitative research designs to understand a phenomenon more fully than is possible using either quantitative or qualitative approaches alone. Chapter 16 will describe in detail six mixed methods research designs (convergent-parallel, explanatory, exploratory, experimental, social justice, and multistage evaluation). However, the basic differences among the designs are related to the priority given to the following areas:

- The type of data collected (i.e., qualitative and quantitative data are of equal weight, or one type of data has greater weight than the other)
- The sequence of data collection (i.e., both types of data are collected during the same time period, or one type of data is collected in each sequential phase of the project)
- The analysis techniques (i.e., either an analysis that combines the data or one that keeps the two types of data separate)

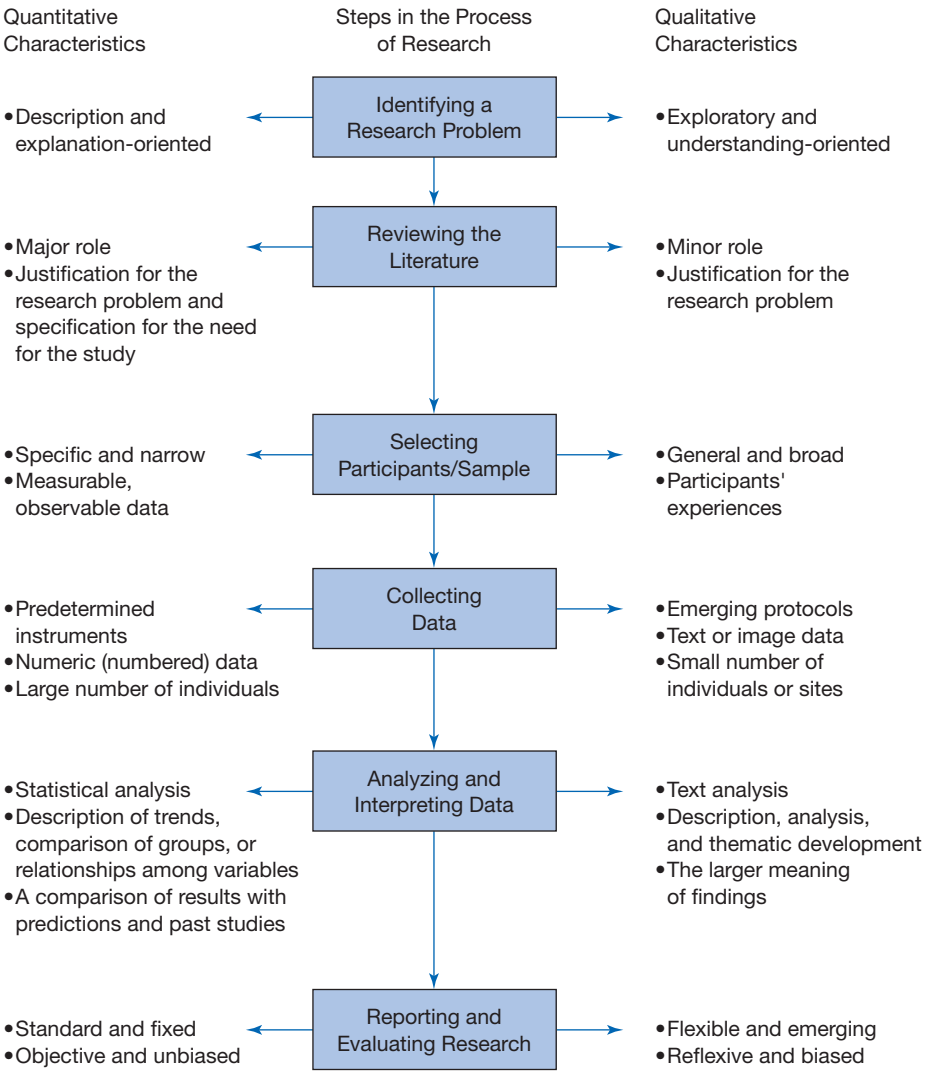
Characteristics of Quantitative and Qualitative Research Approaches

Earlier in this chapter, we presented four general, conceptual research steps used in the scientific method. In this section, we expand the number of steps to six, which are followed by both quantitative researchers and qualitative researchers. As we discuss in subsequent chapters in Part II, however, the application of the steps differs depending on the research design. For example, the research procedures in qualitative research are often less rigid than those in quantitative research. Similarly, although both quantitative and qualitative researchers collect data, the nature of the data differs. Figure 1.1 compares the six steps of qualitative and quantitative research approaches and lists traits that characterize each approach at every step:

1. *Identifying a research topic/problem*
2. *Reviewing the literature*
3. *Selecting participants/sample*
4. *Collecting data*
5. *Analyzing and interpreting data*
6. *Reporting and evaluating the research*

Table 1.1 provides another snapshot of quantitative and qualitative research characteristics. Despite the differences between them, you should not consider quantitative and qualitative research approaches to be oppositional. Taken together, they represent the full range of educational research designs. The terms *quantitative* and *qualitative* are used to differentiate one approach from the other conveniently. If you see yourself as a positivist—the belief that qualities of natural phenomena must be verified by evidence before they can be considered knowledge—that does not mean you cannot use or learn from qualitative research methods. The same holds true for nonpositivist, phenomenologist qualitative researchers. Depending on the nature of the question, topic, or problem to be investigated, one of these approaches will generally be more appropriate than the other, although selecting a primary approach does not preclude borrowing from the other. In fact, both may be utilized in the same studies, as when the administration of a (quantitative) questionnaire is followed by a small number of detailed (qualitative) interviews to obtain deeper explanations for the numerical data.

Figure 1.1 Characteristics of quantitative and qualitative research



SOURCE: *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (6th ed.), (pp. 20, 464, 504, 541), by John W. Creswell & Guetterman, T.C. © 2019. Reprinted by permission of Pearson Education, Inc., Upper Saddle River, NJ.

Table 1.1 Overview of qualitative and quantitative research characteristics

	Quantitative Research	Qualitative Research
Type of data collected	Numerical data	Nonnumerical narrative and visual data
Research problem	Hypothesis and research procedures stated before beginning the study	Research problems and methods evolve as understanding of topic deepens
Manipulation of context	Yes	No
Sample size	Larger	Smaller
Research procedures	Relies on statistical procedures	Relies on categorizing and organizing data into patterns to produce a descriptive, narrative synthesis
Participant interaction	Little interaction	Extensive interaction
Underlying belief	We live in a stable and predictable world that we can measure, understand, and generalize about.	Meaning is situated in a particular perspective or context that is different for people and groups; therefore, the world has many meanings.

Classification of Research by Design

Learning Outcome 1.3 **Briefly define and state the major characteristics of these research designs: survey, correlational, causal–comparative, experimental, single-subject, narrative, ethnographic, case study, mixed methods, and action research.**

A research design comprises the overall strategy followed in collecting and analyzing data. Although there is some overlap, most research studies follow a readily identifiable design. The largest distinction we can make in classifying research by design is the distinction between quantitative and qualitative approaches. Quantitative and qualitative research approaches, in turn, include several distinct types or designs with a focus on unique research problems.

Quantitative Approaches

Quantitative research approaches are applied to describe current conditions, investigate relations, and study cause–effect phenomena. Survey research is often designed to describe current conditions. Studies that investigate the relations between two or more variables are correlational research. Experimental studies and causal–comparative studies provide information about cause–effect outcomes. Studies that focus on the behavior change an individual exhibits as a result of some intervention fall under the heading of single-subject research.

SURVEY RESEARCH **Survey research** determines and reports the way things are; it involves collecting numerical data to test hypotheses or answer questions about the current status of the subject of study. One common type of survey research involves assessing the preferences, attitudes, practices, concerns, or interests of a group of people. A pre-election political poll and a survey about community members' perception of the quality of the local schools are examples. Survey research data are mainly collected through questionnaires, interviews, and observations.

Although survey research sounds very simple, there is considerably more to it than just asking questions and reporting answers. Because researchers often ask questions that have not been asked before, they usually have to develop their own measuring instrument for each survey study. Constructing questions for the intended respondents requires clarity, consistency, and tact. Other major challenges facing survey researchers are participants' failure to return questionnaires, their willingness to be surveyed over the phone, and their ability to attend scheduled interviews. If the response rate is low, then valid, trustworthy conclusions cannot be drawn. For example, suppose you are doing a study to determine the attitudes of principals toward research in their schools. You send a questionnaire to 100 principals and include the question "Do you usually cooperate if your school is asked to participate in a research study?" Forty principals respond, and they all answer "Yes." It's certainly a mistake to conclude that principals in general cooperate. Although all those who responded said yes, those 60 principals who did not respond may never cooperate with researchers. After all, they didn't cooperate with you! Without more responses, it is not possible to make generalizations about how principals feel about research in their schools.

Following are examples of questions that can be investigated in survey research studies, along with typical research designs:

- *How do second-grade teachers spend their teaching time?* Second-grade teachers are asked to fill out questionnaires, and results are presented as percentages (e.g., teachers spent 50% of their time lecturing, 20% asking or answering questions, 20% in discussion, and 10% providing individual student help).
- *How will citizens of Yourtown vote in the next school board election?* A sample of Yourtown citizens complete a questionnaire or interview, and results are presented as percentages (e.g., 70% said they will vote for Peter Pure, 20% named George Graft, and 10% are undecided).

  Survey research is described in more detail in Chapter 8.

CORRELATIONAL RESEARCH **Correlational research** involves collecting data to determine whether, and to what degree, a relation exists between two or more quantifiable variables. A **variable** is a conceptthat can assume any one of a range of values; for example, intelligence, height, and test score are variables. At a minimum, correlational research requires information about at least two variables obtained from a single group of participants.

The purpose of a correlational study may be to establish relations or use existing relations to make predictions. For example, a college admissions director may be interested in answering the question "How do the SAT scores of high school seniors correspond to the students' first-semester college grades?" If students' SAT scores are strongly related to their first-semester grades, SAT scores may be useful in predicting how students will perform in their first year of college. On the other hand, if there is little or no correlation between the two variables, SAT scores likely will not be useful as predictors.

Correlation refers to a quantitative measure of the degree of correspondence. The degree to which two variables are related is expressed as a **correlation coefficient**, which is a number between +1.00 and −1.00. Two variables that are not related have a correlation coefficient near

0.00. Two variables that are highly correlated will have a correlation coefficient near +1.00 or −1.00. A number near +1.00 indicates a positive correlation: As one variable increases, the other variable also increases (e.g., students with high SAT scores may also have high grade point averages [GPAs]). A number near −1.00 indicates a negative correlation: As one variable increases, the other variable decreases (e.g., a high GPA may correlate negatively with the likelihood of dropping out). Because very few pairs of variables are perfectly correlated, predictions based on them are rarely +1.0 or −1.0.

It is very important to note that the results of correlational studies do not suggest cause–effect relations among variables. Thus, a positive correlation between, for example, self-concept and achievement does not imply that self-concept causes achievement or that achievement causes self-concept. The correlation indicates only that students with higher self-concepts tend to have higher levels of achievement and that students with lower self-concepts tend to have lower levels of achievement. We cannot conclude that one variable is the cause of the other.

Following are examples of research questions tested with correlational studies:

- *What is the relation between intelligence and self-esteem?* Scores on an intelligence test and a measure of self-esteem are acquired from each member of a given group. The two sets of scores are analyzed, and the resulting coefficient indicates the degree of correlation.
- *Does an algebra aptitude test predict success in an algebra course?* Scores on the algebra aptitude test are correlated with final exam scores in the algebra course. If the correlation is high, the aptitude test is a good predictor of success in algebra.

Correlational research is described in detail in Chapter 9.

CAUSAL–COMPARATIVE RESEARCH **Causal–comparative research** attempts to determine the cause, or reason, for existing differences in the behavior or status of groups of individuals. The cause is a behavior or characteristic believed to influence some other behavior or characteristic and is known as the **grouping variable**. The change or difference in a behavior or characteristic that occurs as a result—that is, the effect—is known as the **dependent variable**. Put simply, causal–comparative research attempts to establish cause–effect relations among groups.

Following are examples of research questions tested with causal–comparative studies (note that the word is *causal*, not *casual*):

- *How does preschool attendance affect social maturity at the end of the first grade?* The grouping variable is preschool attendance (i.e., the variable can take one of two values—students attending preschool and students not attending); the dependent variable, or effect, is social maturity at the end of the first grade. The researcher identifies a group of first-graders who attended preschool and a group who did not, gathers data about their social maturity, and then compares the two groups.
- *How does having a long-term, uncertified substitute teacher affect a child's school absenteeism?* The grouping variable is the certification status of the teacher (again with two possible values—the teacher is either certified or not); the dependent variable is absenteeism, measured as number of days absent. The researcher identifies a group of students who have long-term, uncertified substitute teachers and a group who has a certifed teacher, gathers information about their absenteeism, and compares the groups.

A weakness of causal–comparative studies is that, because the cause under study has already occurred, the researcher has no control over it. For example, suppose a researcher wanted to investigate the effect of heavy smoking on lung cancer and designs a study comparing the frequency of lung cancer diagnoses in two groups, long-time smokers and nonsmokers. Because the groups are preexisting, the researcher did not control the conditions under which the participants smoked or did not smoke (this lack of researcher control is why the variable is known as a grouping variable rather than an independent variable). Perhaps a large number of the long-time smokers lived in a smoggy, urban environment, whereas only a few of the nonsmokers were exposed to those conditions. In that case, attempts to draw cause–effect conclusions in the study would be tentative at best. Is it smoking that causes higher rates of lung cancer? Is it living in a smoggy, urban environment? Or is it some unknown combination of smoking and environment? A clear cause–effect link cannot be obtained.

Although causal–comparative research produces limited cause–effect information, it is an important form of educational research. True cause–effect relations can be determined only through experimental research (discussed in the next section), in which the researcher maintains control of an independent variable; but in many cases, an experimental study is inappropriate or unethical. The causal–comparative approach is chosen precisely because the grouping variable either cannot be manipulated (e.g., as with gender, height, or year in school) or should not be manipulated (e.g., as with smoking or prenatal care). For example, to conduct the smoking study as an experiment, a researcher would need to select a large number of participants who had never smoked and divide them into two groups, one directed to smoke heavily and one forbidden to smoke. Obviously, such a study is unethical because of the potential harm to those forced to smoke.

A causal–comparative study, which approximates cause–effect results without harming the participants, is the only reasonable approach. Like descriptive and correlational studies, however, causal–comparative research does not produce true experimental research outcomes.

Causal–comparative research is described in detail in Chapter 10.

EXPERIMENTAL RESEARCH In **experimental research**, at least one independent variable is manipulated, other relevant variables are controlled, and the effect on one or more dependent variables is observed. True experimental research provides the strongest results of any of the quantitative research approaches because it yields clear evidence for linking variables. As a result, it also offers **generalizability**, or applicability of findings to settings and contexts different from the one in which they were obtained.

Unlike causal–comparative researchers, researchers conducting an experimental study can control an independent variable. They can select the participants for the study, divide the participants into two or more groups that have similar characteristics at the start of the research experiment, and then apply different treatments to the selected groups. They can also control the conditions in the research setting, such as when the treatments will be applied, by whom, for how long, and under what circumstances. Finally, the researchers can select tests or measurements to collect data about any changes in the research groups. The selection of participants from a single pool of participants and the ability to apply different treatments or programs to participants with similar initial characteristics permit experimental researchers to draw conclusions about cause and effect. The essence of experimentation is control, although in many education settings it is not possible or feasible to meet the stringent control conditions required by experimental research.

Following are examples of research questions that are explored with experimental studies:

- *Is personalized instruction from a teacher more effective for increasing students' computational skills than computer instruction?* The independent variable is type of instruction (with two values: personalized instruction and computer instruction); the dependent variable is computational skills. A group of students who have never experienced either personalized teacher instruction or computer instruction are selected and randomly divided into two groups, each taught by one of the methods. After a predetermined time, the students' computational skills are measured and compared to determine which treatment, if either, produced higher skill levels.
- *Is there an effect of reinforcement on students' attitude toward school?* The independent variable is type of reinforcement (with three values: positive, negative, or no reinforcement); the dependent variable is attitude toward school. The researcher randomly forms three groups from a single large group of students. One group receives positive reinforcement, another negative reinforcement, and the third no reinforcement. After the treatments are applied for a predetermined time, student attitudes toward school are measured and compared for each of the three groups.

Experimental research is described in detail in Chapter 11.

SINGLE-SUBJECT RESEARCH Rather than compare the effects of different treatments (or treatment versus no treatment) on two or more groups of people, experimental researchers sometimes compare a single person's behavior before treatment to behavior exhibited during the course of the experiment. They may also study a number of people together as one group, rather than as individuals. **Single-subject experimental designs** are those used to study the behavior change that an individual or group exhibits as a result of some intervention or treatment. In these designs, the size of the **sample**—the individuals selected from a population for a study—is said to be one.

Following are examples of published studies that used single-subject designs:

- *The effects of eye movement games on the visual capabilities of students with low vision. One student with low vision was studied.*[3]
- *The effects of instruction focused on assignment completion on the homework performance of students with learning disabilities.* A single-subject experiment design was used to determine how instruction in a comprehensive, independent assignment completion strategy affected the quality of homework and the homework completion rate of eight students with learning disabilities.[4]

Single-subject experimental research is described in detail Chapter 12.

Qualitative Approaches

Qualitative research seeks to probe deeply into the research setting to obtain in-depth understandings about the way things are, why they are that way, and how the participants in the context perceive them. To achieve the detailed understandings they seek, qualitative researchers

[3] "Development of Eye Movement Games for Students with Low Vision: Single-Subject Design Research," by M. Donmez and K. Cagiltay, 2018, *Education and Information Technologies, 24*, pp. 295–305.

[4] "Effects of Instruction in an Assignment Completion Strategy on the Homework Performance of Students with Learning Disabilities in General Education Classes," by C. A. Hughes, K. L. Ruhl, J. B. Schumaker, and D. D. Deshler, 2002, *Learning Disabilities Research and Practice, 17*(1), pp. 1–18.

must undertake sustained in-depth, in-context research that allows them to uncover subtle, less overt, personal understandings. The field of qualitative research uses a variety of common qualitative research designs. For example, some qualitative researchers focus on the exploration of phenomena that occur within a bounded system (e.g., a person, event, program, life cycle; in a *case study*); some focus in depth on a group's cultural patterns and perspectives to understand participants' behavior and their context (i.e., using *ethnography*); some examine how multiple cultures compare to one another (i.e., *ethology*); some examine people's understanding of their daily activities (i.e., *ethnomethodology*); some derive theory using multiple steps of data collection and interpretation that link actions of participants to general social science theories or work inductively to arrive at a theory that explains a particular phenomenon (i.e., *grounded theory*); some ask about the meaning of this experience for these participants (i.e., *phenomenology*); some look for common understandings that have emerged to give meaning to participants' interactions (i.e., *symbolic interaction*); some seek to understand the past by studying documents, relics, and interviews (i.e., *historical research*); and some describe the lives of individuals (i.e., *narrative*). Overall, a collective, generic name for these qualitative approaches is *interpretive research*.[5]

NARRATIVE RESEARCH **Narrative research** is the study of how different humans experience the world around them; it involves a methodology that allows people to tell the stories of their "storied lives."[6] The researcher typically focuses on a single person and gathers data by collecting stories about the person's life. The researcher and participant then construct a written account, known as a *narrative*, about the individual's experiences and the meanings the individual attributes to the experiences. Because of the collaborative nature of narrative research, it is important for the researcher and participant to establish a trusting and respectful relationship. Another way to think of narrative research is that the narrative is the story of the phenomenon being investigated, and narrative is also the method of inquiry being used by the researcher.[7] One of the goals of narrative research in education is to increase understanding of central issues related to teaching and learning through the telling and retelling of teachers' stories.

Following is an example of the narrative research approach:

> Qualeya, an assistant professor of education, is frustrated by what she perceives as the gender-biased distribution of resources within the School of Education (SOE). Qualeya shares her story with Winston, a colleague and researcher. In the course of their lengthy recorded conversations, Qualeya describes in great detail her view that the SOE dean, George, is allocating more resources for technology upgrades, curriculum materials, and conference travel to her male colleagues. Qualeya also shares with Winston her detailed journals, which capture her experiences with George and other faculty members in interactions dealing with the allocation of resources. In addition, Winston collects artifacts—including minutes of faculty meetings, technology orders, and lists of curriculum materials ordered for the library at the university—that relate to resource allocation.
>
> After collecting all the data that will influence the story, Winston reviews the information, identifies important elements and themes, and retells Qualeya's story in a narrative form. After constructing the story with attention given to time, place, plot, and scene, he shares the story with Qualeya, who collaborates on establishing its accuracy. In his interpretation of Qualeya's unique story of gender bias, Winston describes themes related to power and influence in a hierarchical school of education and the struggles faced by beginning professors to establish their career paths in a culture that is remarkably resistant to change.

Narrative research is described in detail in Chapter 13.

ETHNOGRAPHIC RESEARCH **Ethnographic research**, or ethnography, is the study of the cultural patterns and perspectives of participants in their natural settings. Ethnography focuses on a particular site or sites that provide the researcher with a context in which to study both the setting and the participants who inhabit it. An ethnographic setting can be defined as anything from a bowling alley to a neighborhood, from a nomadic group's traveling range to an elementary principal's office. The participants are observed as they take part in naturally occurring activities within the setting.

The ethnographic researcher avoids making interpretations and drawing conclusions too early in the study. Instead, the researcher enters the setting slowly, learning how to become accepted by the participants and gaining rapport with them. Then, over time, the researcher collects data in waves, making initial observations and interpretations about the context and participants, then collecting and examining more data in a second wave of refining the initial interpretation, then collecting another wave of data to further refine observations and interpretation, and so on, until the researcher has obtained a deep understanding

[5]For a discussion, see *Qualitative Evaluation and Research Methods* (4th ed.), by M. Q. Patton, 2015, Thousand Oaks, CA: Sage.

[6]"Stories of Experience and Narrative Inquiry," by F. M. Connelly and D. J. Clandinin, 1990, *Educational Research, 19*(5), p. 2.

[7]"Stories," Connelly and Clandinin, pp. 2–14.

of both the context and its participants' roles in it. Lengthy engagement in the setting is a key facet of ethnographic research. The researcher organizes the data and undertakes a cultural interpretation. The result of the ethnographic study is a holistic description and cultural interpretation that represents the participants' everyday activities, values, and events. The study is written and presented as a narrative, which, like the study from which it was produced, may also be referred to as an ethnography.

Following is an example of an ethnographic approach to a research question:

What are the cultural experiences of first-generation college students in an urban community college? After selecting a general research question and a research site in a community college that enrolls many first-generation students, the researcher first gains entry to the college and establishes rapport with the participants of the study. Building rapport can be a lengthy process, depending on the characteristics of the researcher (e.g., the researcher's socioeconomic history, how they are perceived by participants as an academic). As is common in qualitative approaches, the researcher simultaneously collects and interprets data to help focus the general research question initially posed.

Throughout data collection, the ethnographic researcher identifies recurrent themes, integrates them into existing categories, and adds new categories as new themes or topics arise. The success of the study relies heavily on the researcher's skills in analyzing and synthesizing the qualitative data into coherent and meaningful descriptions. The research report includes a holistic description of the culture, the common understandings and beliefs shared by participants, a discussion of how these beliefs relate to life in the culture, and discussion of how the findings compare to literature already published about similar groups. In a sense, the successful researcher provides guidelines that enable someone not in the culture to know how to think and behave in the culture.

Ethnographic research is described in detail in Chapter 14.

CASE STUDY RESEARCH **Case study research** is a qualitative research approach to conducting research on a unit of study or bounded system (e.g., an individual teacher, a classroom, or a school can be a case). Case study research is an all-encompassing method covering design, data collection techniques, and specific approaches to data analysis.[8] A case study is also the name for the product of case study research, which is different from other field-oriented research approaches such as narrative research and ethnographic research.

Following is an example of a study that used the case study research approach:

> Mills (1988)[9] asked, "How do central office personnel, principals, and teachers manage and cope with multiple innovations?" and studied educational change in one American school district. Mills described and analyzed how change functioned and what functions it served in this district. The function of change was viewed from the perspectives of central office personnel (e.g., superintendent, director of research and evaluation, program coordinators), principals, and teachers as they coped with and managed multiple innovations, including the introduction of kindergartens to elementary schools, the continuation of a program for students considered at-risk of academic failure, and the use of the California Achievement Test (CAT) scores to drive school improvement efforts. Mills used qualitative data collection techniques including participant observation, interviewing, written sources of data, and nonwritten sources of data.

Case study research is described in detail in Chapter 14.

Classification of Research by Purpose

Learning Outcome 1.4 **Explain the purposes of basic research, applied research, evaluation research, research and development (R&D), and action research.**

Research designs can also be classified by the degree of direct applicability of the research to educational practice or settings. When the purpose is the classification criterion, all research studies fall into one of two categories: basic research and applied research. Applied research can be subdivided into evaluation research, research and development (R&D), and action research.

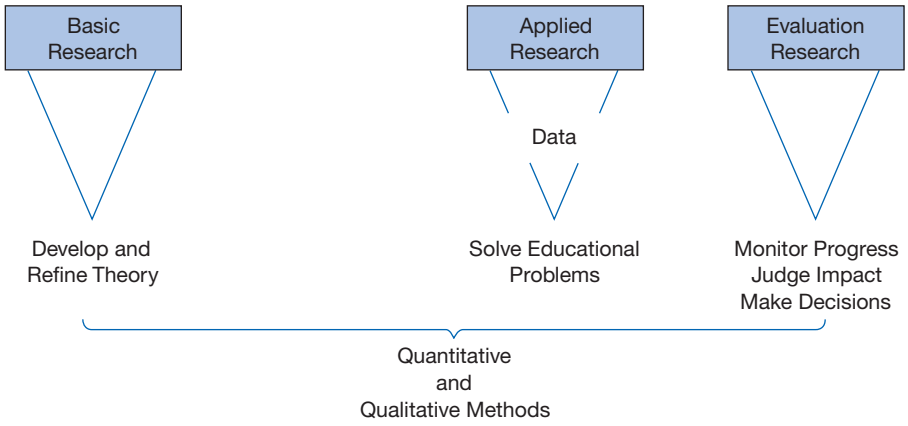
Basic and Applied Research

It is difficult to discuss basic and applied research separately because they are on a single continuum. In its purest form, basic research is conducted solely for the purpose of developing or refining a theory. Theory development is a conceptual process that requires many research studies conducted over time. Basic researchers may not be concerned with the immediate utility of their findings because it may be years before basic research leads to a practical educational application. For example, one of the articles listed at the end of this chapter focuses on basic research to develop and refine theories of children's adaptation to new school settings (Hamre & Pianta, 2005).

[8] *Case Study Research: Design and Methods* (5th ed.) by R. K. Yin, 2014, Thousand Oaks, CA: Sage.

[9] *Managing and Coping with Multiple Educational Changes: A Case Study* by G. E. Mills, 1988, unpublished doctoral dissertation, University of Oregon, Eugene.

Figure 1.2 The educational research continuum



Applied research, as the name implies, is conducted for the purpose of applying or testing a theory to determine its usefulness in solving practical problems. A teacher who asks, "Will the theory of multiple intelligences help improve my students' learning?" is seeking an answer to a practical classroom question. This teacher is not interested in building a new theory or even generalizing beyond her classroom; instead, she is seeking specific helpful information about the impact of a promising practice (i.e., a teaching strategy based on the theory of multiple intelligences) on student learning. For example, one of the articles listed at the end of this chapter focuses on how a beginning teacher integrates university coursework on multicultural education into her classroom teaching and the decision-making process related to the implementation of a multicultural curriculum (Sleeter, 2009).

Educators and researchers sometimes disagree about which end of the basic–applied research continuum should be emphasized. Many educational research studies are located on the applied end of the continuum; they are more focused on what works best than on finding out why it works as it does. However, both basic research and applied research are necessary. Basic research provides the theory that produces the concepts for solving educational problems. Applied research provides data that can help support, guide, and revise the development of theory. Studies located in the middle of the basic–applied continuum seek to integrate both purposes. Figure 1.2 illustrates the educational research continuum.

Evaluation Research

At the applied end of the research continuum is evaluation research, an important, widely used, and explicitly practical form of research. **Evaluation research** is the systematic process of collecting and analyzing data about the quality, effectiveness, merit, or value of programs, products, or practices. Unlike other forms of research that seek new knowledge or understanding, evaluation research focuses mainly on making decisions—decisions about those programs, products, and practices. For example, following evaluation, administrators may decide to continue a program or to abandon it, to adopt a new curriculum or to keep the current one. Some typical evaluation research questions are "Is this special science program worth its costs?" "Is the new reading curriculum better than the old one?" "Did students reach the objectives of the diversity sensitivity program?" and "Is the new geography curriculum meeting the teachers' needs?"

Evaluations come in various forms and serve different functions.[10] An evaluation may be either formative or summative, for example. **Formative evaluation** occurs during the design phase when a program or product is under development and is conducted during implementation so that weaknesses can be remedied. **Summative evaluation** focuses on the overall quality or worth of a completed program or product.

Research and Development (R&D)

Research and development (R&D) is the process of researching consumer needs and then developing products to fulfill those needs. The purpose of R&D efforts in education is not to formulate or test theory but to develop effective products for use in schools. Such products include teacher-training materials, learning materials, sets of behavioral objectives, media materials, and management systems. R&D efforts are generally quite extensive

[10]See *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*, by D. Stufflebeam, G. Madaus, and T. Kellaghan, 2000, Norwell, MA: Kluwer Academic; *The Program Evaluation Standards: A Guide for Evaluators and Evaluation Users* (3rd ed.), by Yarbrough, D. B., Shulha, L. M., Hopson, R. K., and Caruthers, F. A., by Joint Committee on Standards for Educational Evaluation, 2011, Thousand Oaks, CA: Sage.

in terms of objectives, personnel, and time to completion. Products are developed according to detailed specifications. Once completed, products are field-tested and revised until a prespecified level of effectiveness is achieved. Although the R&D cycle is expensive, it results in quality products designed to meet specific educational needs. School personnel who are the consumers of R&D endeavors may, for the first time, really see the value of educational research.

Action Research

Action research in education is any systematic inquiry conducted by teachers, principals, school counselors, or other stakeholders in the teaching–learning environment to gather information about the ways in which their particular schools operate, the teachers teach, and the students learn. Its purpose is to provide teacher-researchers with a method for solving everyday problems in their own settings. Because the research is not characterized by the same kind of control evident in other categories of research, however, study results cannot be applied to other settings. The primary goal of action research is the solution of a given problem, not contribution to science. Whether the research is conducted in one classroom or in many classrooms, the teacher is very much a part of the process. The more research training the teachers have had, the more likely it is that the research will produce valid results. Action research can use quantitative, qualitative, or mixed methods research designs depending on the nature of the research problem.

Following are examples of action research:

- *A study to determine how mathematics problem-solving strategies are integrated into student learning and transferred to real-life settings outside the classroom.* An elementary teacher conducts the study in his or her own school.
- *A study on how a school grading policy change affects student learning.* A team of high school teachers works collaboratively to determine how replacing number and letter grades with narrative feedback affects student learning and attitudes toward learning.

The value of action research is confined primarily to those conducting it. Despite this limitation, action research represents a scientific approach to problem solving that is considerably better than change based on the alleged effectiveness of untried procedures and infinitely better than no change at all. It is a means by which concerned school personnel can attempt to improve the educational process, at least within their environment.

Action research is described in detail in Chapter 17.

This chapter has provided a general introduction to fundamental aspects of the scientific method, along with examples of both quantitative and qualitative approaches. Included are overviews of educational research methods and research purposes. If the number of new terms and definitions seems overwhelming, remember that most will be revisited and reviewed in succeeding chapters. In those chapters we present more specific and detailed features needed to carry out, understand, and conduct useful educational research.

Write Like a Researcher!

Sometimes educational researchers are criticized for being incredibly boring, especially in the way they write about their research (not engaging like this textbook!). After all, while the content of our research is perhaps far more "black and white" than E. L. James's *Fifty Shades of Grey*, perhaps it is still possible to write in a style that is consistent with meeting the requirements of the scientific method while still being engaging for our readers. After all, how many of us can claim that we have recently completed a real "page turner" of a research paper?! Like some of the pages of this text, I suspect that we ultimately find ourselves trapped with the writing conventions of the academy, but I encourage you not to give up the goal of writing about your research in an engaging fashion.

If we look at the characteristics of the scientific method presented in this chapter, we can start to think about writing an initial research narrative that includes the following four sections and compare what we may write with what the authors of the research articles at the end of this chapter have written:

1. *Selection and definition of a problem.* A problem is a question of interest that can be tested or answered through the collection and analysis of data.
    - "The current study was designed to extend work related to school effects by following children identified in kindergarten as being at risk of school failure and examining whether the classroom environment to which they were exposed during the first grade moderated these risks by the end of the first grade" (Hamre & Pianta).
    - "To examine the relationship between a teacher's learning and my teaching strategies in the university" (Sleeter)

2. *Execution of research procedures.* The procedures reflect all the activities involved in collecting data related to the problem (e.g., how data are collected and from whom).

    - "Children included in this study took part in the National Institute of Child Health and Human Development (NICHD) Study of Early Child Care. The children's mothers were recruited from hospitals located in or near Little Rock, AK; Irvine, CA; Lawrence, KS; Boston, MA; Philadelphia, PA; Pittsburgh, PA; Charlottesville, VA; Morganton, NC; Seattle, WA; and Madison, WI. In 1991, research staff visited 8,986 mothers giving birth in these hospitals. Of these mothers, 5,416 met eligibility criteria and agreed to be contacted after returning home from the hospital. A randomly selected subgroup (with procedures to ensure economic, educational, and ethnic diversity) were contacted and enrolled in the study. This resulted in a sample of 1,364 families with healthy newborns. Details of this selection procedure are published in the study manuals (NICHD ECCRN, 1993)" (Hamre & Pianta).
    - "Case study research typically uses a variety of methods to collect data, with an objective toward triangulating findings across methods (Creswell, 2008; Stake, 2000). Data for this study included (1) several papers that Ann completed during the course, including a unit that she designed as a course requirement; (2) a journal that I kept after each class session; (3) notes on two observations of Ann teaching the unit that she designed after the course had ended; and (4) a 40-minute recorded interview with Ann following my observations" (Sleeter).

3. *Analysis of data.* Data are analyzed in a way that permits the researcher to test the research hypothesis or answer the research question.

    - "In order to establish whether instructional and emotional support in the first grade may moderate risk, we first had to establish two preconditions: (1) the existence of a natural experiment, in which children with varying risk backgrounds in kindergarten would sort into first-grade classrooms offering different levels of emotional and instructional support, and (2) whether the hypothesized risk factors were associated with poorer outcomes in first grade. The first precondition was assessed through examining the distribution of children in each risk group into classrooms offering high, moderate, and low support. The second precondition was assessed by conducting ANCOVAs in which risk status was used to predict first-grade outcomes, after adjusting for children's previous performance on these outcomes measures" (Hamre & Pianta).
    - "Early in the semester (September), I guided teachers in analyzing epistemological assumptions in various documents related to curriculum, such as curriculum standards and school reform proposals available on the Internet. Teachers examined documents in relationship to questions such as the following: *Who produced this document (if it is possible to tell)? How is it intended to be used? By whom? What is its purpose? Whose view of the world does it support? Whose view does it undermine or ignore? Whose knowledge isn't here?* In addition, they analyzed textbooks from their classrooms, with guidance from a textbook analysis instrument" (Sleeter).

4. *Drawing and stating conclusions.* The conclusions, which should advance our general knowledge of the topic in question, are based on the results of data analysis.

    - "The study provides evidence that across two important domains of child functioning in the early grades of school, achievement, and relationships with teachers, the quality of everyday classroom interactions in the form of instructional and emotional support moderates the risk of early school failure" (Hamre & Pianta).
    - "This case study showed how one novice teacher began to question institutionalized assumptions in the context of a graduate course and how she began to think more complexly. The case study reinforced for me the importance of creating contexts in which teachers can examine their own backgrounds and beliefs, interact with one another, and interact with ideas that stretch them intellectually. Of course, no two teachers bring the same prior experiences, beliefs, and commitments. The challenge for an instructor lies in planning a course that activates a variety of experiences and enables uncomfortable questions and disagreements to take place so that teachers can grow. This inquiry into learning has helped me make sense of that challenge" (Sleeter).

Are there characteristics of each of these papers that you find engaging? If so, what are they and how might they find their way into your own writing about your research? Are you up to the challenge? Chapter 22 contains considerable details about guidelines for writing a research report, and formats and styles for theses, dissertations, and journal articles.

Summary

The Scientific Method

1. The goal of all scientific endeavors is to describe, explain, predict, and/or control phenomena.
2. Compared to other sources of knowledge, such as experience, authority, inductive reasoning, and deductive reasoning, application of the scientific method is the most efficient and reliable.
3. The scientific method is an orderly process that entails recognition and definition of a problem, formulation of hypotheses, collection and analysis of data, and statement of conclusions regarding confirmation or disconfirmation of the hypotheses.

Limitations of the Scientific Method

4. Four main factors put limitations on the use of a scientific and disciplined inquiry approach: inability to answer some types of questions, inability to capture the full richness of the research site and the complexity of the participants, limitations of measuring instruments, and the need to address participants' needs in ethical and responsible ways.

Application of the Scientific Method in Education

5. Research is the formal, systematic application of the scientific method to the study of problems; educational research is the formal, systematic application of the scientific method to the study of educational problems.
6. The major difference between educational research and some other types of scientific research is the nature of the phenomena studied. It can be quite difficult to explain, predict, and control situations involving human beings, by far the most complex of all organisms.
7. The research process usually comprises four general steps:
   a. Selection and definition of a problem
   b. Execution of research procedures
   c. Analysis of data
   d. Drawing and stating conclusions

Different Approaches to Educational Research

The Continuum of Research Philosophies

8. Certain philosophical assumptions underpin an educational researcher's decision to conduct research. These philosophical assumptions address issues related to the nature of reality (ontology), how researchers know what they know (epistemology), and the methods used to study a particular phenomenon (methodology).

Quantitative Research

9. Quantitative research is the collection and analysis of numerical data to explain, predict, and/or control phenomena of interest.
10. Key features of quantitative research are hypotheses that predict the results of the research before the study begins; control of contextual factors that may influence the study; collection of data from sufficient samples of participants; and use of numerical, statistical approaches to analyze the collected data.
11. The quantitative approach assumes that the world is relatively stable, uniform, and coherent.

Qualitative Research

12. Qualitative research is the collection, analysis, and interpretation of comprehensive narrative and visual (nonnumeric) data to gain insights into a particular phenomenon of interest.
13. Key features of qualitative research include defining the problem, but not necessarily at the start of the study; studying contextual factors in the participants' natural settings; collecting data from a small number of purposely selected participants; and using nonnumeric, interpretive approaches to provide narrative descriptions of the participants and their contexts.
14. An important belief that underlies qualitative research is that the world is not stable, coherent, or uniform, and therefore there are many truths.

Mixed Methods Research

15. Mixed methods research combines quantitative and qualitative approaches by including both quantitative and qualitative data in a single study. The purpose of mixed methods research is to build on the synergy and strength that exists between quantitative and qualitative research approaches to understand a phenomenon more fully than is possible using either quantitative or qualitative approaches alone.

Characteristics of Quantitative and Qualitative Research Approaches

16. Qualitative and quantitative researchers follow the same basic six steps in conducting research, although application of the steps may differ depending on the research design.

Classification of Research by Design

Quantitative Approaches

17. Quantitative research approaches are intended to describe current conditions, investigate relations, and study cause–effect phenomena.

18. Survey research involves collecting numerical data to answer questions about the current status of the subject of study.

19. Correlational research examines the relation between two or more variables. A variable is a placeholder—such as age, IQ, or height—that can take on different values.

20. In correlational research, the degree of relation is measured by a correlation coefficient. If two variables are highly related, one is not necessarily the cause of the other.

21. Causal–comparative research seeks to investigate differences between two or more different programs, methods, or groups. The activity thought to make a difference (e.g., the program, method, or group) is called the grouping variable. The effect is called the dependent variable.

22. In most causal–comparative research studies, the researcher does not have control over the grouping variable because it already has occurred or cannot be manipulated. Causal–comparative research is useful in those circumstances when it is impossible or unethical to manipulate an independent variable.

23. True experimental research investigates causal relations among variables.

24. The experimental researcher controls the selection of participants by choosing them from a single pool and assigning them at random to different causal treatments. The researcher also controls contextual variables that may interfere with the study. Because participants are randomly selected and assigned to different treatments, experimental research permits researchers to make true cause–effect statements.

25. Single-subject experimental designs are a type of experimental research that can be applied when the sample is one individual or group. This type of design is often used to study the behavior change that an individual or group exhibits as a result of some intervention or treatment.

Qualitative Approaches

26. Qualitative approaches include narrative research, ethnographic research, and case study research. The focus of these methods is on deep description of aspects of people's everyday perspectives and context.

27. Narrative research is the study of how individuals experience the world. The researcher typically focuses on a single person and gathers data through the collection of stories.

28. Ethnographic research is the study of the cultural patterns and perspectives of participants in their natural setting. Ethnography focuses on a particular site or sites that provide the researcher with a context in which to study both the setting and the participants who inhabit it.

29. Case study research is a qualitative research approach to conducting research on a unit of study or bounded system (e.g., classroom, school).

Classification of Research by Purpose

Basic and Applied Research

30. Basic research is conducted to develop or refine theory, not to solve immediate practical problems. Applied research is conducted to find solutions to current practical problems.

Evaluation Research

31. The purpose of evaluation research is to inform decision making about educational programs and practices.

Research & Development

32. The major purpose of R&D efforts is to develop effective products for use in schools.

Action Research

33. The purpose of action research is to provide teacher-researchers with a method for solving everyday problems in their own settings.

Performance Criteria Task 1

Tasks 1A and 1B

Reprints of two published research reports appear on the following pages (Task 1A Quantitative Example and Task 1B Qualitative Example). Read the reports and then state the following for each study:

- Research design
- Rationale for the choice of the design
- Major characteristics of the design including: research problems, methods of data analysis, and major conclusions

One sentence should be sufficient to describe the research design and the rationale for the choice of the design. For the major characteristics of the design, one or two sentences will usually be sufficient to state the research problems and the method of data analysis. You are expected only to identify the analysis, not explain it. The major conclusion that you identify and state (one or two sentences should be sufficient) should directly relate to the original topic. Statements such as "more research is needed in this area" do not represent major conclusions.

Suggested responses to these tasks appear in Appendix B of this text. If your responses differ greatly from those suggested, study the reports again.

Task 1C

Brief descriptions of five research studies follow these instructions. Read each description and decide whether the study represents a survey, correlational, causal–comparative, experimental, single-subject, narrative, ethnographic, or case study approach. State the research approach for each topic statement, and indicate why you selected that approach. Your reasons should be related to characteristics that are unique to the type of research you have selected.

1. In this study, researchers administered a questionnaire to determine how social studies teachers felt about teaching world history to fifth-graders.
2. This study was conducted to determine whether the Acme Interest Test provided similar results to the Acne Interest Test.
3. This study compared the achievement in reading of fifth-graders from single-parent families and those from two-parent families.
4. This study divided fifth-grade students in a school into two groups at random and compared the results of two methods of conflict resolution on students' aggressive behavior.
5. This study examined the culture of recent Armenian emigrants in their new setting.

Suggested responses appear in Appendix B.

Task 1A Quantitative Example

Can Instructional and Emotional Support in the First-Grade Classroom Make a Difference for Children at Risk of School Failure?

BRIDGET K. HAMRE
University of Virginia

ROBERT C. PIANTA
University of Virginia

ABSTRACT This study examined ways in which children's risk of school failure may be moderated by support from teachers. Participants were 910 children in a national prospective study. Children were identified as at risk at ages 5–6 years on the basis of demographic characteristics and the display of multiple functional (behavioral, attention, academic, social) problems reported by their kindergarten teachers. By the end of first grade, at-risk students placed in first-grade classrooms offering strong instructional and emotional support had achievement scores and student–teacher relationships commensurate with their low-risk peers; at-risk students placed in less supportive classrooms had lower achievement and more conflict with teachers. These findings have implications for understanding the role that classroom experience may play in pathways to positive adaptation.

Identifying the conditions under which experiences in school settings can alter the early trajectories of children's social or academic functioning has important implications for understanding pathways to children's positive adaptation. Of particular interest is whether experiences in high-quality classrooms can help close the gap between children at risk of school failure and their low-risk peers, particularly in the early grades when small increments in achievement play a large role in eventual outcomes (Alexander, Entwisle, & Kabbani, 2001; Ferguson, 1998; Phillips, Crouse, & Ralph, 1998; Ross, Smith, Slavin, & Madden, 1997). Two bodies of work are relevant to this question. The first examines everyday classroom interactions between teachers and children that predict more positive development for all children (Brophy & Good, 1986; Gage & Needel, 1989; Howes et al., 2005; NICHD ECCRN, 2003; Pianta, LaParo, Payne, Cox, & Bradley, 2002; Rimm-Kaufman, LaParo, Pianta, & Downer, in press; Ritchie & Howes, 2003; Skinner & Belmont, 1993; Stipek et al., 1998). The second area of research provides evidence of specific school-based interventions that may alter trajectories for students with various risk factors (Battistich, Schaps, Watson, & Solomon, 1996; Durlak & Wells, 1997; Elias, Gara, Schuyler, Branden-Muller, & Sayette, 1991; Greenberg et al., 2003; Weissberg & Greenberg, 1998; Wilson, Gottfredson, & Najaka, 2001). At the intersection of these areas of education and developmental science is the question of whether students' everyday instructional and social interactions with

(01)

Application of the scientific method: selection and definition of a problem—whether experiences in high-quality classrooms can help close the gap between children at risk of school failure and their low-risk peers in the early grades.

The work reported herein was supported in part by the National Institute of Child Health and Human Development (NICHD) Study of Early Child Care (U10-HD25449), NICHD R21-43750 and by American Psychological Association/Institute of Education Sciences Postdoctoral Education Research Training fellowship under the Department of Education, Institute of Education Sciences grant number R305U030004.

Correspondence concerning this article should be addressed to Bridget K. Hamre, University of Virginia, PO Box 800784, Charlottesville, VA 22908-0784. Electronic mail may be sent to bkh3d@virginia.edu.

0009-3920/2005/7605-0001]

teachers in the classroom may themselves ameliorate the risk of school failure. If this were the case, focused efforts related to teacher training and support, curriculum implementation, and assessments of classroom settings could be used more strategically to counter the tendency toward poor outcomes for such children (see Pianta, in press, for a discussion). The current study used data from a large, national prospective study of children and families to examine ways in which risk of school failure may be moderated by strong support from teachers in the first-grade classroom. Specifically, we examined whether children at risk of early school failure experiencing high levels of instructional and emotional support in the first grade displayed higher achievement and lower levels of student–teacher conflict than did their at-risk peers who did not receive this support.

Everyday Classroom Interactions and Student Outcomes

(02) Research on everyday classroom processes that may alter trajectories for students at risk has its foundations in the process–product research from the 1960s to 1980s that focused attention on observable teacher behaviors (Brophy & Good, 1986; Gage & Needel, 1989) and in developmentally informed theories of schooling that focus attention on socio-emotional, motivational (Connell & Wellborn, 1991; Deci & Ryan, 1985; Eccles, 1993; Wentzel, 2002) and instructional (e.g., Resnick, 1994; Stevenson & Lee, 1990) experiences in classrooms that trigger growth and change in competence. Although it posited the type of interactions between student characteristics and teacher behaviors that are now beginning to be reported in the literature (e.g., Morrison & Connor, 2002; Rimm-Kaufman et al., 2002) and has resulted in frameworks for describing classroom processes that inform educational research (e.g., Brophy, 2004), the process–product research tradition did not yield a body of empirical findings that provide a strong case for classroom effects, particularly in relation to issues such as moderation of child characteristics. Reviews of the contribution of this literature in large part note the lack of grounding in developmental and psychological research as well as the complex and interactive nature of students' classroom experiences (Gage & Needel, 1989; Good & Weinstein, 1986). Within developmental psychology, the focus on proximal processes in ecological models (Bronfenbrenner & Morris, 1998; Lerner, 1998; Sameroff, 1995, 2000) and the extension of these perspectives to school settings (Connell & Wellborn, 1991; Pianta, 1999; Resnick, 1994; Stevenson & Lee, 1990) have advanced efforts to understand the interactive processes through which children and adolescents experience, the classroom environment (Pianta, in press). Roeser, Eccles, and Sameroff (2000) extend the linkage between developmental studies and education, even further when arguing, with respect to understanding middle school effects, for research "linking the study of adolescents' experience, motivation, and behavior in school with the study of their teachers' experience, motivation, and behavior at school" (p. 466). This explicit need to focus on the interaction of child characteristics with types or categories of resources available in classroom (and school) settings is consistent with Rutter and Maughan's (2002) analysis of shortcomings in the school-effects literature. However, if such an approach is to yield more fruitful results than the process–product work, it is in large part predicated on more sophisticated understandings of the developmental needs of children vis-à-vis experiences in school (e.g., Reid, Patterson, & Snyder, 2002) and parallel efforts to understand and measure developmentally relevant assets in school environments (see Morrison & Connor, 2002; Rimm-Kaufman et al., 2002, as recent examples).

(03) One avenue for advancing the understanding of schooling as a moderator of child (or background) characteristics is the assessment of variation in the nature, quality, and quantity of teachers' interactions with students (e.g., Burchinal et al., 2005). Recent large-scale observational studies indicate that these types of interaction within classrooms are highly variable (e.g., National Institute of Child Health and Human

Development, Early Child Care Research Network) [NICHD ECCRN], 2002b, in press). Even the most well-described, manualized, standardized, scientifically based classroom intervention programs are enacted in practice in ways that vary widely from child to child or classroom to classroom (e.g., Greenberg, Doitrovich, & Bumbarger, 2001). In descriptions of less-tightly prescribed classroom interactions, the degree to which classroom teachers make productive use of time or classrooms are well-managed ranges across the full spectrum of possibilities, even though kindergartens and first-grade classes appear, on average, to be positive and supportive social settings (NICHD ECCRN, 2002b, in press; Pianta et al., 2002).

(04) In recent large-scale observational studies of pre-k to elementary classrooms, two dimensions consistently emerge: instructional support and emotional support (NICHD ECCRN, 2002b, in press; Pianta et al., 2002; Pianta, LaParo, & Hamre, 2005). Interestingly, these two dimensions, to some extent, predict differentially children's social and academic outcomes, confirming theoretical views that various developmental needs of children may interact differentially with the qualities of school settings (Connell & Wellborn, 1991; Morrison & Connor, 2002; Rutter & Maughan, 2002). For example, when evaluated in the same prediction model, instructional support for learning predicts achievement outcomes to a significantly greater degree than emotional support predicts these same outcomes (Howes et al., 2005). On the other hand, children's anxious behavior reported by mothers (but not academic performance) is predicted by the degree of classroom structure and instructional press in the first grade (NICHD ECCRN, 2003), while higher levels of emotional support predict a very broad range of social and task-oriented competencies such as following directions (Howes et al., 2005). Morrison and Connor (2002) argue that the effects of schooling on development have to be modeled at the level of specific forms of input and resource that are matched to specific child needs, abilities, and skills. Thus, according to Morrison and Connor (2002), it is not only necessary to conceptualize and measure the classroom setting (or school) in terms of specific aspects of instructional or social environment, but also to gauge the effects of those experiences relative to how well they match the child's capacities and skill. In this view, school effects are predominantly in the form of interactions between specific inputs from the classroom and the characteristics of the child.

(05) These two broad dimensions of everyday teacher–student classroom interactions—emotional and instructional support—with theoretical and empirical links to student development, can be a starting point for examining interactions with child and background characteristics, particularly attributes that place children at risk for school failure. In global observations reported in the literature, emotional support encompasses the classroom warmth, negativity, child-centeredness as well as teachers' sensitivity and responsivity toward specific children (NICHD ECCRN, 2002b, in press). This should not be surprising as a number of developmentally informed theories suggests that positive and responsive interactions with adults (parents, teachers, child-care providers) contribute to regulation of emotional experience and social behavior, the development of skills in social interactions, and emotional understanding (Birch & Ladd, 1998; Connell & Wellborn, 1991; Eccles, 1993; Howes, 2000; Howes, Matheson, & Hamilton, 1994; Pianta, 1999; Wentzel, 2002). Confirming this perspective are results indicating that exposure to positive classroom climates and sensitive teachers is linked to greater self-regulation among elementary and middle school students (Skinner, Zimmer-Gembeck, & Connell, 1998), greater teacher-rated social competence (Burchinal et al., 2005; Howes, 2000; Pianta et al., 2002), and decreases in mother-reported internalizing problems from 54 months to the end of the first grade (NICHD ECCRN, 2003).

(06) From a somewhat different theoretical perspective, teachers' emotional support directly provides students with experiences that foster motivational and learning-related processes important to academic functioning (Crosnoe, Johnson, & Elder,

2004; Greenberg et al., 2003; Gregory & Weinstein, 2004; Pianta et al., 2002; Rimm-Kaufman et al., in press; Roeser et al., 2000; Zins, Bloodworth, Weissberg, & Walberg, 2004). Theories of motivation suggest that students who experience sensitive, responsive, and positive interactions with teachers perceive them as more supportive and are more motivated within the academic contexts of schooling (Connell & Wellborn, 1991; Deci & Ryan, 1985; Eccles, 1993). In the early grades, Pianta et al. (2002) found that when teachers offered a more child-centered climate, kindergarten children were observed to be more often on-task and engaged in learning. Among older students, perceptions of positive relatedness to teachers predict gains in student engagement over the course of the school year (Furrer & Skinner, 2003), increased motivation to learn (Roeser et al., 2000), and greater academic achievement (Crosnoe et al., 2004; Gregory & Weinstein, 2004). Consistent with this link between motivation and support from adults, teacher support was related to sixth graders' school and class-related interests and pursuit of social goals (Wentzel, 2002), which in turn predicted pursuit of social goals and grades in the seventh grade. For children at risk of problems in school, Noam and Herman's (2002) school-based prevention approach emphasizes the primary importance of relationships with a school-based mentor (Noam, Warner, & Van Dyken, 2001), based explicitly on the rationale that such relationships function as resources and resilience mechanisms in counteracting the effects of risk mechanisms attributable to problems in family relationships.

(07) Notwithstanding the importance of relationships and social support, the nature and quality of instruction is of paramount importance for the value of classroom experience that is intended to produce gains in learning; in elementary school, instruction is under great scrutiny as a result of standards and performance evaluations (Pianta, in press). Although the apparent dichotomy between child-centered and direct instruction has for some years dominated discussions of learning in the early grades (see Stipek et al., 1998), there is accumulating evidence that teachers' instructional interactions with children have the greatest value for students' performance when they are focused, direct, intentional, and characterized by feedback loops involving student performance (Dolezal, Welsh, Pressley, & Vincent, 2003; Juel, 1996; Meyer, Wardrop, Hastings, & Linn, 1993; Pianta et al., 2002; Torgesen, 2002). Torgesen (2002) provides an explicit example of this type of instruction applied to the area of reading by suggesting three primary ways in which everyday teaching can contribute to growth in reading skills: the provision of explicit teaching experiences and practice (i.e., phonemic skills, vocabulary); more productive classroom time in which there are more opportunities for teaching and learning; and intensive scaffolding and feedback to students about their progress. The value of intentional, focused interaction and feedback is not limited to reading, but appears to be a key component in other skill domains such as writing (Matsumura, Patthey-Chavez, Valdes, & Garnier, 2002) that may extend to cognition and higher order thinking (Dolezal et al., 2003).

(08) In addition, these instructional inputs are also associated with more positive and fewer negative interactions between students and teachers, and higher levels of attention and task-oriented behavior (NICHD ECCRN, 2002a; Pianta et al., 2002). Yet, as was the case for emotional support in classrooms, large-scale studies document great variation in the frequency and quality of these instructional procedures within early elementary school classrooms (Meyer et al., 1993; NICHD ECCRN, 2002a, in press). For example, within the NICHD Study of Early Child Care sample (NICHD ECCRN, 2002b, in press), teachers provided specific academic instruction in an average of 8% of all observed intervals over the course of a morning-long observation. However, the range was remarkable, with some classrooms providing no explicit instruction and others providing this instruction in almost 70% of observed intervals. This variability provides an opportunity to examine ways in which exposure to these classroom processes may impact student achievement.

(09) Taken together, research on the nature and quality of early schooling experiences provides emerging evidence that classroom environments and teacher behaviors are

associated in a "value-added" sense with student outcomes. Yet, until recently, few researchers have specifically examined the possibility that these everyday processes in elementary school classrooms may help close (or increase) the gap in student achievement observed among students at risk of school failure because of demographic characteristics (low income, minority status) or functional risks such as serious behavioral and emotional problems. Although there is increasing evidence from well-designed and highly controlled studies that school-based interventions that prescribe certain desired teacher–child interactions can succeed in ameliorating some risks (Catalano et al., 2003; Greenberg et al., 2001; Ialongo et al., 1999; Walker, Stiller, Severson, Feil, & Golly, 1998), there is little available evidence on whether features of classrooms and child–teacher interactions such as emotional or instruction support, present in everyday classroom interactions in naturally varying samples, are sufficiently potent to counteract risk for school failure.

Everyday Interactions and Risk for Early School Failure

Recent evidence from developmentally informed studies of naturally occurring variation in classroom environments directly tests the hypothesis that everyday experiences within elementary classrooms may moderate outcomes for children at risk (Peisner-Feinberg et al., 2001). In one such study, Morrison and Connor (2002) demonstrate that children at risk of reading difficulties at the beginning of the first grade (identified on the basis of test scores) benefited from high levels of teacher-directed explicit language instruction—the more teacher-directed, explicit instruction they received, the higher were their word-decoding skills at the end of the first grade. In contrast, teacher-directed explicit instruction made no difference in decoding skills for children with already high skills on this dimension upon school entry. These highly skilled children made the strongest gains in classrooms, with more child-led literacy-related activities. (10)

In another study providing evidence of the moderating effect of teachers' classroom behaviors on outcomes for at-risk children, Rimm-Kaufman et al. (2002) examined whether teacher sensitivity predicted kindergarten children's behavior for groups of socially bold and wary children, with the bold children demonstrating high levels of off-task behavior and negative interactions with peers and teachers. Although there was no relation between teachers' sensitivity and child classroom behavior among the socially wary children, socially bold children who had more sensitive teachers were more self-reliant and displayed fewer negative and off-task behaviors than did bold children with less sensitive teachers. Similarly, two recent studies suggest that student–teacher conflict is a stronger predictor of later problems for children who display significant acting out behaviors than for their peers who do not display these behavior problems (Hamre & Pianta, 2001; Ladd & Burgess, 2001). Taken together, these studies suggest that positive social and instructional experiences within the school setting may help reduce children's risk, while negative interactions between teachers and children may be particularly problematic for those children displaying the highest risk of school failure. In the present study, we follow and extend the work of Morrison and Connor (2002) and Rimm-Kaufman et al. (2002) to examine effects of two dimensions of classroom process (instructional and emotional quality) on moderating the association(s) between two forms of risk for failure in achievement and social adjustment in the first grade. (11)

Defining School-Based Risk

Although conceptualizations of risk vary, two central categories of children's risk for early school failure relate to *demographic* and *functional* risks. Prior to entering school, it is largely family and demographic factors that place children at risk of failure. One of the most robust of these demographic risk indicators is low maternal education (e.g., Christian, Morrison, & Bryant, 1998; Ferguson, Jimerson, & Dalton, 2001; (12)

NICHD ECCRN, 2002a; Peisner-Feinberg et al., 2001; Shonkoff & Phillips, 2000). One reason posited for this is that children of mothers with low levels of education are less likely to be exposed to frequent and rich language and literacy stimulation (Bowman, Donovan, & Burns, 2001; Christian et al., 1998; Hart & Risley, 1995; U.S. Department of Education, 2000) and thus may come to kindergarten with fewer academic skills (Pianta & McCoy, 1997). These early gaps are often maintained throughout children's school careers (Alexander et al., 2001; Entwisle & Hayduk, 1988; Ferguson et al., 2001).

(13) In addition to demographic factors that signal risk, indicators reflecting children's general functioning and adaptation in the classroom as they enter school (behavioral, attention, social, and academic problems) are established predictors of success or failure in the next grade(s). Children identified by their teachers as displaying difficulties in these domains in the early school years are at higher risk of problems throughout their school careers (Alexander et al., 2001; Flanagan et al., 2003; Hamre & Pianta, 2001; Ladd, Buhs, & Troop, 2002; Lewis, Sugai, & Colvin, 1998). Although problems in individual domains of functioning predict future difficulties, research suggests that the accumulation of multiple risks is typically a much stronger indicator of later problems (Gutman, Sameroff, & Cole, 2003; Seifer, Sameroff, Baldwin, & Baldwin, 1992) and therefore our approach to conceptualizing and assessing functional risk will rely on multiple indicators.

Current Study

(14) The current study was designed to extend work related to school effects by following children identified in kindergarten as being at risk of school failure and examining whether the classroom environment to which they were exposed during the first grade moderated these risks by the end of the first grade. Rutter and Maughan (2002) suggest that effectively testing environmental influences on child development requires attending to several methodological issues. First, they suggest using longitudinal data to measure change within individuals. We were interested in assessing achievement and relational functioning in first grade as a function of the support these children received from teachers; therefore, we needed to adjust for previous performance on these outcomes. Ideally, we would adjust for performance at the beginning of the first-grade year; however, because multiple assessments were not available within the first-grade year, we adjusted for earlier performance on the outcomes (completed at either 54 months or kindergarten). Secondly, Rutter and Maughan (2002) suggest using some form of a *natural experiment* that "pulls apart variables that ordinarily go together" (p. 46). Within this study, the classroom process itself served as the natural experiment, in which children with differing risk backgrounds in kindergarten were placed in first-grade classrooms offering varying levels of emotional and instructional support. Their third recommendation suggests quantified measurement of the postulated causal factor; here we use observations of teachers' instructional and emotional support conducted within classrooms, a notable difference from most previous research on classroom effects, which relies on structural features of the classroom or teacher-reported practices. Two of Rutter and Maughan's (2002) last three recommendations, testing for a dose response gradient and controlling for social selection, initial level, and self-perpetuating effects, were also attended to within this study. The last recommendation, explicitly testing the hypothesized mechanism against some competing explanations, was beyond the scope of this study, although the implications of not testing competing explanations are addressed in the discussion.

Selection of a problem—identification of children in kindergarten as being at risk of school failure and examining whether the classroom environment to which they were exposed during first grade moderated these risks by the end of first grade.

Classification of research by approach—"a natural experiment"—or causal–comparative research.

Causal–comparative research "grouping variable" is the classroom process offering varying levels of emotional and instructional support.

(15) Because of an interest in examining both academic and social functioning, we examined two major outcomes—performance on an individually administered, standardized achievement battery, and first-grade teacher ratings of conflict with the student. Although student–teacher conflict could be viewed as a classroom process, when assessed via the teachers' perspective, it is best conceptualized as an outcome derived

in part from the teachers' social or instructional interactions toward the child. Teachers' rating of their relationship with children measure the extent to which students are able to successfully use the teacher as a resource in the classroom. Thus, although teachers' interactions with students are expected to influence relationships in important ways, these relationships are themselves key indicators of school adaptation. This conceptualization of relationships as outcomes was validated by a study showing that kindergarten teachers' perceptions of conflict with students were stronger predictors of behavioral functioning through the eighth grade than were these same teachers' ratings of behavior problems (Hamre & Pianta, 2001).

Dependent variables are "performance on an individually administered standardized achievement battery" and "first-grade teacher ratings of conflict with the student."

Globally, we expected that children in the risk groups would be more likely than (16)
children at low risk to benefit from placement in classrooms offering high levels of support and that placement in high-quality classrooms would help at-risk students catch up to their low-risk peers. More specific hypotheses require a consideration of the mechanisms through which we expect the risk factors to operate. For example, children whose mothers have low levels of education tend to have less exposure to pre-academic experiences within the home (Bowman et al., 2001; U.S. Department of Education, 2000); thus, we expected that these children would benefit academically from high levels of instructional support within the classroom. In contrast, children displaying behavioral and social problems in kindergarten may require higher levels of emotional support to adjust to the demands of the first grade. However, by responding to children's social and emotional needs, teachers may not only help children adapt socially, but may allow these children to more successfully access the instructional aspects of classrooms; thus, we expected that high levels of emotional support would be associated with more positive academic experiences and lower levels of teacher–child conflict for children displaying multiple functional risks in kindergarten.

Method

Participants

Children included in this study took part in the NICHD Study of Early Child Care. (17)
The children's mothers were recruited from hospitals located in or near Little Rock, AK; Irvine, CA; Lawrence, KS; Boston, MA; Philadelphia, PA; Pittsburgh, PA; Charlottesville, VA; Morganton, NC; Seattle, WA; and Madison, WI. In 1991, research staff visited 8,986 mothers giving birth in these hospitals. Of these mothers, 5,416 met eligibility criteria and agreed to be contacted after returning home from the hospital. A randomly selected subgroup (with procedures to ensure economic, educational, and ethnic diversity) were contacted and enrolled in the study. This resulted in a sample of 1,364 families with healthy newborns. Details of this selection procedure are published in the study manuals (NICHD ECCRN, 1993).

Scientific method, Selection of a sample.

Classroom observations were conducted in the children's second year of school, (18)
which for the majority was the first grade. Of the original sample of 1,364 children, 910 had complete data and were included in the current study. Analyses comparing the children included in this investigation with the entire sample indicate selected attrition: among all children who began the study, White children and those with mothers with higher education were more likely to have data collected in the first grade, $\chi^2(3, N = 1{,}364) = 18.14$, $p < .001$ and $\chi^2(3, N = 1{,}364) = 16.75$, $p < .001$, respectively. Among the children in the present study, 49% were female. The majority were White ($n = 723$), followed in frequency by African American ($n = 96$), Hispanic ($n = 50$), and Other ($n = 39$). Maternal education ranged from 7 to 21 years, with a mean of 14.45 years. The income-to-needs ratio, used to measure income relative to the number of household members, was average across the period of study (54 months, kindergarten, and first grade) and ranged from .15 to 33.77, with an average of 3.73. These factors indicate a largely nonpoverty sample, although there was considerable range.

Overview of Data Collection

(19) Children in this study were followed from birth through the first grade. Maternal education and child ethnicity were reported when children were one-month old. Child outcomes and measures of classroom process were collected in the spring of the children's first-grade year. The 827 classrooms were distributed across 747 schools, in 295 public school districts, in 32 states. Earlier assessments, conducted when the children were 54 months and in kindergarten, provided measures of children's risk status as well as a measure of children's prior functioning on the outcomes of interest. Further documentation about all data collection procedures, psychometric properties of measures, and descriptions of how composites were derived are documented in the Manuals of Operations of the NICHD Study of Early Child Care (NICHD ECCRN, 1993).

Risk Indicators

(20) Children in this study were grouped based on their status on *functional* and *demographic* indicators of risk. Functional indicators of risk included measures of children's attention, externalizing behavior, social skills, and academic competence. The last three measures were collected through teacher report when the study children were in kindergarten. Unfortunately, individual child assessments were not conducted when children were in kindergarten. Because of an interest in including a non-teacher-reported risk variable and based on data showing the links between sustained attention and school failure (Gordon, Mettelman, & Irwin, 1994), the attention risk variable used in this investigation was collected during child assessments conducted when children were 54 months old. Students whose mothers had less than a 4-year college degree were placed in the demographic risk group. Information on the measures and procedures used to identify children at risk of school failure is provided below.

Functional Risk

(21) *Sustained attention.* Sustained attention was assessed using a continuous performance task (CPT) based on the young children's version described by Mirsky, Anthony, Duncan, Aheani, and Kellam (1991). This measure consisted of a computer-generated task in which children are asked to push a button each time a target stimulus appears. The number of omission errors was used as the unit of analysis for this study. The CPT has adequate test–retest reliability $(r = .65 - .74)$ and has high content and predictive validity (Halperin, Sharman, Greenblat, & Schwartz, 1991).

(22) *Externalizing behaviors*. Externalizing behaviors were assessed with the teacher report form (TRF; Achenbach, 1991), a widely used measure of problem behaviors that has been standardized on large samples of children. This measure lists 100 problem behaviors and has teachers rate them as not true (0), somewhat true (1), or very true (2) of the student. The externalizing problems standard score was used for these analyses. This scale contains teachers' reports on children's aggressive (e.g., gets in many fights; cruelty, bullying, or meanness to others; physically attacks people), attention (e.g., cannot concentrate; fails to finish things he/she starts), and defiant behaviors (e.g., defiant, talks back to staff; disrupts class discipline). The reliability and validity of the TRF has been widely established (see Bérubé & Achenbach, 2001, for a review).

(23) *Social skills and academic competence*. Students' social skills and academic competence were assessed with the social skills rating system–teacher form (SSRS; Gresham & Elliot, 1990). This measure consists of three scales: social skills, problems behaviors, and academic competence. Because the TRF is a more established measure of problem behaviors, only the social skills and academic competence scales were used in these analyses. The social skills composite asks teachers to rate the frequency of classroom behaviors $(0 = \text{never}, 1 = \text{sometimes}, \text{two} = \text{very often})$ in three areas related

to positive social adjustment in school settings: cooperation (e.g., paying attention to instructions, putting away materials properly), assertion (e.g., starting conversations with peers, helping peers with classroom tasks), and self-control (e.g., responding to peer pressure appropriately, controlling temper). Within this sample, the coefficient α for the social skills composite was .93. The academic competence composite asks teachers to judge children's academic or learning behaviors in the classroom on a 5-point scale that corresponds to the percentage clusters of the students in the class ($1 = \text{lowest}\,10\%$, $5 = \text{highest}\,10\%$). Within this sample, the coefficient α for this scale was .95. Scores are standardized based on norms from a large, national sample of children. The SSRS has sufficient reliability and has been found to correlate with many other measures of adjustment (Gresham & Elliot, 1990).

Functional risk status. Students' risk status was determined for each of these four indicators. Children with standardized scores at least one standard deviation below the mean (85 or lower) on the social skills and academic competence scales were placed in the social risk ($n = 83$; 10%) and academic risk groups ($n = 112$; 13%), respectively. Similarly, children who fell one standard deviation above the mean on the number of omission errors on the CPT were included in the attention risk group ($n = 144$; 17%). Consistent with recommendations in the TRF manual (Achenbach, 1991), children in the externalizing problems risk group had *T* scores at or above 62 on the externalizing problems factor ($n = 80$; 9%). Given previous research indicating that multiple, rather than isolated, risks are most predictive of later problems (Gutman et al., 2003; Seifer et al., 1992), each child was given a risk score created by summing the number of risks. The children were then split into two groups, those with zero or one risk ($n = 811$; 89%), referred to within the remainder of this report as displaying "low functional risk," and those with multiple risks ($n = 99$; 11%), referred to as displaying "high functional risk." Among children in the low functional risk group, 73% had no risk factors and 25% had one risk factor. Among children in the high functional risk group, 73% had two risk factors, 21% had three risk factors, and 6% had all four risk factors. Among this high functional risk group, academic problems were most common (72%), followed by social skills problems (63%), attention problems (59%), and externalizing problems (36%). (24)

Demographic Risk

We were also interested in following the trajectory of children who have typically been identified as at risk of school failure—children whose mothers have low levels of education. Among this sample, 249 children (27%) had mothers with less than a 4-year college degree. This cutpoint was chosen to provide an adequate sample size and is validated as a risk indicator in later analyses; implications of the moderate level of risks in this sample are included in the discussion. Ways in which school processes may moderate this risk factor were hypothesized to differ from the functional risk factor; thus, rather than composting demographic risk with those manifest in child behavior or skills, demographic risk was maintained as a separate indicator. Although low maternal education children were more likely than other children to display functional risks, the majority (78%) of those with low maternal education were in the low functional risk group. (25)

Child Outcomes

Achievement. Children's achievement was assessed with the Woodcock–Johnson Psycho-educational Battery-Revised (WJ-R; Woodcock & Johnson, 1989), a standardized measure of young children's academic achievement with excellent psychometric properties (Woodcock & Johnson, 1989). At each assessment point, several subtests were given out of the cognitive and achievement batteries. The cognitive battery included an assessment of long-term retrieval (Memory for Names), short-term (26)

Scientific method, Execution of research procedures. Data collection strategies are used to measure three dependent variables: (1) children's achievement, (2) student-teacher relationships, and (3) classroom processes.

memory (Memory for Sentences), auditory processing (Incomplete Words), and comprehensive knowledge (Picture Vocabulary). The achievement battery included measures of reading (Letter–Word Identification and Word Attack) and mathematics (Applied Problems). Memory for Names and Word Attack were only administered in first grade; all other tests were given at both 54 months and first grade. Because of the high levels of association between measures of cognitive ability and achievement, all subtests were composited at each time point, and are referred to for the remainder of this report as achievement scores. The coefficient α at 54 months was .80, and at first grade it was .83. Descriptives on the achievement battery are provided in Table 1.

(27) **Student–Teacher Relationships.** Children's relational functioning was assessed with the Student–Teacher Relationship Scale (Pianta, 2001), a 28-item rating scale, using a Likert-type format, designed to assess teachers' perceptions of their relationship with a particular student. This scale has been used extensively in studies of preschool-age and elementary-age children (e.g., Birch & Ladd, 1997, 1998; Hamre & Pianta, 2001; Howes & Hamilton, 1992). The conflict scale assesses the degree of negative interactions and emotions involving the teacher and child and contains items such as, "This child easily becomes angry at me" and "This child and I always seem to be struggling with each other." Coefficient α for conflict was .93 among this sample. Descriptives on the conflict scores are provided in Table 1.

(28) **Classroom Process.** Classroom process was measured using the Classroom Observation System for First Grade (COS-1; NICHD ECCRN, 2002b). Trained data collectors observed each classroom on 1 day during the spring of the first-grade year. Classrooms were observed for approximately 3 hrs during a morning-long period beginning with the official start of the school day on a day the teacher identified as being focused on academic activities. Observers made global ratings of classroom quality and teacher behavior using a set of 7-point rating scales. Some of the scales focused on global classroom quality and others focused specifically on the teacher's interaction with the study child. Global ratings of classroom-level dimensions included overcontrol, positive emotional climate, negative emotional climate, effective classroom management, literacy instruction, evaluative feedback, instructional conversation, and encouragement of child responsibility. Rating scales for the teacher's behavior toward the target child included sensitivity/responsivity, instrusiveness/overcontrol, and detachment/disengagement. A summary of these ratings is provided in Table 2. A rating of 1 was assigned when that code was "uncharacteristic," a 3 was assigned when the description was "minimally characteristic," a 5 was assigned when the description of the code was "very characteristic" of the classroom, and a 7 was assigned under circumstances in which the code was "extremely characteristic" of the observed classroom or teacher–child interactional pattern.

Table 1 Mean (Standard Deviation) on Academic Achievement (Woodcock–Johnson) and Student–Teacher Conflict by Time and Risk Status

	Kindergarten functional risk		Demographic risk (maternal education)	
	Low (n = 881)	High (n = 99)	Low (n = 661)	High (n = 249)
Woodcock–Johnson composite				
54 months	100.40 (10.79)	87.81 (10.42)	101.56 (10.50)	92.33 (11.24)
First	106.45 (9.78)	94.93 (10.42)	107.39 (9.58)	99.37 (10.54)
Student–teacher conflict				
K	9.80 (4.47)	15.74 (7.15)	10.00 (4.76)	11.74 (6.05)
First	10.28 (4.63)	14.59 (6.19)	10.32 (4.67)	11.91 (5.64)

Table 2 Summary of COS-1 Rating of Emotional and Instructional Climate

Composite Construct	Description (at high end)
Emotional support	
Teacher sensitivity	The sensitive teacher is tuned in to the child and manifests awareness of the child's needs, moods, interests, and capabilities, and allows this awareness to guide his/her behavior with the child
Intrusiveness (reversed)	An intrusive teacher imposes his/her own agenda on the child and interactions are adult-driven, rather than child-centered
Detachment (reversed)	A detached teacher shows a lack of emotional involvement and rarely joins in the child's activities or conversations
Positive climate	A positive classroom is characterized by pleasant conversations, spontaneous laughter, and exclamations of excitement. Teachers demonstrate positive regard and warmth in interactions with students
Classroom management	In a well-managed classroom, the teacher has clear yet flexible expectations related to the classroom rules and routines. Children understand and follow rules and the teacher does not have to employ many control techniques
Negative climate (reversed)	A negative classroom is characterized by hostile, angry, punitive, and controlling interactions in which the teacher displays negative regard, disapproval, criticism, and annoyance with children
Over-control (reversed)	The over-controlled classroom is rigidly structured and children are not given options for activities but instead must participate in very regimented ways
Instructional support	
Literacy instruction	This rating captures the amount of literacy instruction in the classroom. At the high end, the teacher frequently reads and teaches phonics and comprehension
Evaluative feedback	This rating focuses on the quality of verbal evaluation of children's work comments or ideas. At the high end feedback focuses on learning, mastery, developing understanding, personal improvement, effort, persistence, or trying new strategies
Instructional conversation	This scale focuses on the quality of cognitive skills or concepts elicited during the teacher-led discussions. At the high end children are encouraged to engage in conversations and expand on their ideas and perceptions of events. Teachers ask open-ended questions such as "what do you think?"
Encouragement of child responsibility	Children in classrooms high on this scale are encouraged to take on jobs, asked to offer solutions to classroom problems, and take responsibility for putting away materials, etc.

(29) Observers from all 10 sites trained on practice videotapes using a standardized manual that provided extensive descriptions of codes and anchor points. They trained on these videotaped observations prior to attending a centralized training workshop. After the training workshop, coders returned to their sites, conducted pilot observations, and trained on one to two more videotaped cases. All observers had to pass a videotaped reliability test involving six cases. Criteria for passing were an 80% match (with 1 scale point) on the global rating scales. All coders passed at these levels on a reliability test before being certified to conduct observations in the field.

(30) These scales were factor analyzed and averaged into two composite indicators of the classroom environment: emotional support and instructional support. The emotional support composite included ratings of overcontrol (reflected), positive emotional climate, negative emotional climate (reflected), effective classroom management, teacher sensitivity, intrusiveness (reflected), and detachment (reflected). The instructional support composite included ratings of literacy instruction, evaluative feedback, instructional conversation, and encouragement of child responsibility. These two composites are moderately associated with one another ($r = .57$). Table 2 provides a summary of these scales. For details on these composites and the training of observers, refer to NICHD ECCRN (2002b). Of note is the fact that, although only one observation was made for the majority of classrooms (one visit per child enrolled in the study), for almost 60 classrooms there was more than one child enrolled and hence more than one observation was conducted. For these classrooms, the correlations between pairs of the global ratings described above was, on average, higher than .70, indicating that these ratings reflect quite stable features of the classroom environment (NICHD ECCRN, 2004).

(31) The COS-1 composites were used to categorize classrooms into offering high, moderate, and low support (using 33% cutpoints). We used these cutoffs, rather than continuous measures of classroom process, because of our interest in creating a natural experiment and decided on cutting the sample in thirds to capture adequate range while allowing for ease of interpretation and analysis. For emotional support, the 303 classrooms in the Low category ranged from a score of 15.33 to 38.83 ($M = 33.15$; $SD = 5.16$), the 313 in the Moderate category ranged from a score of 39 to 44 ($M = 41.83$; $SD = 1.58$), and the 294 in the High category ranged from a score of 44.33 to 49.00 ($M = 46.53$; $SD = 1.45$). For instructional support, the 289 classrooms in the Low category ranged from a score of 4 to 13 ($M = 11.13$; $SD = 1.76$), the 328 in the Moderate category ranged from a score of 14 to 17 ($M = 15.41$; $SD = 1.07$), and the 293 in the High category ranged from a score of 18 to 28 ($M = 20.47$; $SD = 2.15$).

Results

Data Analysis Plan

Scientific method: Analysis of the data

(32) In order to establish whether instructional and emotional support in the first grade may moderate risk, we first had to establish two preconditions: (1) the existence of a natural experiment, in which children with varying risks backgrounds in kindergarten would sort into first-grade classrooms offering different levels of emotional and instructional support, and (2) whether the hypothesized risk factors were associated with poorer outcomes in first grade. The first precondition was assessed through examining the distribution of children in each risk group into classrooms offering high, moderate, and low support. The second precondition was assessed by conducting ANCOVAs in which risk status was used to predict first-grade outcomes, after adjusting for children's previous performance on these outcomes measures.

(33) Following these analyses, we turned to answering the main questions of this study: does classroom support moderate children's risk of school failure? First, the instructional and emotional support variables were entered into the ANCOVA models to assess whether classroom support had a main effect on children's outcomes. Next, following the recommendations of Kraemer, Stice, Kazdin, Offord, and Kupfer (2001) regarding testing the moderation of risk, a series of interactions were added to the model to test whether functional and demographic risks were moderated by classroom support variables. The relatively small *n*s among the risk groups provides for unbalanced ANCOVA designs. This situation may inflate Type I errors and thus increase the likelihood that true effects are not statistically significant (Keselman, Cribbie, & Wilcox, 2002). Although not ideal, this analytic approach was determined to be most appropriate for testing the natural experiment described above and provides a stringent test of potential effects for placement in high-quality classrooms. Further details on these analyses are provided below.

Selection into High- and Low-Support Classrooms

(34) The distribution of classroom support among the risk groups is presented in Table 3. Children displaying high functional risk in kindergarten were as likely as those with low functional risk to be in classrooms offering high instructional or emotional support. Children of mothers with less than a 4-year college degree were somewhat more likely than their peers to be in first-grade classrooms offering low instructional or emotional support. Despite this differential placement based on maternal education levels, there were enough low and high maternal education students placed in each of the three levels of classrooms to exploit a natural experiment. The implication of this differential placement will be considered in the discussion.

Table 3 Percentage Placement in First-Grade Classroom Support (Instructional and Emotional) by Risk Status

	Kindergarten functional risk			Demographic risk (maternal education)		
	Low (*n* = 881)	High (*n* = 99)	χ^2	Low (*n* = 661)	High (*n* = 249)	χ^2
Instructional support						
Low	31.6	33.3	0.28	28.6	40.2	11.76**
Moderate	36.5	32.3		37.1	33.3	
High	31.9	34.3		34.3	26.5	
Emotional support						
Low	32.4	40.4	3.54	29.8	42.6	13.87**
Moderate	35.3	27.3		35.6	31.3	
High	32.3	32.3		34.6	26.1	

*$p < .05$. **$p < .01$. ***$p < .001$.

Risks as Indicators of First-Grade Achievement and Relational Functioning Achievement

In order to provide a robust test of associations between risk and outcomes, we adjusted for children's prior scores on outcomes. Descriptive information on both previous and first-grade outcomes are presented for each risk group in Table 1. Consistent with hypotheses, results of ANCOVAs suggest that after adjusting for children's achievement at 54 months, children whose mothers had less than a 4-year college degree and those with high functional risk in kindergarten had lower achievement scores at the end of first grade (see Table 4). This suggests that not only do children at risk start school behind their low-risk peers, but the gap increases by the end of the first-grade year.

(35)

Table 4 Results of ANCOVAs Predicting First-Grade Achievement, Controlling for Previous Performance, From Risk and Classroom Process

	Achievement Woodcock–Johnson[a] (*n* = 908)				Teacher–child conflict[b] (*n* = 881)			
	Main effects		Moderation		Main effects		Moderation	
	F	Partial η^2	*F*	Partial η^2	*F*	Partial η^2	*F*	Partial η^2
Corrected model	152.17***	.57	78.45***	.58	31.38***	.22	22.03***	.23
Intercept	389.85***	.30	389.39***	.30	415.75***	.32	396.41***	.31
54 months WJ/K conflict	774.03***	.46	789.39***	.47	103.68***	.11	106.14***	.11
Female	12.80***	.01	13.59***	.01	20.77***	.02	20.64***	.02
Risk factors								
Maternal education—some college or less	8.97**	.01	8.335**	.01	0.74	.00	0.77	.00
High functional risk—kindergarten	14.92***	.02	13.20***	.02	23.58***	.03	19.27***	.02
Classroom process								
Instructional support	0.34	.00	0.13	.00	0.03	.00	0.60	.00
Emotional support	1.29	.00	3.20*	.01	2.30	.00	5.69**	.01
Risk: classroom process								
Maternal education: instructional support			6.68**	.02				
Maternal education: emotional support			1.82	.00				
Functional risk: instructional support			1.22	.00			0.69	.00
Functional risk: emotional support			4.57*	.01			3.62*	.01

*p 6.05. **p 6.01. ***p 6.001.

(36) To test whether these risks may operate differently for boys and girls, interactions between each risk and child gender were initially included in the ANCOVA model. Because none of these interactions were statistically significant at the $p < .05$ level, they were removed from the model.

Relational Functioning

(37) An identical set of analyses was performed to assess children's relational adjustment at the end of the first grade. Risk status was used to predict first-grade teachers' ratings of conflict with children, adjusting for kindergarten teacher ratings on this measure. Children in the high functional risk group had higher levels of teacher-rated conflict at the end of the first grade (see Table 4). Low maternal education did not arise as a significant risk factor for poor relational adjustment. As in the case of analyses on achievement, associations between risk and outcomes were not different between boys and girls and therefore these interactions were not included in the final model.

(38) These analyses provide support for the conceptualization of risk within this study. Even after controlling for previous performance, children at risk were not performing as well by the end of the first grade as were their peers without these risks, suggesting that these are indicators of increasing gaps between children at risk and those who are not at risk. Furthermore, the analyses provide evidence of the independence of each domain of risk; in the case of achievement, both functional and demographic risk independently predicted poorer outcomes. Only functional risk predicted higher rates of conflict with first-grade teachers.

Role of Instructional and Emotional Support in Moderating Risk Achievement

(39) Results presented in Table 4 suggest that neither support variable had a significant main effect on children's achievement. Because both risk indicators significantly predicted poorer achievement in the first grade, interactions between maternal education and functional risk status with each of the classroom support variables were entered into the final ANCOVA model. The two-way interactions between instructional support and maternal education and between emotional support and functional risk status both explained significant variance in the final model (Table 4). Effect sizes (partial η^2) were small; however, an examination of the estimated marginal means, presented in Figures 1 and 2, suggests that differences were meaningful, particularly considering that these models controlled for previous performance on very stable measures of academic functioning and are attributable to a relatively short period of time, that is, 1 school year. Figure 1 shows that, consistent with hypotheses, among children whose mothers had less than a 4-year college degree, those in classrooms with moderate and high instructional support had achievement performance in the first grade (controlling for 54-month achievement) equal to their peers whose mothers had more education. In contrast, children at high demographic risk who were in low instructionally supportive classrooms were performing significantly below their peers with low demographic risk.

(40) The main effect for the presence of high functional risk on achievement was moderated by the level of emotional support in the first-grade classroom (Table 4). Among children displaying high functional risk in kindergarten, those who were in highly emotionally supportive first-grade classrooms had similar scores on the first-grade Woodcock–Johnson as did their peers with low functional risk (see Figure 2). Children displaying high functional risk in kindergarten who were in low or moderately emotionally supportive classrooms had lower Woodcock–Johnson scores than did children in the low functional risk group.