10

Research Design in Qualitative/Quantitative/ Mixed Methods

Michael R. Harwell

University of Minnesota

The editors of this handbook assert that educational inquiry is weakly connected to the study of promising ideas. The implication is that much educational research has a "nuts and bolts" character that limits creative problem generation or problem solving and discourages risk taking. The consequences of this practice are missed opportunities to advance our understanding of important educational problems and solutions. This is a bold assertion, but it is one that is supported by an examination of published and unpublished educational studies.

A key player in this process is a study's research design. In many educational studies, research design reflects a "cookie cutter" approach in which the same designs are repetitively and narrowly applied in ways that may limit what is studied and how it is studied. Research-based ideas that cannot be easily mapped onto a small number of designs acceptable to funders and professional journals are likely to be abandoned or modified to fit widely used research designs.

The assertion of this chapter is that the study of promising ideas in education would be well served by the increased use of research designs that draw on multiple traditions of inquiry. In some cases, these designs are currently available, and the challenge is for more researchers to employ them (or continue to employ them) in a rigorous manner. In other instances, new designs will need to be developed. The net effect is that the catalog of research designs available to educational researchers is expanding dramatically. This change does not threaten or undermine the value of designs that are currently widely used, but rather it increases the pool of research designs available to support rigorous inquiry.

This chapter explores the role of research design in the study of promising ideas in education, assuming two conditions are present. First is that the idea being pursued is worth pursuing, that is, an idea that could reasonably lead to, and support, the formation of important educational questions and identification of solutions to important educational problems. Second is that every phase of the research process linked to studying a promising idea is rigorous. These conditions simplify an examination of the role of research design in the study of promising ideas.

Characteristics of three research methodologies (qualitative methods, quantitative methods, mixed methods) and their role in studying ideas

believed to be worth studying are described. I use two studies to illustrate strengths in the research design, as well as opportunities to enhance the results, and give greater attention to mixed methods because of their relative newness and potential. Three ways that researchers can enhance the ability of research designs to better support the study of promising ideas in educational studies are described. I conclude by arguing that mixed methods offer an especially promising path toward using research design in ways that support rigorous inquiry.

RESEARCH DESIGN

In educational research, it is usually possible (and certainly popular) to characterize a research study's methodology as qualitative; as quantitative; or as involving both qualitative and quantitative methods, in which case it is typically referred to as mixed methods. The term *research design* is widely used in education, yet it takes on different meanings in different studies. (The terms *research method* and *research design* will be used interchangeably in this chapter.) For example, in one study, research design may reflect the entire research process, from conceptualizing a problem to the literature review, research questions, methods, and conclusions, whereas in another study, research design refers only to the methodology of a study (e.g., data collection and analysis). Perhaps not surprisingly, there is variation within and between methodologies in how research design is defined. However, this variation does not affect an examination of the role of research design in promoting rigorous study of promising ideas and, thus, a single definition of research design is not adopted in this chapter. I assume that research questions are the driving force behind the choice of a research design and any changes made to elements of a design as a study unfolds.

Identifying a study's research design is important because it communicates information about key features of the study, which can differ for qualitative, quantitative, and mixed methods. However, one common feature across research designs is that at one or more points in the research process, data are collected (numbers, words, gestures, etc.), albeit in different ways and for different purposes. Thus, qualitative studies are, among other things, studies that collect and analyze qualitative data; quantitative studies are, among other things, studies that collect and analyze quantitative data; and so on.

Crotty (1998) described four key features to consider in research design: the epistemology that informs the research, the philosophical stance underlying the methodology in question (e.g., post-positivism, constructivism, pragmatism, advocacy/participatory; see Morgan, 2007), the methodology itself, and the techniques and procedures used in the research design to collect data. These features inform the descriptions of research designs below.

Qualitative Research Methods

Qualitative research methods focus on discovering and understanding the experiences, perspectives, and thoughts of participants—that is, qualitative research explores meaning, purpose, or reality (Hiatt, 1986). In other words,

> qualitative research is a situated activity that locates the observer in the world. It consists of a set of interpretive, material practices that make the world visible. These practices transform the world. They turn the world into a series of representations, including field notes, interviews, conversations, photographs, recordings, and memos to the self. At this level, qualitative research involves an interpretive, naturalistic approach to the world. This means that qualitative researchers study things in their natural settings, attempting to make sense of, or interpret, phenomena in terms of the meanings people bring to them. (Denzin & Lincoln, 2005, p. 3)

Central to this inquiry is the presence of multiple "truths" that are socially constructed (Lincoln & Guba, 1985). Qualitative research is usually described as allowing a detailed exploration of a topic of interest in which information is collected by a researcher through case studies, ethnographic work, interviews, and so on. Inherent in this approach is the description of the interactions among participants and researchers in naturalistic settings with few boundaries, resulting in a flexible and open research process. These unique interactions imply that different results could be obtained from the same participant depending on who the researcher is, because results are created by a participant and researcher in a given situation

(pp. 39–40). Thus, replicability and generalizability are not generally goals of qualitative research.

Qualitative research methods are also described as inductive, in the sense that a researcher may construct theories or hypotheses, explanations, and conceptualizations from details provided by a participant. Embedded in this approach is the perspective that researchers cannot set aside their experiences, perceptions, and biases, and thus cannot pretend to be objective bystanders to the research. Another important characteristic is that the widespread use of qualitative methods in education is relatively new, dating mostly to the 1980s, with ongoing developments in methodology and reporting guidelines (Denzin, 2006). The relative newness of this methodology also means that professional norms impacting research, including evidence standards, funding issues, and editorial practices, are evolving (see, e.g., Cheek, 2005; Freeman, deMarrais, Preissle, Roulston, & St.Pierre, 2007). Good descriptions of qualitative methods appear in Bogdan and Biklen (2003), Creswell (1998), Denzin and Lincoln (2005), Miles and Huberman (1994), and Patton (2002).

There are several categorizations of research designs in qualitative research, and none is universally agreed upon (see, e.g., Denzin & Lincoln, 2005). Creswell (2003) listed five strategies of inquiry in qualitative research that I treat as synonymous with research design: narratives, phenomenological studies, grounded theory studies, ethnographies, and case studies. Creswell also described six phases embedded in each research design that are more specific than those suggested by Crotty (1998), but still encompass virtually all aspects of a study: (1) philosophical or theoretical perspectives; (2) introduction to a study, which includes the purpose and research questions; (3) data collection; (4) data analysis; (5) report writing; and (6) standards of quality and verification.

Journals that publish qualitative methodology papers and qualitative research studies in education include *Qualitative Research, Qualitative Inquiry, Field Methods, American Educational Research Journal, Educational Researcher,* and the *International Journal of Qualitative Studies in Education.* Examples of the use of qualitative research designs are provided by Stage and Maple (1996), who used a narrative design to describe the experiences of women who earned a bachelor's or master's degree in mathematics and opted to earn a doctorate in education; Gaines (2005), who explored the process of interpreting interviews and media collected during the author's visits to India in ways that took into account his identity; Harry, Sturges, and Klingner (2005), who used the methods of grounded theory to develop a theory providing a new perspective on ethnic representation in special education; Brown (2009), who studied the perspectives of university students identified as learning disabled; and Chubbuck and Zembylas (2008), who examined the emotional perspective and teaching practices of a White novice teacher at an urban school.

These studies reflect several important features of qualitative research, including a focus on discovering and understanding the experiences, perspectives, and thoughts of participants through various strategies of inquiry. The studies were also conducted in naturalistic settings in which inquiry was flexible and guided by participants' comments, which in some instances were used to construct explanations of their views and perspectives. An important feature of several of these studies is their use of elements of different strategies of inquiry.

Quantitative Research Methods

Quantitative research methods attempt to maximize objectivity, replicability, and generalizibility of findings, and are typically interested in prediction. Integral to this approach is the expectation that a researcher will set aside his or her experiences, perceptions, and biases to ensure objectivity in the conduct of the study and the conclusions that are drawn. Key features of many quantitative studies are the use of instruments such as tests or surveys to collect data, and reliance on probability theory to test statistical hypotheses that correspond to research questions of interest. Quantitative methods are frequently described as deductive in nature, in the sense that inferences from tests of statistical hypotheses lead to general inferences about characteristics of a population. Quantitative methods are also frequently characterized as assuming that there is a single "truth" that exists, independent of human perception (Lincoln & Guba, 1985).

Trochim and Land (1982) defined quantitative research design as the

> glue that holds the research project together. A design is used to structure the research, to show how all of the major parts of the research project—the samples or groups, measures,

> treatments or programs, and methods of assignment—work together to try to address the central research questions. (p. 1)

Definitions of quantitative research design are complicated by the fact that this term is often used to identify the experimental design reflecting the arrangement of independent and dependent variables associated with data collection. Older categorizations of experimental designs tend to use the language of analysis of variance in describing these layouts—for example, a single-factor, completely between-subjects, fixed-effects design or a factorial split-plot design with between- and within-subject effects (Kempthorne, 1952; Winer, 1962). These descriptions were typically linked to an experimental design in which subjects were randomly assigned to treatment conditions, although quasi-experimental designs in which intact groups (e.g., African American and White students) are compared, and correlational designs in which no definable groups are present, also received attention (e.g., Campbell & Stanley, 1963).

More recent categorizations of research designs largely abandon the analysis-of-variance framework and instead rely on a trichotomy: randomized designs or, equivalently, randomized controlled trials in which participants are assigned at random to treatment conditions, quasi-experimental designs, and correlational designs (Pedhazur & Schmelkin, 1991; What Works Clearinghouse, 2008). These newer categorizations also tend to focus on conditions needed to justify strong causal inferences (Schneider, Carnoy, Kilpatrick, Schmidt, & Shavelson, 2007; Shadish, Cook, & Campbell, 2002).

A quantitative research design also involves phases that are superficially similar to those offered by Crotty (1998) and Creswell (2003) for qualitative research, but that are quite different in purpose and execution: (1) introduction to a study that includes the purpose and research questions; (2) theoretical perspectives or models; (3) methodology that encompasses sampling and an evaluation of external validity, instrumentation that may include an evaluation of construct validity, experimental design that includes an evaluation of internal validity and data collection, and data analysis that includes an evaluation of statistical conclusion validity; (4) reporting the results; and (5) conclusions and implications (Pedhazur & Schmelkin, 1991; Shadish et al., 2002).

Quantitative methods have a long history, dating to at least the 1930s, that has produced strong professional norms that impact research activities, such as the criteria used to make funding decisions (What Works Clearinghouse, 2008) and decisions about the kinds of studies and results likely to be published (Bozarth & Roberts, 1972; Leech, Morgan, Wang, & Gliner, 2010; Rosenthal, 1979).

Good descriptions of quantitative methods appear in Bryman (2004); Kerlinger (1964); Balnaves and Caputi (2001); Gay, Mills, and Airasian (2005); and the Research Methods Knowledge Base (Trochim, 2006). There are also several journals that publish quantitative methodology papers and quantitative research studies in education, including *Psychological Methods, Journal of Educational and Behavioral Statistics, British Journal of Mathematical and Statistical Psychology, American Educational Research Journal, American Journal of Evaluation,* and *Educational Evaluation and Policy Analysis.*

Examples of quantitative research designs in the educational literature are provided by Howell, Wolf, Campbell, and Peterson (2002), who used a randomized design to study the effectiveness of school vouchers in improving achievement; Jacob and Lefgren (2004), who used a quasi-experimental (regression discontinuity) design to study the relationship between receiving a remedial education and achievement; and Garner and Raudenbush (1991), who used a correlational design to examine the relationship between neighborhood characteristics and achievement. These studies reflect several important features of quantitative research, including an assumption that researchers set aside their experiences, perceptions, and biases in conducting the study and drawing conclusions, data collection that did not treat the researcher as the data collection instrument, and the use of hypothesis testing to make deductive inferences about characteristics of a population.

In sum, research design in education has moved from an almost exclusive reliance on quantitative methods to a more varied approach that includes qualitative research methods. As Lincoln and Guba (1985) pointed out, both qualitative and quantitative research methods emphasize truth, consistency, applicability, and neutrality while taking different procedural approaches to assure quality. Thus, it might be imagined that qualitative and quantitative methods have been

promoted to educational researchers in the spirit of "Use the methodology or methodologies that is (are) most appropriate and helpful for your purposes," but that is not quite the case. Rather, educational research in the last two decades has been home to considerable conflict fueled by advocates of each methodology who have emphasized weaknesses of the other in terms of epistemology, importance for the field, permissible inferences, and so forth (see, e.g., Berkenkotter, 1991; Denzin & Lincoln, 2005; Gage, 1989; Howe, 2009; Maxwell, 2004; Morse, 2006; Pegues, 2007). Much of the debate has been useful in encouraging researchers to think about features of these methodologies important for their work and for the field (e.g., Moss et al., 2009), but its prolonged and sometimes acrimonious nature has likely prompted many researchers to simply sit it out.

Mixed Methods

The qualitative versus quantitative debate has coincided with the rapid development of mixed methods, which combine qualitative and quantitative methods in ways that ostensibly bridge their differences in the service of addressing a research question. The roots of mixed methods are typically traced to the multi-trait, multi-method approach of Campbell and Fiske (1959, cited in Teddlie & Tashakkori, 2009, p. 31), although it is considered a relatively new methodology whose key philosophical and methodological foundations and practice standards have evolved since the early 1990s (Tashakkori, 2009).

Johnson and Turner (2003) have argued that the fundamental principle of mixed methods research is that multiple kinds of data should be collected with different strategies and methods in ways that reflect complementary strengths and non-overlapping weaknesses, allowing a mixed methods study to provide insights not possible when only qualitative or quantitative data are collected. Put another way, mixed methods research allows for the "opportunity to compensate for inherent method weaknesses, capitalize on inherent method strengths, and offset inevitable method biases" (Greene, 2007, p. xiii).

While mixed methods research combines qualitative and quantitative methods in ways that draw on the strengths of both traditions of inquiry, it is a clear step away from the boundaries and practices of those traditions, especially those linked to quantitative methods. According to Johnson and Onwuegbuzie (2004),

> Mixed methods research is formally defined here as the class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts or language into a single study. Mixed methods research also is an attempt to legitimate the use of multiple approaches in answering research questions, rather than restricting or constraining researchers' choices (i.e., it rejects dogmatism). It is an expansive and creative form of research, not a limiting form of research. It is inclusive, pluralistic, and complementary, and it suggests that researchers take an eclectic approach to method selection and the thinking about and conduct of research. (pp. 17–18)

This definition highlights the potential value of mixing multiple elements of qualitative and quantitative methods, as well as the potential complexity of doing so.

Caracelli and Greene (1997) identified three typical uses of a mixed methods study: (1) testing the agreement of findings obtained from different measuring instruments, (2) clarifying and building on the results of one method with another method, and (3) demonstrating how the results from one method can impact subsequent methods or inferences drawn from the results. These purposes appear in some form in many mixed methods studies in diverse fields including education (Taylor & Tashakkori, 1997), psychology (Todd, Nerlich, McKeown, & Clarke, 2004), criminology (Maruna, 2010), nursing and health sciences (O'Cathain, 2009), family research (Greenstein, 2006), and business (Bryman & Bell, 2007). In part, recent increases in the number of mixed methods studies can be attributed to increases in funding (Creswell & Plano Clark, 2007; Plano Clark, 2010). Still, journal articles and funding programs offer abundant evidence that widespread acceptance and funding of mixed methods in educational research is a work in progress (e.g., Alise & Teddlie, 2010; National Research Council, 2002, 2005; O'Cathain, Murphy, & Nicoll, 2007; What Works Clearinghouse, 2008).

Despite their growing popularity, there is not widespread agreement on exactly what constitutes a mixed methods study (Morse, 2010). For

example, some authors insist that a mixed methods study is any study with both qualitative and quantitative data, whereas other authors say a mixed methods study must have a mixed methods question, both qualitative and quantitative analyses, and integrated inferences (Tashakkori, 2009). There is also disagreement regarding various aspects of mixed methods, such as when mixing should occur (e.g., at the point of designing a study, during data collection, during data analyses, and/or at the point of interpretation).

Still other authors have criticized the whole idea of mixed methods (Denzin, 2006; Sale, Lohfeld, & Brazil, 2002; Smith & Hodkinson, 2005), criticism which is sometimes framed in terms of the response of advocates of a particular "stance" to arguments for mixing methods: the purist stance, the pragmatic stance, and the dialectical stance (Greene & Caracelli, 1997; Johnson & Onwuegbuzie, 2004; Lawrenz & Huffman, 2002). Those adopting a purist stance argue that mixed methods are inappropriate because of the incompatibility of the worldview or belief system (paradigms) (Tashakkori & Teddlie, 2003) underlying qualitative and quantitative methods, i.e., qualitative and quantitative methods are studying different phenomena with different methods (Smith & Hodkinson, 2005). Some purists have also raised concerns that mixed methods designs leave qualitative methods in the position of being secondary to quantitative methods (Denzin, 2006; Giddings, 2006; Yin, 2006).

Researchers who adopt a pragmatic stance argue that paradigm differences are independent of, and hence can be used in conjunction with, one another in the service of addressing a question (Johnson & Onwuegbuzie, 2004; Morgan, 2007). Wheeldon (2010) summarizes this view:

> Instead of relying on deductive reasoning and general premises to reach specific conclusions, or inductive approaches that seek general conclusions based on specific premises, pragmatism allows for a more flexible abductive approach. By focusing on solving practical problems, the debate about the existence of objective "truth," or the value of subjective perceptions, can be usefully sidestepped. As such, pragmatists have no problem with asserting both that there is a single "real world" and that all individuals have their own unique interpretations of that world. (p. 88)

However, the pragmatic stance also has its critics (Mertens, 2003; Sale et al., 2002). Dialectical researchers argue that multiple paradigms are compatible and should be used, but their differences and the implications for research must be made clear (Greene & Caracelli, 1997). It is important to emphasize that a researcher who adopts a dialectic stance would, other things being equal, draw on the same procedures for mixing as one adopting a pragmatic stance. The issue of stances is certainly not settled, and additional developments on this topic continue to appear in the mixed methods literature (e.g., design stance of Creswell & Plano Clark, 2007).

Mixed methods designs seem especially firmly rooted in the evaluation literature. An early important paper in this area was Greene, Caracelli, and Graham (1989), which highlighted five major purposes of (or justifications for) a mixed methods evaluation. One is triangulation, which examines the consistency of findings, such as those obtained through different instruments, and which might include interviews and surveys. According to Green et al., triangulation improves the chances that threats to inferences will be controlled. A second purpose is complementarity, which uses qualitative and quantitative data results to assess overlapping but distinct facets of the phenomenon under study (e.g., in-class observations, surveys), and a third is development, in which results from one method influence subsequent methods or steps in the research; for example, interviews with teachers might suggest that an additional end-of-year assessment be added. A fourth purpose of a mixed methods evaluation is initiation, in which results from one method challenge other results or stimulate new directions for the research; for example, teacher interviews might challenge results provided by administrators in a school district. The fifth and last purpose is expansion, which may clarify results or add richness to the findings.

A number of frameworks for mixed methods have appeared in this literature, many of which have built on the work of Greene et al. (1989) (Caracelli & Greene, 1997; Creswell, 2003; Johnson & Onwuegbuzie, 2004; Lawrenz & Huffman, 2002; Leech & Onwuegbuzie, 2007; Morse, 2010; Newman, Ridenour, Newman, & DeMarco, 2003; Tashakkori & Teddlie, 2003). These frameworks differ in many ways, but they all successfully convey a sense of the large number of methodological tools available to

researchers. However, none of these frameworks has been widely adopted.

The framework of Creswell (2003) is used here to categorize research designs in mixed methods, but this choice should not be interpreted to mean that it has been widely endorsed by mixed methods theoreticians and practitioners. This framework is described next, and studies by Howell et al. (2002) and Norman (2008) are used to illustrate strengths in their research design, as well as opportunities to enhance their results by employing a mixed methods approach. In doing so, I assume that different paradigms can be matched to answer research questions of interest, and thus I adopt a pragmatic stance. It is important to emphasize that there are many complexities not captured by these examples, including the nature and timing of the mixing and the use of elements from multiple mixed methods approaches in a given research design.

Creswell (2003) described six somewhat overlapping mixed methods research designs, referred to as strategies of inquiry, that guide the construction of specific features of a mixed methods study. The designs vary in whether qualitative and quantitative data are collected sequentially or concurrently, the weight given one kind of data or another, when the mixing is done, and the extent to which a theoretical perspective (e.g., post-positivism, constructivism) is present and guides the research design.

The first mixed methods approach described in Creswell (2003) is the *sequential explanatory design,* in which qualitative data are used to enhance, complement, and in some cases follow up on unexpected quantitative findings. In this approach, the focus is on interpreting and explaining relationships among variables and may or may not be guided by a particular theoretical perspective. Quantitative data are collected and analyzed first, followed by the collection and analysis of qualitative data, meaning that qualitative and quantitative data are not combined (mixed) in the data analysis; rather, integration takes place when the findings are interpreted. In general, results are interpreted in ways that usually give more weight to the quantitative component. The separate phases of design, data collection, and reporting for qualitative and quantitative data are considered strengths, because this arrangement is relatively easy to implement. The weaknesses of this approach are the time and resources needed for separate data collection phases (as opposed to only collecting qualitative or quantitative data) and the expertise needed to integrate the qualitative and quantitative findings. Morgan (1998) suggested that the sequential explanatory design is the most frequently used mixed methods approach.

As an example, consider the quantitative study of Howell et al. (2002). These authors were interested in learning if students who received vouchers allowing them to enroll in a private school (and who subsequently enrolled) showed improved achievement compared to students who did not receive a voucher and attended a public school. To explore this question, these authors examined achievement data for more than 4,000 students in three U.S. cities in which vouchers were randomly awarded via a lottery. This arrangement produced a randomized design in which receiving a voucher (yes, no) was the key independent variable, and student achievement was the key dependent variable. No theoretical model was offered. The results showed an average increase in the achievement of African American students who received a voucher and subsequently enrolled in a private school compared to African American students who did not receive a voucher, but not for any other groups in the study. From a quantitative perspective, this study has many strengths, particularly random assignment, that enhance causal arguments.

Howell et al. (2002) also commented that the reasons vouchers improved the academic performance of African American students and not other students were unclear. Is this difference related to a student's view of his or her academic skills, peer related, due to motivation differences or strong parental support for education, the student's interactions with school staff, or other factors? The Howell et al. comment suggests that a sequential explanatory research design would allow these authors to search for explanations of the finding that vouchers improved the academic performance of African American students but not other students. For example, collecting qualitative data at the end of the study through interviews of a purposively sampled group of parents of students who had and had not received a voucher could be helpful in enhancing and complementing the quantitative findings.

A second mixed methods approach, the *sequential exploratory design,* is essentially the reverse of the sequential explanatory design, with quantitative data used to enhance and

complement qualitative results. This approach is especially useful when the researcher's interest is in enhancing generalizability, and it may or may not be guided by a theoretical perspective. Creswell (2003) pointed out that instrument construction is an example of this approach, in that a draft of an instrument (survey, test) is piloted with a small number of cases who often provide important qualitative feedback about their experience with the instrument, followed, after appropriate modifications of the instrument, by using the instrument to collect quantitative data. The quantitative results are then used to enhance, complement, and possibly extend the earlier pilot results. The strengths and weaknesses of this approach are similar to those of the sequential explanatory design.

As an example, consider the mixed methods study of Norman (2008). Norman was interested in uncovering the perspectives of college academic counselors when advising newly admitted students on the appropriate mathematics course with which to begin their college mathematics career, and integrating this information with quantitative information reflecting the high school mathematics curriculum students completed and their scores on a college mathematics placement test. Norman's work was motivated by anecdotal evidence that students who completed a standards-based mathematics curriculum in high school were more likely to be advised to begin college with a less difficult mathematics course than comparable students who completed a traditional high school mathematics curriculum. Standards-based curricula in high school focus on a variety of mathematics topics every school year in a way that emphasizes problem solving and small-group work, and de-emphasizes algorithmic manipulation. Traditional curricula typically focus on algebra, algorithms, and repetition and depend heavily on the teacher for student learning (Schoenfeld, 2004).

Differential advising of students as a function of the high school mathematics curriculum they completed implies that some students are advised to begin their college mathematics with a course they should not take (e.g., precalculus rather than Calculus I). The implication of this practice is that students and colleges may be spending time, money, and effort on unneeded courses.

Norman (2008) adapted a theoretical decision-making model (Miller & Miller, 2005) to the process of how academic counselors make decisions, and used this model and a case study approach (Creswell, 1998). Norman purposively sampled 6 counselors and 24 students and collected data via in-depth interviews with the counselors and students that included their observations of the advising process, and high school and college records of students' mathematics course taking and grades. Norman reported that the counselors generally misinterpreted the mathematics portions of high school transcripts, which had important implications for advising a student on which college mathematics course to begin with. For example, a student whose transcript indicated that his or her highest completed high school mathematics course was "Integrated Mathematics IV," a standards-based course, was generally advised to start with a precalculus mathematics course, even though Integrated Mathematics IV is a precalculus course and the student should have been advised to enroll in Calculus I. Norman also reported that counselors who looked at transcripts of students who completed a traditional high school mathematics curriculum, in which the highest mathematics course completed was listed as precalculus, were more likely to recommend that a student enroll in Calculus I. Norman suggested that counselor views toward standards-based curricula may be related to working in a mathematics department, because mathematics departments have generally been quite critical of standards-based curricula (Roitman, 1999; Schoenfeld, 2004).

Norman (2008) also collected and analyzed quantitative data for a sample of more than 1,000 college freshmen that included information on the high school mathematics curriculum they completed; their score on a college mathematics placement exam; and the difficulty of their first college mathematics course, which was captured using a 4-point Likert variable (1 = a course that should have been completed in high school, which is sometimes referred to as a developmental course; 4 = a course whose difficulty exceeded that of Calculus I). The results of these analyses suggested that the curriculum a student completed was related to his or her mathematics placement score and the difficulty level of the student's first college mathematics course. In particular, students who completed a standards-based high school mathematics curriculum were more likely to enroll in a less difficult

college mathematics course compared to students who completed a traditional curriculum.

Norman's (2008) study reflects several exemplary features of mixed methods research, including using a case study approach that focused on discovering and understanding the experiences, perspectives, and thoughts of counselors and students, and the presence of quantitative data highlighting the possible consequences of counselors' misunderstanding the information on a student's high school transcript. While the exact timing of the quantitative and qualitative data collection in Norman's study was unclear, it appears the data were collected relatively close in time, suggesting that the results of one data collection did not influence the other data collection. However, the main question focused on counselors, and collecting and analyzing data for counselors first, intentionally using this information to guide the collection and analysis of quantitative data for students, and integrating the two sets of findings suggest that a sequential exploratory design would have been of value. For example, information obtained from the qualitative data about counselors' misunderstanding of high school mathematics course taking could have been used in developing questions for a student survey, such as on whether students thought their counselor misunderstood their high school course taking.

A third mixed methods approach is the *sequential transformative design,* in which either qualitative or quantitative data may be collected first. Here, the theoretical perspective underlying the methodology is critical to the conduct of the study, and the chosen methods should serve the theoretical perspective. Once again, qualitative and quantitative data are analyzed separately, and the findings are integrated during the interpretation phase. This approach is often used to ensure that the views and perspectives of a diverse range of participants are represented or when a deeper understanding of a process that is changing as a result of being studied is sought. Its strengths and weaknesses are similar to those of the sequential explanatory design.

For example, the decision-making model used by Norman (2008) suggested that collecting data for students who are affected (sometimes adversely) by the recommendations provided by counselors is important in evaluating and improving the advising process. Norman commented that most students follow the recommendation of which mathematics course to take, even if they disagree with it. Here, the study could focus on discovering and understanding students' experiences with the advising process and its implications for their college experience. A case study approach, using a purposively chosen sample of students who began their college mathematics course taking at different difficulty levels, would be appropriate. Information obtained from interviews and student academic records could be used to inform the construction of a survey to be sent to a representative sample of students. The survey results could be used to improve decision making by counselors and to enhance generalizability.

A fourth mixed methods approach is the *concurrent triangulation design,* which is used when the focus is on confirming, cross-validating, or corroborating findings from a single study. Qualitative and quantitative data are collected concurrently, such that weaknesses of one kind of data are ideally offset by strengths of the other kind. Typically, equal weight is given to the two kinds of data in mixing the findings, although one kind of data can be weighted more heavily. The qualitative and quantitative data are analyzed separately, and mixing takes place when the findings are interpreted. Important strengths of this approach are the ability to maximize the information provided by a single study, for example, when interest is in cross-validation, and a shorter data collection period compared to the sequential data collection approaches. Important weaknesses include the additional complexity associated with collecting qualitative and quantitative data at the same time and the expertise needed to usefully apply both methods. Discrepancies between the qualitative and quantitative findings may also be difficult to reconcile.

In the Howell et al. (2002) study, the primary finding was that the achievement of African American students who received a voucher to attend private school was on average higher than that of African American students who did not receive a voucher, and that this difference did not emerge for other student groups. Adopting a concurrent triangulation design could provide an explanation for these findings by collecting qualitative data in the form of interviews with parents of students who did and did not receive a voucher, and quantitative data in the form of student test scores and background information. This would offer an opportunity to corroborate findings from this study with respect to the improved achievement of African American students but

not other students. For example, a plausible outcome of concurrent data collection is that the qualitative data suggest that parents of African American students appeared to be more committed to, and enthusiastic about, their students' education in general and the voucher program in particular, than parents of other students, and that this enthusiasm persisted throughout the school year (Achievement tests used in this study were given at the end of the school year.). In this instance, the qualitative and quantitative information would provide corroborating evidence that the improved achievement of African American students could be attributed in part to receiving a voucher and enrolling in a private school, and in part to the support, encouragement, and motivation of students' parents.

The fifth approach is the *concurrent nested design,* in which qualitative and quantitative data are collected concurrently and analyzed together during the analysis phase. Greater weight is given to one kind of data, in the sense that one kind of data is typically embedded in the other. However, there may or may not be a guiding theoretical perspective. A popular application of this approach is with multilevel structures (Tashakkori & Teddlie, 2003), in which different levels or units of an organization are studied. Strengths of this approach include the shorter data collection period and the multiple perspectives embedded in the data, whereas weaknesses include the level of expertise needed to execute the study successfully, especially in mixing the qualitative and quantitative data within the data analysis, and difficulties in reconciling conflicting results from the qualitative and quantitative analyses.

In this design, qualitative and quantitative data are mixed in the analysis phase, a process that can take many different forms (see, e.g., Bazeley, 2009; Tashakkori & Teddlie, 2003). Caracelli and Greene (1993) described four strategies to mix qualitative and quantitative data in the analysis. One is data transformation, in which qualitative data are transformed to quantitative data or qualitative data are transformed into narrative, and the resulting data are analyzed. In Norman's (2008) study, this could involve transforming (i.e., rescaling) qualitative data in the form of interviews, field notes, and so on to a quantitative form that captures key themes in these data. Typically, the transformed qualitative data exhibit a nominal or ordinal scale of measurement.

A second data-mixing strategy described by Caracelli and Greene (1993) is typology development, in which the analysis of one kind of data produces a typology or set of categories that is used as a framework in analyzing the other kind of data. In Norman's (2008) study, analyses of the qualitative data could produce themes that allow a variable with nominally scaled categories to be developed, in which the categories provide an explanation of why participants became counselors. This variable could then be used in the quantitative analysis.

A third data-mixing strategy is extreme case analysis, in which extreme cases identified with one kind of data are examined with the other kind, with the goal of explaining why these cases are extreme. For example, multilevel analyses of quantitative data in Norman's (2008) study may suggest that some counselors are, with respect to the sample of counselors, statistical outliers (e.g., students linked to these counselors disproportionately began their college mathematics study with courses that should have been completed in high school). Qualitative data could be used to try to explain why these counselors appeared to disproportionately steer students to developmental college mathematics courses.

The fourth data-mixing strategy described by Caracelli and Greene (1993) is data consolidation/merging, in which a careful review of both kinds of data leads to the creation of new variables or data sets expressed in a qualitative or quantitative metric. The merged data are then used in additional analyses. In Norman's (2008) study, a review of the qualitative and quantitative data may suggest new variables. For example, a review of the counselor and student data may suggest constructing a variable capturing the extent to which students assert themselves in the advising process. Examples of data mixing appear in Caracelli and Greene (1993); Sandelowski, Voils, and Knafl (2009); and Tashakkori and Teddlie (2003).

Norman's (2008) study seems ready-made for a concurrent nested design, given the inclusion of both counselors and students. A key outcome in this study was the difficulty level of a student's first college mathematics course. Quantitative evidence of a relationship between a student's high school mathematics curriculum (standards-based, traditional) and the difficulty level of his or her first college mathematics course could be mixed with qualitative data obtained concurrently from

information provided by counselors. For example, qualitative data could be obtained using a narrative approach, in which counselors talked about the events in their lives that led them to become counselors and to continue to advise students. Transforming the qualitative counselor data such that the transformed variables reflect important themes allows this information to be directly included (mixed) with student variables in a quantitative multilevel data analysis, in which students are treated as nested within counselors (Raudenbush & Bryk, 2002). These analyses could explore the impact of nesting and the impact of counselor variables on student data and provide a powerful explanation of how and potentially why students enrolled in a particular college mathematics course.

The sixth mixed methods approach is the *concurrent transformative design.* As with the sequential transformative design, there is a clearly defined theoretical perspective that guides the methodology. In this approach, qualitative and quantitative data are collected concurrently and can be weighted equally or unequally during the integration of findings. Qualitative and quantitative data are typically mixed during the analysis phase. Strengths include a shorter data collection period, whereas weaknesses include the need to transform data so that it can be mixed in the analysis phase and difficulties in reconciling conflicting results using qualitative and quantitative data.

As an example of applying a concurrent transformative design, Norman (2008) reported evidence that some students were unhappy with the recommendations they received but felt powerless to change the process. Suppose that the goal of the study was to develop a theory explaining this sense of powerlessness in ways that could improve the advising process. Norman's adaptation of the Miller and Miller (2005) decision-making model includes a component that calls for students who are concerned about their recommendation to be heard. This could take the form of collecting qualitative information from students reflecting their experiences with the advising process and their view of its impact on their mathematics course taking, and simultaneously collecting qualitative data for a group of counselors. Developing a theory explaining the student's feelings of powerlessness would require that student and counselor responses be carefully examined and categorized to suggest themes to guide and inform the construction of the theory. This information could then be used to inform the construction of a survey that would provide quantitative information for a representative sample of newly admitted college students to enhance generalizability. The strengths and weaknesses of this approach are similar to those of the other concurrent approaches.

Other examples of mixed methods studies include Buck, Cook, Quigley, Eastwood, and Lucas (2009); Day, Sammons, and Gu (2008); Onwuegbuzie, Bustamante, and Nelson (2010); and Onwuegbuzie et al. (2007). Good descriptions of mixed methods can be found in Bazeley (2009), Creswell (2003), Creswell and Plano Clark (2007), Greene (2007), Reichardt and Rallis (1994), and Tashakkori and Teddlie (2003). The relative newness of mixed methods means that journals that concentrate on publishing mixed methods methodology papers and mixed methods studies in education are evolving. They currently include the *Journal of Mixed Methods Research, International Journal of Multiple Research Approaches, Qualitative Research Journal, American Educational Research Journal, Educational Researcher,* and *Educational Evaluation and Policy Analysis.*

In sum, the mixed methods approach offers a collection of flexible research designs that seem well suited to support rigorous examinations of promising ideas. The six designs of Creswell (2003) draw on the strengths of qualitative and quantitative methods to enhance inquiry in ways unlikely to occur with singular applications of these methods. Still, it is important to emphasize that mixed methods design continues to face a number of significant challenges (Bryman, 2007; Creswell, 2009; Lawrenz & Huffman, 2002; Tashakkori, 2009).

One important challenge is resolving outstanding disagreements over appropriate paradigms. The mixed methods literature contains a number of "mission accomplished" statements implying that important philosophical differences have been resolved, primarily by adopting a pragmatic stance (e.g., Carey, 1993; Creswell & Plano Clark, 2007; Haase & Meyers, 1988; Tashakkori & Teddlie, 2003). Yet it is clear that important differences remain (Sandelowski, 2001; Yin, 2006). A related challenge is achieving better agreement on what characterizes a mixed methods study and its components. For example, *what* are the components of a mixed methods study (research questions, design,

data collection, data analyses, interpretation), *when* should the components be mixed, and *how* (and how much) should they be mixed to justify the label mixed methods (see, e.g., Bazeley, 2009; Bryman, 2007; Johnson & Onwuegbuzie, 2004; Morse, 2010; Tashakkori & Teddle, 2003; Wheeldon, 2010)? Perhaps a single comprehensive framework outlining the characteristics and components of mixed methods studies will emerge; on the other hand, it is entirely possible that a crazy quilt pattern of frameworks will continue to define mixed methods studies.

THE STUDY OF PROMISING IDEAS AND RESEARCH DESIGN

Earlier descriptions of qualitative, quantitative, and mixed methods suggest three related ways for researchers to enhance the ability of research designs to better support the study of promising ideas. These represent both suggestions and challenges.

First, and most important, is for researchers to consciously shed narrow definitions of research design and to embrace more flexible designs that support rather than constrain the study of promising ideas. This may mean using a familiar research design in an unfamiliar way, such as substantially modifying elements of a design, for example, the way that data are collected, analyzed, or interpreted in a qualitative study, or changing features of the intervention being studied or the hypotheses being tested in the middle of a quantitative study because of preliminary evidence that a different research direction would be more productive. It may mean using a well-known research design to support exploratory rather than confirmatory work, adopting key elements of research designs used in other fields, such as dosage-escalation studies in medicine (Whitehead, Patterson, Webber, Francis, & Zhou, 2001), or abandoning the singular use of qualitative or quantitative methods in a study in favor of a mixed methods approach.

Part of the challenge of embracing more flexibility in research designs is technical, to the extent that it requires additional expertise on the part of a researcher or suggests the need to develop new research designs. For example, a researcher may plan to conduct a quantitative study to compare learning outcomes of a treatment group whose members receive a promising intervention against those of a control group using longitudinal data. However, in the service of developing a better understanding of the intervention, the researcher may decide to add a preliminary component to the study in which single-subject methods (Kratochwill & Levin, 1992) will be used to examine the learning trajectories of a small number of purposively sampled participants. This would require expertise in single-subject methodology that a researcher may or may not possess. Similar examples can be constructed for qualitative and mixed methods studies.

Still, the greater challenge for many researchers is likely to be modifying personal norms defining scholarship and the adequacy of a contribution to the field. This may require a researcher to step outside the methodological boundaries he or she has been trained to honor and, in some cases, enforce, and to embrace research designs the researcher may have been taught, and have taught others, are inferior. Embracing other methodologies in ways that cross disciplinary, funding, and publication lines, for example, serving as a member of an multidisciplinary research team that includes (and values) individuals with expertise in qualitative or quantitative methods, will require a certain amount of risk taking but will help to move the field toward more flexible designs.

Second, researchers can work (or continue to work) to modify professional norms in their roles as authors, manuscript reviewers, journal editors, panel members evaluating grant proposals, and so forth, to allow a greater range of studies and findings to be supported. This will require an artful balancing between encouraging risk taking (e.g., studies employing innovative interventions or emerging methods of qualitative inquiry) and the need to satisfy standards of research design and reporting that help to ensure the integrity of study results. Interestingly, there is growing evidence of support for doing just this. Some of this evidence appears in studies published in journals, such as the *American Educational Research Journal, Educational Evaluation and Policy Analysis,* and the *Journal of Mixed Methods Research,* that rely on multidisciplinary research teams. Other evidence is in the form of new funding programs at the U.S. Department of Education (primarily

supporting quantitative research), such as the Race to the Top and Investing in Innovation, which focus on innovative and ambitious approaches to educational change that draw on multiple disciplines.

An especially important challenge for modifying professional norms lies in the practices of educational research journals that typically publish quantitative studies. The filters enforced by the editorial process of these journals often reinforce existing and narrow professional norms. Evidence of their success can be found in many quantitative meta-analyses. In a typical quantitative meta-analysis, a sample of published and unpublished studies of a common phenomenon are collected, examined, and combined—for example, studies that have introduced interventions (vouchers, increased teacher professional development) designed to reduce the achievement gap between African American and White students. Variables capturing key features of each study, such as sample size, percentage of the sample that was African American, and characteristics of the interventions are developed, and each study is coded using these variables (Cooper, Hedges, & Valentine, 2009). Quantitative analyses of the resulting data typically provide evidence of the impact of editorial practices.

For example, limiting the analyses to published studies will typically reveal that all (or virtually all) reported at least one statistically significant result (see, e.g., Borman, Hewes, Overman, & Brown, 2003; Fukkink, & de Glopper, 1998; Purdie, Hattie, & Carroll, 2002), a reflection of the well-documented practice of publishing studies with statistically significant findings (Bozarth & Roberts, 1972; Hewitt, Mitchell, & Torgerson, 2008; Leech et al., 2010; Rosenthal, 1979). Other analyses will often reveal that only a handful of research designs were used in the pool of sampled studies.

The absence of published studies that failed to find evidence of a statistically significant effect and the prevalence of studies employing similar research designs send a powerful message to researchers: Risk taking is risky, and small successes are valued more than large failures. This reinforces a "cookie cutter" approach to research designs and provides further evidence of the need to strengthen the connection between educational inquiry and the study of promising ideas.

Third, researchers can consciously use techniques and procedures that enhance the depth and breadth of a study's findings. Advances in data collection, data analysis, report writing, and standards of inquiry and verification have been rapid—and in some cases stunning—and employing the techniques and procedures that reflect these advances should enhance the role of research design in supporting the study of promising ideas. For example, advances in computer hardware and software have significantly expanded the ways that data can be collected and analyzed, especially non-numerical data such as words and gestures in qualitative research (Bazeley, 2009; Creswell & Plano Clark, 2007), and the accessibility of techniques to control selection bias and take missing data into account in quantitative research (What Works Clearinghouse, 2008). Advances in standards of quality and verification enhance the transparency and integrity of results (see, e.g., Creswell & Plano Clark, 2007; Shadish et al., 2002).

Conclusion

The pursuit of promising ideas in educational research sounds noble, and it is. In this spirit, the task of this chapter was not to reexamine or reignite the conflict between qualitative and quantitative methods, nor to assess the peacemaking capacity of mixed methods, but rather to examine the role of research design in supporting the rigorous study of ideas that are believed to be worth studying.

An examination of qualitative and quantitative methods in education suggests that singular applications of these methodologies will continue to play an important role in research studying new ideas. Still, there is good reason to believe that mixed methods do indeed represent, as Teddlie and Tashakkori (2003) argued, a "third methodological movement" (p. 5), which is only now beginning to mature "as a well-established methodological alternative with agreed-on foundations, design, and practices" (p. 287).

In arguing for the value of mixed methods, Creswell and Plano Clark (2007) wondered, what would happen if

> quantitative researchers paid more attention to the incredible range of hypotheses that qualitative researchers have generated for them? And what if qualitative researchers spent more

> time exploring the range of phenomena that quantitative researchers have sought to define and test? (p. 59)

These are important and intriguing questions, and it seems clear that flexible research designs that aggressively support the study of promising ideas are needed to help answer them. Mixed methods designs seem especially well suited for this task. Whether mixed methods will someday enjoy an equal partnership with qualitative and quantitative research in education or supplant these methodologies is unclear and in many ways unimportant. What is important is expanding the catalog of research designs available to support rigorous inquiry. Still, change that leads to more researchers embracing more flexible research designs more conspicuously depends on modifying norms in ways that broaden the range of studies and findings supported in education. That change is underway is undeniable; what is less clear is the scope and pace of the change.

In sum, mixed methods research offers an especially promising path toward using research design in ways that support rigorous examinations of promising educational ideas. The time to fully embrace mixed methods designs has come.

REFERENCES

Alise, M. A., & Teddlie, C. (2010). A continuation of the paradigm wars? Prevalence rates of methodological approaches across the social/behavioral sciences. *Journal of Mixed Methods Research, 4,* 103–126.

Balnaves, M., & Caputi, P. (2001). *Introduction to quantitative research methods: An investigative approach.* Thousand Oaks, CA: Sage.

Bazeley, P. (2009). Mixed methods data analysis. In E. Halcomb & S. Andrew (Eds.), *Mixed methods research for nursing and the health sciences* (pp. 84–118). London: Wiley-Blackwell.

Berkenkotter, C. (1991). Paradigm debates, turf wars, and the conduct of sociocognitive inquiry in composition. *College Composition and Communication, 42,* 151–169.

Bogdan, R. C., & Biklen, S. K. (2003). *Qualitative research for education: An introduction to theories and methods* (4th ed., pp. 7–42). Boston: Allyn & Bacon.

Borman, G. D., Hewes, G. M., Overman. L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review* of *Educational Research, 73,* 125–230.

Bozarth, J. D., & Roberts, R. R. (1972). Signifying significant significance. *American Psychologist, 27*(8), 774–775.

Brown, S. (2009). Learning to read: Learning disabled post-secondary students talk back to special education. *International Journal of Qualitative Studies in Education, 22,* 85–98.

Bryman, A. (2004). *Social research methods* (2nd ed.). Oxford, UK: Oxford University Press.

Bryman, A. (2007). Barriers to integrating quantitative and qualitative research. *Journal of Mixed Methods Research, 1,* 8–22.

Bryman, A., & Bell, E. (2007). *Business research methods.* New York: Oxford University Press.

Buck, G., Cook, K., Quigley, C., Eastwood, J., & Lucas, Y. (2009). Profiles of urban, low SES, African American girls' attitudes toward science: A sequential explanatory mixed methods study. *Journal of Mixed Methods Research, 3,* 386–410.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

Caracelli, V. J., & Greene, J. C. (1993). Data analysis strategies for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 15,* 195–207.

Caracelli, V. J., & Greene, J. C. (1997). Crafting mixed-method evaluation designs. In J. Greene & V. Caracelli (Eds.), *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms. New directions for evaluation* (No. 74, pp. 19–32). San Francisco: Jossey-Bass.

Carey, J. W. (1993). Linking qualitative and quantitative methods: Integrating cultural factors into public health. *Qualitative Health Research, 3,* 298–318.

Cheek, J. (2005). The practice and politics of funded qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), The *SAGE handbook of qualitative research* (3rd ed., pp. 387–409).

Chubbuck, S. M., & Zembylas, M. (2008). The emotional ambivalence of socially just teaching: A case study of a novice urban schoolteacher. *American Educational Research Journal, 45,* 274–318.

Cochran, W. G., & Cox, G. M. (1950). *Experimental designs.* New York: Wiley.

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and*

meta-analysis (2nd ed.). New York: Russell Sage Foundation.

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

Creswell, J. W. (2003). *Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: Sage.

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

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

Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process.* London: Sage.

Day, C., Sammons, P., & Gu, Q. (2008). Combining qualitative and quantitative methodologies in research on teachers' lives, work, and effectiveness: From integration to synergy. *Educational Researcher, 37,* 330–342.

Denzin, N. K. (2006). The elephant in the living room: Or extending the conversation about the politics of evidence. *Qualitative Research, 9,* 139–160.

Denzin, N. K., & Lincoln, Y. S. (2005). Introduction. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (3rd ed., pp. 1–29). Thousand Oaks, CA: Sage.

Freeman, M., de Marrais, K., Preissle, J., Roulston, K., & St.Pierre, E. A. (2007). Standards of evidence in qualitative research: An incitement to discourse. *Educational Researcher, 36,* 25–32.

Fukkink, R. G., & de Glopper, K. (1998). Effects of instruction in deriving word meaning from context: A meta-analysis. *Review of Educational Research, 68,* 450–469.

Gage, N. L. (1989). The paradigm wars and their aftermath: A "historical" sketch of research on teaching since 1989. *Educational Researcher, 18,* 4–10.

Gaines, E. (2005). Interpreting India, identity, and media from the field: Exploring the communicative nature of the exotic other. *Qualitative* Inquiry*, 11,* 518–534.

Garner, C., & Raudenbush, S. J. (1991). Neighborhood effects on educational attainment: A multilevel analysis of the influence of pupil ability, family, school, and neighborhood. *Sociology of Education, 64,* 251–262.

Gay, L. R., Mills, G., & Airasian, P. W. (2005). *Educational research: Competencies for analysis and applications* (8th ed.). Upper Saddle River, NJ: Merrill.

Giddings, L. S. (2006). Mixed-methods research: Positivism dressed in drag? *Journal of Research in Nursing, 11*(3), 195–203.

Greene, J. C. (2007). *Mixed methods in social inquiry.* New York: Wiley.

Greene, J. C., & Caracelli, V. J. (1997). *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms. New directions for evaluation* (No. 74, pp. 19–32). San Francisco: Jossey-Bass.

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

Greenstein, T. N. (2006). *Methods of family research* (2nd ed.). Thousand Oaks, CA: Sage.

Haase, J. E., & Myers, S. T. (1988). Reconciling paradigm assumptions of qualitative and quantitative research. *Western Journal of Nursing Research, 10,* 128–137.

Harry, B., Sturges, K. M., & Klingner, J. K. (2005). Mapping the process: An exemplar of process and challenge in grounded theory analysis. *Educational Researcher, 34,* 3–13.

Hewitt, C., Mitchell, N., & Torgerson, D. (2008). Listen to the data when results are not significant. *British Medical Journal, 336,* 23–25.

Hiatt, J. F. (1986). Spirituality, medicine, and healing. *Southern Medical Journal, 79,* 736–743.

Howe, K. R. (2009). Positivist dogmas, rhetoric, and the education science question. *Educational Researcher, 38,* 428–440.

Howell, W. G., Wolf, P. J., Campbell, D. E., & Peterson, P. E. (2002). School vouchers and academic performance: Results from three randomized field trials. *Journal of Policy Analysis and Management, 21,* 191–217.

Jacob, B. A., & Lefgren, L. (2004). Remedial education and student achievement: A regression discontinuity analysis. *Review of Economics and Statistics, 86,* 226–244.

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

Johnson, R. B., & Turner, L. A. (2003). Data collection strategies in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 297–319). Thousand Oaks, CA: Sage.

Kempthorne, O. (1952). *Design and analysis of experiments.* New York: Wiley.

Kerlinger, F. (1964). *Foundations of behavioral research.* New York: Holt.

Kratochwill, T. R., & Levin, J. R. (1992). *Single-case research design and analysis: New directions for psychology and education.* Hillsdale, NJ: Lawrence Erlbaum.

Lawrenz, F., & Huffman, D. (2002). The archipelago approach to mixed method evaluation. *American Journal of Evaluation, 23,* 331–338.

Leech, N. L., Morgan, G. A., Wang, J., & Gliner, J. A. (2010). Statistical significance and evidence-based approach: Implications of non-significant results. *Journal of Non-significant Results in Education, 1,* 1–12.

Leech, N. L., & Onwuegbuzie, A. J. (2007). A typology of mixed methods research designs. *Quality & Quantity: International Journal of Methodology, 43,* 265–275.

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

Maruna, S. (2010). Mixed method research in criminology: Why not go both ways? In A. R. Piquero & D. Weisburd (Eds.), *Handbook of quantitative criminology* (pp. 123–140). New York: Springer.

Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher, 33,* 3–11.

Mertens, D. M. (2003). Mixed methods and the politics of human research: The transformative-emancipatory perspective. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 135–164). Thousand Oaks, CA: Sage.

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

Miller, M., & Miller, T. (2005). Theoretical application of Holland's theory to individual decision-making styles: Implications for career counselors. *Journal of Employment Counseling, 41,* 20–28.

Morgan, D. L. (1998). Practical strategies for combining qualitative and quantitative methods: Applications to health research. *Qualitative Health Research, 3,* 362–376.

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

Morse, J. M. (2006). The politics of evidence. *Qualitative Health Research, 16,* 395–404.

Morse, J. M. (2010). Simultaneous and sequential qualitative mixed method designs. *Qualitative Inquiry, 16,* 483–491.

Moss, P. A., Phillips, D. C., Erickson, F. D., Floden, R. E., Lather, P. A., & Schneider, B. L. (2009). Learning from our differences: A dialogue across perspectives on quality in education research. *Educational Researcher, 38,* 501–517.

National Research Council. (2002). *Scientific research in education* (R. J. Shavelson & L. Towne, Eds.). Committee on Scientific Principles for Education Research. Washington, DC: National Academies Press.

National Research Council. (2005). *Advancing scientific research in education* (L. Towne, L. L. Wise, & T. M. Winters, Eds.). Committee on Research in Education. Washington, DC: National Academies Press.

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

Norman, K. (2008). *High school mathematics curriculum and the process and accuracy of initial mathematics placement for students who are admitted into one of the STEM programs at a research institution.* Unpublished dissertation, University of Minnesota, Twin Cities.

O'Cathain, A. (2009). Mixed methods research in the health sciences: A quiet revolution. *Journal of Mixed Methods Research, 3,* 3–6.

O'Cathain, A., Murphy, E., & Nicoll, J. (2007). Why, and how, mixed methods research is undertaken in health services research in England: A mixed methods study. *BMC Health Services Research, 7.* Available at http://www.biomedcentral.com/1472-6963/7/85

Onwuegbuzie, A. J., Bustamante, R. M., & Nelson, J. A. (2010). Mixed research as a tool for developing quantitative instruments. *Journal of Mixed Methods Research, 4,* 56–78.

Onwuegbuzie, A. J., Witcher, A. E., Collins, K. M. T., Filer, J. D., Wiedmaier, C. D., & Moore, C. W. (2007). Students' perceptions of characteristics of effective college teachers: A validity study of a teaching evaluation form using a mixed-methods analysis. *American Educational Research Journal, 44,* 113–160.

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach.* Hillsdale, NJ: Lawrence Erlbaum.

Pegues, H. (2007). Of paradigm wars: Constructivism, objectivism, and postmodern stratagem. *Educational Forum, 71,* 316–330.

Plano Clark, V. L. (2010). The adoption and practice of mixed methods: U.S. trends in federally funded health-related research. *Qualitative Inquiry, 16,* 428–440.

Purdie, N., Hattie, J., & Carroll, A. (2002). A review of the research on interventions for attention deficit hyperactivity disorder: What works

best? *Review of Educational Research, 72,* 61–99.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed). Newbury Park, CA: Sage.

Reichardt, C. S., & Rallis, S. F. (Eds.). (1994). The qualitative–quantitative debate: New perspectives. *New Directions for Program Evaluation* (no. 61). San Francisco: Jossey-Bass.

Roitman, J. (1999). Beyond the math wars. *Contemporary Issues in Mathematics Education, 36,* 123–134.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86,* 638–641.

Sale, J. E. M., Lohfeld, L. H., & Brazil, K. (2002). Revisiting the quantitative–qualitative debate: Implications for mixed methods research. *Quality & Quantity, 36,* 43–53.

Sandelowski, M. (2001). Real qualitative researchers do not count: The use of numbers in qualitative research. *Research in Nursing & Health, 24,* 230–240.

Sandelowski, M., Voils, C. I., & Knafl, G. (2009). On quantitizing. *Journal of Mixed Methods Research, 3,* 208–222.

Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). *Estimating causal effects using experimental and observational designs.* Washington, DC: American Educational Research Association.

Schoenfeld, A. H. (2004). The math wars. *Educational Policy, 18,* 253–286.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference* (2nd ed). Boston: Houghton Mifflin.

Smith, J. K., & Hodkinson, P. (2005). Relativism, criteria, and politics. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (3rd ed., pp. 915–932). Thousand Oaks, CA: Sage.

Stage, F. K., & Maple, S. A. (1996). Incompatible goals: Narratives of graduate women in the mathematics pipeline. *American Educational Research Journal, 33,* 23–51.

Tashakkori, A. (2009). Are we there yet? The state of the mixed methods community. *Journal of Mixed Methods Research, 3,* 287–291.

Tashakkori, A., & Teddlie, C. (2003). *Mixed methodology: Combining qualitative and quantitative approaches.* Thousand Oaks, CA: Sage.

Taylor, S. J., & Tashakkori, A. (1997). Toward an understanding of teachers' desire for participation in decision making. *Journal of School Leadership, 7,* 1–20.

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

Todd, Z., Nerlich, B., McKeown, S., & Clarke, D. D. (2004). *Mixing methods in psychology: The integration of qualitative and quantitative methods in theory and practice.* East Sussex, UK: Psychology Press.

Trochim, W. M. K. (2006). *Research methods knowledge base* (2nd ed.). Available at http://www.socialresearchmethods.net/kb/

Trochim, W. M. K., & Land, D. A. (1982). Designing designs for research. *The Researcher, 1,* 1–6.

What Works Clearinghouse. (2008). Procedures and standards handbook (version 2.0). Washington, DC: U.S. Department of Education. Available at http://ies.ed.gov/ncee/wwc/help/idocviewer/doc.aspx?docid=19&tocid=1

Wheeldon, J. (2010). Mapping mixed methods research: Methods, measures, and meaning. *Journal of Mixed Methods Research, 4,* 87–102.

Whitehead, J., Patterson, S., Webber, D., Francis, F., & Zhou, Y. (2001). Easy-to-implement Bayesian methods for dose-escalation studies in healthy volunteers. *Biostatistics, 2,* 47–61.

Winer, B. J. (1962). *Statistical principles in experimental design.* New York: McGraw-Hill.

Yin, R. K. (2006). Mixed methods research: Are the methods genuinely integrated or merely parallel? *Research in the Schools, 13,* 41–47.

11

Intellect, Light, and Shadow in Research Design

John P. Bean
Indiana University

As people are known by the company they keep, scholars should be known by the ideas they keep. A poultice against stale ideas resides in the generation of new ones. Creating valuable new ideas, ideas that transform and energize a discipline, is the province of researchers. Ideas, which are directly related to research topics, are also directly constrained or expanded by research design. Big ideas change the way a generation of researchers thinks and works, and these ideas transform practice. Little ideas are refinements of big ideas, if they are of any use at all. The tension between studying big ideas that do not fit neatly into existing methods and studying safe and small ideas that produce predictable but trivial results creates a backdrop against which scholarship evolves.

Research, as a form of scholarship and creative work, is at the core of the academic enterprise. It is through research that universities contribute knowledge to society. Research provides the basis for what is taught in the disciplines and how members of a discipline understand their professional work. The design of research has a direct effect on what is discovered, the ideas that are created, and what forms of research contain legitimate professional information that is then passed on to the next generation of scholars in a field. The promise of research is that it will give us, as a society, what we need to know to improve our lives.

An axiological question arises for those planning to do research: Is research of value because it is true or because it is useful? Truth, the meaning of which is contested by philosophers, the existence of which is contested by postmodernists, and the use of which is contested by critical theorists, might be unattainable. I use the term *truth* as shorthand to mean that which is consistent with observation—if observations can be made—and is identified through procedures accepted in the discipline.

The Bright Promise of Knowledge

Basic research emphasizes the search for disciplinary truth, whereas applied research emphasizes finding out something useful. Both goals are attractive, and one does not preclude the other. Conjoined with the axiological question is a metaphysical one: Is there an objective reality

out there that we can discover, or is the world a product of the imagination, constructed in the minds of individuals and groups? Researchers design studies based on what they believe knowledge to be. The search for an objective truth involves a different path from the search for individual meaning or a consensus about intersubjective meaning.

In the professions, as opposed to the basic disciplines, utility is necessary. Professionals provide service to a client based on superior knowledge developed from a long study of the disciplinary research. According to Shils (1984), what separates academic knowledge from common knowledge is that academic knowledge is developed by a rigorous methodology. Researchers in a pure discipline (Biglan, 1973) attempt to establish truth in that discipline. Researchers in a profession have as their purpose not to attain pure knowledge but rather praxis, that is, to attain the best knowledge that can be applied in service of their clients' needs. To offer education as a societal good, money is spent, programs are funded, and teachers are trained and hired. If research is to inform these processes, then it is difficult to escape from pragmatism and positivism.

Research that advances methodology is of value to a field by developing better researchers, but such research is not always of direct use to a researcher's clientele. A field that emphasizes internal debates about philosophy, methodology, and definitions can be very lively but is in danger of becoming irrelevant. Well-designed research should deliver new understandings and new theories. The ultimate test of its value to the public will not rest with internal elaboration or with faculty members charming other faculty members; rather, it will be seen with improving understanding, teaching, learning, and organizing in a heterogeneous society. In what follows, I discuss some of the primary considerations that should inform research designs.

Theories

Educational researchers are interested in finding out how one thing is related to another; describing a set of phenomena; and establishing a basis on which to make claims, predictions, and explanations. Braithwaite (1955) writes that the purpose of science is theory and that, by extension, the purpose of research is to contribute theories or refinements of existing theories to science. Theory is a kind of abstraction, a simplification of reality that applies in similar circumstances and not just to the specific case at hand. For researchers, theories focus attention, limit choices, and provide explanations—characteristics that give good theories a central role in research design. For actors in the educational environment, they are practical for the same reasons.

Theories about social behavior have inherent limits. These are identified by Thorngate (1976) and elaborated on by Weick (1979). Thorngate developed a postulate of commensurate complexity in which there are trade-offs among a theory being *general,* a theory being *accurate,* and a theory being *simple.* A theory cannot be all three simultaneously; general accurate theories are not simple, accurate simple theories are not general, and simple general theories are not accurate. Weick provides examples of each. In developing a research design, the theory used, or the one the researcher is trying to develop, has more important effects on research design than does anything except the topic chosen for the research. Theory drives hypotheses. The choice of a theory to use or develop reflects the researcher's interest in being general, simple, or accurate and shapes the study accordingly. From my observations, I would suggest that educational research errs on the side of being simple and, with luck, accurate.

Theory is emphasized in research because it provides explanation. Without meaningful descriptions of the situation—that is, without identifying new ideas to be understood and related to each other by theories—research would not move forward. Designing research that identifies the ways in which people in a given situation view their worlds is a sensible starting place for meaningful research. Without important things to be studied, no theories would need to be developed, and a rigorous methodology to estimate relationships based on theory would not be necessary.

Topics and Ideas

How does one go about selecting substantive issues connected by a theory? Texts covering research in education or the behavioral sciences

have tended to either be mute on the question or provide only minimal advice (Gall, Borg, & Gall, 2002; Kerlinger, 1973; Krathwohl, 1988), with some improvement lately (Creswell, 2007). The selection of topics for study is neither innocent nor rational. It is not innocent, because being selected gives a topic legitimacy, creates the power of knowledge for those affected by the topic, and creates invisibility for those excluded. It is not rational, because researchers choose topics to study based on professional interests, not professional mandates, and on self-interest based on what researchers value, are curious about, or perceive they will profit from. What is studied in a discipline becomes what is taught in the discipline. Just like what is included in the curriculum is a political decision as well as an educational decision, what is studied is not neutral; it implies that what is studied is valuable.

It is axiomatic that the most important part of any study is the choice of a topic; that is, research findings depend on what is being researched. A good topic, when well studied, improves descriptions in a field, better explains how theories operate in the discipline, and shows how this knowledge can be applied to benefit clients and society as a whole. The research problem to be addressed and the statement of the purpose of the study focus research activities and limit the scope of the study. If the purpose is too broad, then the research cannot be accomplished with reasonable effort. If the purpose is too narrow, then the study is trivial. The statement of purpose is the most important sentence in a research proposal. Researchers need to avoid making Type III errors—asking the wrong question or not asking the right question—or what I consider to be Type IV errors, that is, studying the wrong thing.

Because we are well trained, we academics can justify studying practically anything. Politicians can mock this attribute of the academy, in the past handing out "Golden Fleece" awards for research that seemed to be fleecing the taxpayer. Rather than focusing exclusively on what we *do* study, it is important to recognize what we choose *not* to study. Nearly 30 years ago, Gardner (1983) identified eight areas of intelligence: verbal/linguistic, logical/mathematical, visual/special, physical/kinesthetic/intrapersonal/social, musical, intrapersonal/emotional, naturalistic, and spiritual. Of these, verbal/linguistic and logical/mathematical topics have dominated educational research, while the others have received little study or have been the province of specialized fields (e.g., music).

The topics we choose reflect a worldview of educational research: study the specific rather than the general, emphasize topics related to linguistic and logical cognitive development, and use methods that favor such studies. We also favor topics that reflect social justice, cognitive development that leads students into economically productive lives, and those topics that fit within predetermined disciplinary specialties. If education deals with the whole student, as Bowen (1977) suggests, it would behoove researchers to look for topics that consider more kinds of intelligence than those associated with linguistic and mathematical reasoning. Having become obsessed with certainty in our research, we have eschewed broad research topics that require integrating various components of human intelligence in favor of being able to say with great certainty something of diminutive importance. Highly specialized methodologies require highly refined topics, which provide little opportunity for transformational outcomes and the development of big ideas.

Choosing Topics

Identifying a *research problem* is the starting place for research. Research problems involve unresolved real-world conditions or situations, and it is the research problem that is nested between the topic and the purpose. Theoretical research problems deal with "we don't know why," descriptive research problems deal with "we don't know what," and practical research problems deal with "we don't know how." The best researchers are attracted to uncertainty, paradoxes, anomalies, contradictions, and ambiguities in the field. The significance of a problem is often based on the way in which it intersects with theoretical uncertainty and practical importance.

Probably the best way of finding a problem to study is to do extensive reading in an important topical area and find out what is poorly understood. Many articles contain sections that give suggestions for future research, and many have

glaring shortcomings that suggest a problem should be reexamined from a different perspective. Some researchers choose a methodology and then try to find a problem to match it. This approach creates an unnecessary constraint on what is to be studied, and the foolishness of this sequence cannot be emphasized enough. The cart does not pull the horse.

Other researchers are told what to study by a superior, such as an advisor, a provider of resources, an admired scholar in the field, or a coinvestigator. So long as the relationship is not exploitative, joining an ongoing research agenda has the clear advantage of the person being clear about what he or she will study. It has the disadvantage of not learning how to define one's own research problem. Typically, a study involves the following iterative process: Approach a topic of interest, read in the area, write a brief statement of purpose, discuss this statement with others, reflect, read, write, talk, and so on, until a compelling problem with a realistic scope has been identified and clearly expressed.

A hot topic is one where there is a lot of interest and a great likelihood of getting a project funded, finding others to participate in the study, and publishing the results. Good topics allow young scholars to demonstrate their research skills and increase the likelihood of getting further support for their research and publications in their fields. For better or worse, it usually means becoming more specialized. With luck, a good topic is one the researcher loves. Here, *love* is a passionate attachment to the research and an enjoyment of the research process. This attachment should not be confused with a lack of objectivity, but it involves caring about an increased understanding of a topic and a willingness to put forward the effort that results in influential research.

After deciding what to study and why such a study is worthwhile, the final part of designing research is to decide on the processes by which the research is to be accomplished. Many people, in thinking about research design, think that they need only be concerned with research methodology (Kerlinger, 1973, cited in Daniel, 1996). The following three questions always affect research design: (1) What will the researcher study? (2) Why is the research important? (3) How will the researcher carry out the research? Only the third question is related to methodology.

Different Approaches to Research

There are many ways in which to study the same topic, and these produce different results. Mitroff and Kilmann (1978) describe four approaches to research based on a Jungian analysis of our predispositions to approach decision making and obtain information, similar to the Myers-Briggs Type Indicator tests (Myers-Briggs, 1962). The four approaches include the following: the scientist, the conceptual theorist, the conceptual humanist, and the individual humanist. Given the topic *college student retention,* consider the way in which each methodology is chosen based on the researcher's interests, is different from the other methods, and produces different results. Descriptive phrases below are taken from tables in Mitroff and Kilmann on the pages indicated.

For the *scientist,* the approach should be objective, causal, cumulative, and progressive, emphasizing reliability and external validity and separating the scientist from the observed. It aims at precise, unambiguous empirical knowledge using strict logic (Mitroff & Kilmann, 1978, p. 34). The norms of this approach are known as the CUDOS:

> **C**ommunism, indicating that scientific knowledge is common property;
>
> **U**niversalism, indicating that scientific knowledge should be independent of the personality of the individual scientist;
>
> **D**isinterestedness, such that the scientist should observe what happens and not advocate a theory or experimental outcome; and
>
> **O**rganized
>
> **S**kepticism, where scientists should be critical of their own and others' ideas. (Merton, 1942/1973, p. 269)

An example of the scientific study of retention would be an organizational experiment based on the hypothesis that higher-achieving students are more likely to remain enrolled in college. Students would be randomly assigned to a treatment group or a control group. In the treatment group, students would participate in a

retention program, such as developing study skills, but otherwise would have experiences no different from those of the control group. After a given time period, the researcher would find out whether the retention rate for students who participated in the retention program was significantly different from that for students who did not. This information would be used to support or negate the hypothesis.

The *conceptual theorist* is involved with research that is impersonal, value-free, disinterested, imaginative, and problematic, involving multiple causation, purposeful ambiguity, and uncertainty. The theorist is interested in the conflict between antithetical imaginative theories, comprehensive holistic theories, and ever expanding research programs to produce conflicting schemas using dialectical and indeterminate logics (Mitroff & Kilmann, 1978, p. 56). A theorist conducting retention research would provide at least two theoretical explanations of retention behavior, use survey methods to gather information, analyze the data using statistics, find out whether the data supported one theory more than the other, and use that information to make more theories. Much of the empirical research reported in educational journals is a combination of the scientist and theorist—theory guiding social science inquiry into educational structures and processes.

The *conceptual humanist* (although I find *social humanist* to be a more accurate description) approaches research as a value-constituted, interested activity that is holistic, political, and imaginative; where multiple causations are present in an uncertain and problematic social environment; and with a deep concern for humanity. This approach recognizes the importance of the relationship between the inquirer and the subject and has the aim of promoting human development on the widest possible scale. The normative outcomes of such research would be economic plenty, aesthetic beauty, and human welfare. Similar to an action researcher, the social humanist prefers small-group dynamics where both the inquirer and the participants learn to know themselves better and work together to improve the situation (Mitroff & Kilmann, 1978, p. 76). A retention researcher using this approach could develop an ongoing program of action-oriented ethnographic research studies, where the researcher comes to better understand how the issues facing students contribute to their leaving, and tries to alleviate those conditions. The purpose is to increase the overall retention rate with the belief that students who complete college lead richer lives.

An *individual humanist* addresses inquiry as a personal, value-constituted, interested, and partisan activity, engaging in poetic, political, acausal, and nonrational discourse in pursuit of knowledge. Intense personal knowledge and experience are highly valued, aiming to help *this* person to know himself or herself uniquely and to achieve her or his own self-determination. The logic of the unique and singular has mythical, mystical, and transcendental overtones that operate as counternorms to the CUDOS (Mitroff & Kilmann, 1978, p. 95). A retention study from this perspective would try to develop a detailed understanding of a single student in the full context of his or her life. It could take the form of an "*N* of 1" case study: a phenomenological inquiry into who the student is, what the student finds at school, and how staying or leaving school would be better for this particular individual.

The purpose of presenting these four perspectives—and many more can be imagined—is to illustrate that there is no best way in which to study a topic. Different kinds of studies make different kinds of assumptions about what is important to know, serve different needs for different people involved in the studies, and produce different kinds of outcomes. The four perspectives were presented in what was once considered the normative order of acceptability: science-based research, theory development, action research, and phenomenology. One may be no more correct than the others. Some are more acceptable to certain audiences than to others, and each produces a particular outcome that favors some stakeholders more than it does others.

Methodology and the Scientific Approach

Methodology is often considered the core of research design. Kerlinger (1973, cited in Daniel, 1996) described as one of the research myths the idea that research design and research methods were synonymous, even though many researchers held this view. Methodology is the tool used to accomplish part of the study, specifically, how

to obtain and analyze data. It is subservient to choosing an important topic to study, matching the research problem and the methodology, and knowing what the results mean and how they can be applied. To do good research, the methodology used should be appropriate for the problem addressed. This is a necessary condition but not a sufficient one. An elegantly analyzed data set that was composed of ambiguously measured data that addressed a question of trivial importance is not likely to enter the annals of great research.

Educational research is part of the social science research tradition, a tradition that was influenced by research in the natural sciences. The natural sciences use the scientific method to solve research problems or support a perspective. The method contains a series of sequential steps similar to the following: Identify a problem, gather information from the literature about this question, develop a hypothesis in the context of a theory, collect data related to the hypothesis, analyze the data, and draw a conclusion related to the truthfulness of the hypothesis and correctness of the theory.

Scientists, as logical purists, build arguments on *falsifiability* and the *law of the excluded middle*. This law states that *A* and *not-A* cannot exist simultaneously. But if *A* stands for "this program helps students to learn" and *not-A* stands for "this program does *not* help students to learn," then both can be true, as in the case of aptitude–treatment interactions; that is, a treatment could be effective for a high-aptitude student but not effective for a low-aptitude student. If both are true, then the law of the excluded middle is violated and falsifiability cannot be demonstrated. This situation is problematic for scientific research in education.

Education is not a scientifically based process, partly because the term *education* is ideological and idiosyncratic, much different from the term *temperature*. At best, scientific research can shed light on narrowly defined educational behaviors, and researchers can hope for—but cannot guarantee—a cumulative effect. When a government policy assumes that education is equivalent to improving the score on a test, the society will not have a moral compass and will not be educated. Feyerabend (1993) holds the view that if we do not separate scientific research and the state, as we have separated the church and the state, irreparable harm will be done.

In the same way that social science research has imitated the natural sciences, educational research has imitated social science research. Research in these areas may be separated more by the topic studied than by the rigor of the methodology. As with social science research in general, educational research might pretend a level of control so as to carry out an experiment. Researchers give themselves solace that "other things are equal," or "other effects are random," or "spuriousness is not a problem," and proceed as if the social world were simple and understandable in the same way that the traditional world of the pure sciences can be.

A Traditional Approach to Designing Educational Research

Graduate programs in education typically are arranged in departments that reflect academic subspecialties such as history of education, sociology of education, anthropology of education, counseling psychology, experimental psychology, higher education, school administration, and curriculum and instruction. Most of these programs require a course in research design appropriate for their field. My discussion revolves around quantitative and qualitative approaches to research, terms that reify categories that are themselves overlapping and arbitrary, but are commonly used to describe research courses.

A simple way of looking at a proposal is to see how it answers the three questions posed earlier: What is this study about? Why is the study important? How will the researcher conduct the study? The study itself must cover these three questions and answer two additional questions: What did the researcher find? What do the findings mean? Research designs revolve around a limited number of elements. Their exact use and exposition vary depending on the particular approach taken. The purpose and promise of these elements have been identified and discussed by a number of textbooks such as those of Gall and colleagues (2002) and Creswell (2007). These texts identify many of the issues facing researchers, especially those who are new to the process.

Although not suitable for all studies, well-designed quantitative research usually addresses the areas presented in the following outline:

I. Introduction to the topic of the research
   a. Background and context in which that topic has been studied
   b. Importance of studying the topic, including
   c. Practical value of the study
   d. Research problem to be addressed
   e. Purpose of the study
   f. Objectives or questions to be addressed
   g. Definitions and related constructs
   h. Assumptions used in the study
   i. Limitations of the study
   j. Scope of the study

II. Using the literature to build conceptual arguments
   a. Relevant theories to guide the study
   b. The findings of other researchers that identify important variables related to the purpose
   c. Identification of the dependent variable, independent variables, and theories linking these variables

III. Methodology to be used in the study
   a. Site, sample, or selection procedure for respondents, including
   b. How the data will be gathered
   c. How the data will be measured
   d. How the data will be analyzed
   e. Why these methods are appropriate. [This information usually completes the design of a research proposal and appears as the first part of a finished study. Finished studies also include the following.]

IV. Findings
   a. A description of the sample actually analyzed in the study
   b. Description of the data
   c. Treatment of missing cases
   d. Possible or known biases
   e. Description of how the researcher met the assumptions required to use the chosen statistics
   f. Presentation of the data
   g. Support or lack of support for the hypotheses or theories used
   h. Discussion of the findings
   i. Conclusions

V. Summary of the study
   a. Practical implications of the study
   b. Areas for future research

Qualitative research can involve using the five research traditions identified by Creswell (1998)—biography, ethnography, grounded theory, case study, and phenomenology—which can be used singly or in combination. General headings appropriate for qualitative studies include the topic, focus and purpose, significance, related literature, methodology, presentation of the data, interpretation of the data, and conclusions. Detailed headings would be similar to the following:

I. Topic to be studied
   a. The overall interest focusing on what will be explained or described
   b. Organizing metaphor (like grounded theory)
   c. The mystery and the detective
   d. Hermeneutic elements
   e. The significance of the study
   f. Why the reader should be interested.

II. Getting information away from the source
   a. Relevant literature
   b. Theories
   c. Findings for content area, themes, foci, or analogous situations

III. The selected qualitative method
   a. How information is obtained
   b. How sense is made from it
   c. Site selection
   d. Informant selection
   e. Data collection
   f. Data analysis
   g. Data evaluation
   [The methodology can be in a separate chapter, part of the first chapter, in an appendix, or woven into the chapters that present respondent information. This section is followed by one or more chapters which does the following.]

IV. Present the text as natural answers to natural questions.
   a. Presentation formats include stories, tables, interviews, documents, narratives, photographs, videotapes, vignettes, texts of various kinds, descriptions, and routines (see Schwartzman [1993], from which some of these headings were taken).
   b. Raw data can be presented without comment, presented and interpreted simultaneously, or presented and interpreted in a later section.

V. Findings
   a. Conclusions
   b. Recommendations for others

The structure of qualitative studies is more idiosyncratic than the more formal structure of traditional quantitative studies. In qualitative research, the sample *is* the study, and the reasons for selecting the sample need to be emphasized.

Most researchers, when approaching a topic they care about, have tentative hypotheses about what causes what or predispositions to thinking that the world operates according to certain principles that also apply in this area. People bias their observations based on their experience. All of us know that we have certain biases, and we can try to counter those biases in our research by looking for evidence that directly contradicts what we expect. In the unconscious, there is a second set of biases of which, by definition, we are not aware. Sometimes peer readers can help the researcher to discover what is missing or what is inappropriately under- or overemphasized in the study.

After analyzing the data in a quantitative study, the researcher presents the findings. Typically, it is rather straightforward, because the data to be gathered and the analyses proposed for the data were specified in the proposal for the study. For qualitative researchers, the data, the findings, and the method might not be distinct. The narrative that presents selected questions and answers can represent findings based on data that came from the method by which questions were developed. As the previous sentence suggests, it is a convoluted process. The presentation might revolve around respondents' experiences and understandings, a chronology of events, or themes supported by respondents' statements. In ethnographic studies, the descriptions of lives in context can stand on their own (Lawrence-Lightfoot, 1995). Thick descriptions (Geertz, 1973) might provide greater insight into the education of a student in a school than would an analysis of variables. In most studies, some analysis of the descriptions is expected. This predisposition is part of the legacy of pragmatism; researchers in education are expected to identify how knowledge gained from the study can improve professional practice.

In designing a study, the background, context, importance of the topic, and presumed practical value of the study come from the literature written about the topic or analogous literatures in similar fields. For example, the study of college student retention can be considered to be analogous to the study of turnover in work organizations, and the literature in one area can be used to reinforce the literature in the other area (Bean, 1980). The use of literature in quantitative studies, however, can differ substantially from the use of literature in qualitative studies.

In a quantitative study, the literature is used to identify the importance of the dependent variable, relevant independent variables, and theories that bind these factors together, to justify the use of statistical procedures and to provide a context for the discussion. In qualitative studies, a premium is placed on the ability to see what is before the researcher. Our ability to observe is both heightened and diminished by our prior knowledge and expectations (Bean, 1997). It is heightened by making ourselves aware of important details to observe, and it is diminished because we focus only on those details. Due to the preconceived notions of the researcher, those factors actually influencing the respondents' world might not be identified. When the literature shapes the way in which we view the world, what is actually before us is replaced by what we expect to see.

A review of the literature, as a stand-alone section summarizing research in the topical area, makes little sense. The literature, as a compendium of related information, should be used to advance arguments related to the importance of the subject. It should identify topical areas that are either well or poorly understood, identify and describe relevant theories, identify and describe appropriate methodologies to study the topic, describe dependent and independent variables if relevant, provide definitions, and provide a context to discuss the findings from the study.

Google Scholar, Dissertation Abstracts International, ERIC Documents, the ISI Web of Knowledge, and the proceedings of relevant professional organizations all can be used to access current research. A condemning retort is that the literature in a study is dated. This phrase has some interesting subtexts. The first assumption is that the most recent research is the best research and that previous research is irrelevant. A second assumption is that all research is of limited generalizability over time so that if it is older than, say, 5 years, it is irrelevant. In either case, dated research is of marginal value. By extension, the research that a person is currently conducting is also of marginal value because it

will be useful for only 5 years. This planned obsolescence of research becomes a justification for a frenzied increase in the rate of publication and is counterproductive in terms of identifying important and durable ideas in the field. Literature should not be weighed nor considered dated after 5 years.

In traditional quantitative research, the topic contains the dependent variable, and the factors associated with it identify the independent variables that have been found to have important effects on the dependent variable. In studies that are not codifications—that is, not extensive reviews of the literature for the heuristic purpose of organizing what is known about a topic—citing the literature should be done for the purpose of building an argument, not simply to show familiarity with the canon.

Since the 1960s, the number of statistical analyses available for researchers to include in their designs has increased dramatically. Five commercial statistical packages bear the initials SAS, SPSS, BMDP, GLIM, and HLM. The development of these statistical packages has allowed ever more complex analyses to be performed. National data sets from the National Center for Educational Statistics (NCES) and other sources have provided the opportunity to bring order to vast amounts of data.

For the description of large-scale phenomena, these data sets can be very valuable. For analyzing the causes of behavior, however, the attempt to gain a broad vision masks individual or small-group differences. Longitudinal studies almost always suffer from decay; that is, measures may differ from year to year and respondents drop out of the study. So the comparisons from year to year might not be the result of what people report; rather, they might be the result of changes in who is doing the reporting.

The availability of data and the means to analyze them raised the level of expectation in some journals that such analyses should be the norm. What is certain is that during the past 50 years, the sophistication of analyses has increased. The literature shifted from normed surveys that reported frequencies, to chi squares, to analyses of variance (ANOVAs) and simple correlations, to factor analysis and multiple regression, to causal modeling with ordinary least squares path analysis, to maximum likelihood used in linear structural relations (LISREL) modeling, and to generalized linear modeling (GLIM) and hierarchical linear modeling (HLM).

The increase in complexity is associated with an increase in agitated exchanges between statisticians about whose method is correct. An improved methodology has not been matched by these studies becoming more influential in policy making or practice (Kezar, 2000). The debate is sometimes invisible to the public, taking place between the author of a piece of research and the consulting editors who review the research, and sometimes it appears in journals such as the *Educational Researcher.*

Data

Quantitative studies require data that can be used in statistical analyses. The sources of data can vary widely—historical documents, governmental records, organizational records, interviews, standardized surveys, questionnaires developed as part of the research protocol for a particular study, unobtrusive measures, observations, participant observation, and so on. The quality of the research depends on the quality of the data analyzed; data analysis has only a secondary influence.

The quality of the data varies greatly. Good research design requires that the researcher understand the strengths and weaknesses of the data. Historical data can reflect the biases and ideological preferences of those who recorded it. People who provide data can intentionally distort it to put themselves in a better light, for example, reporting that they had higher grades than they actually did. Survey data might come from a biased sample reflecting only the experiences of high–socioeconomic status respondents. Questions in a survey might be ambiguously written, or a single item might contain two questions with different answers, for example, "How satisfied are you with your salary and fringe benefits?" Survey data that require a forced-choice response might not represent the real interests of the respondent. A respondent might have no opinion on most of the questions and refuse to answer them. Other respondents might not want to reveal personal information and so might misrepresent their actual incomes, whether they have ever plagiarized, or how much they use drugs or alcohol. Although the questionnaire is not missing any data, the data provided might be intentionally inaccurate.

In other cases, respondents might not understand the questions, might not care about the answers given, or might become fatigued while filling out the questionnaire, so that the accuracy of the responses is different for the beginning and the end of the questionnaire. A well-written question should reflect one bit of information about the respondent unambiguously and reliably, and the answer to the question should match observable facts.

It is acceptable to use problematic data if the analyst understands and acknowledges the problems that exist in the data. For example, a data set might not be random but might be completely representative of one subgroup in the population studied. The bias in this sample can make conclusions drawn about the well-represented group accurate, but the conclusions would not apply to the whole population. Although not representative, the data might be useful to see whether a hypothesized relationship exists at all, that is, as a test of theory.

Data gathered from face-to-face interviews for qualitative research have the potential to yield a gold mine of insights into the people's lives and situations. There is no substitute for prolonged and focused conversations between trusted parties to discover what is important to the interviewees and how respondents understand key elements in their own lives. When mishandled, interview data can reflect what the interviewees think the interviewers want to hear, normatively appropriate responses, the fears and biases of the interviewees, and the fears and biases of the interviewers. Data flaws become limitations of the study for which the only response is to caution the reader that the results are far from certain.

Ethics and Institutional Review Boards

Before proceeding with an examination of research methods, there are some ethical and legal considerations that have obtrusively entered the development of a research protocol. In line with designing research to be useful, it should also be designed to be ethical. The most obvious ethical problems arise when a research procedure causes harm to those who are asked or forced to participate in the process. There are several well-known cases of abuse, including psychological studies where participants were put in unusually stressful situations (Baumrind, 1964; Milgram, 1974) and medical research where participants were given diseases or intentionally denied treatment (Jones, 1993).

The bureaucratic response to these ethical violations was to create rules that would include everybody doing any kind of research that involved contact with living people. Bureaucratic actors, evaluating research they are not conducting themselves, become the gatekeepers of ethical behavior. This responsibility is misplaced; researchers themselves should be responsible for protecting the interests of participants in their studies. I am not naïve enough to think that all researchers are ethical or that institutional review boards (IRBs) or protection of human subjects committees will go away. The problem is that ethical judgments about research have been made extrinsic to the research process. Researchers need to design research that does just what the committees want—to protect the participants of a study from harm. If researchers are not socialized to provide these protections, IRBs might not help. The enforcement system used, which involves taking federal support away from ethical researchers because they happen to be at an institution where one person did not comply with the guidelines, is a collective punishment which is itself unethical. IRBs have the enormous power of being able to block research, and the potential for abusing power must be kept under constant scrutiny.

For qualitative researchers especially, complying with a written informed consent form can damage the trust required to conduct a study. The study of any group that dislikes authority is made impossible, or at least less reliable, by asking participants at the outset to sign a form that says, "You should know that this researcher does not intend to hurt you." A journalist and an ethnographer can conduct and publish identical studies. However, the journalist needs no informed consent from those who are interviewed for the story, whereas the ethnographer at a research institute needs IRB permission to ask the same questions. The journalist is protected by freedom of speech, whereas academic freedom, according to IRB rules, provides no such protection for the researcher.

While much of the antagonism between IRBs and the faculty involve what are seen as nuisance parameters, as hurdles to be jumped, IRB guidelines constitute a direct attack on academic freedom. When a faculty member has

to get permission to "do research as he or she sees fit," there is no academic freedom. There are three traditional threats to academic freedom: economics, religion, and politics. One could argue that IRBs are simply a new form of the economic threat, being that failure to follow IRB guidelines will result in the loss of substantial federal funding for research at the institution. Too often, I fear, these offices displace their goals of protecting human subjects and replace them with the goal of making researchers follow their rules. The direct and opportunity costs to institutions in wasting faculty time is real, but the greater problem is researcher self-censorship—not doing research because of the fear that it will not be approved by the IRB. Academic freedom is the most central value in American higher education, and anything that puts it at risk needs greater justification than IRBs have provided for their behavior. Regardless of what happens to IRBs, research should be designed to protect everyone, to benefit the participants in the research, and to protect society from ignorance.

Generalizability

Generalizability is the central bulwark of the scientific research in education approach. In a 2002 National Research Council report, the editors observed, "Regularity in the patterns across groups and across time—rather than replication per se—is a source of generalization. The goal of such scientific methods, of course, remains the same: to identify generalized patterns" (p. 82).

Generalizability is a powerful statistical tool that allows researchers to make predictions about patterns of behavior in a population, such as the percentage of people who will vote as independents, based on a measure of that behavior taken from a sample of the population. It is attractive to policy makers because it suggests the extent to which a particular solution will work everywhere in the population. As the behavior in question gets more complicated, such as how students learn ethical behavior, generalization is of more limited value. Lincoln and Guba (1985) describe this limitation well:

> Generalizations are nomothetic in nature, that is, lawlike, but in order to use them—for purposes of prediction or control, say—the generalizations must be applied to particulars. And it is precisely at that point that their probabilistic, relative nature comes into sharpest focus. (p. 116)

Does what works 90% of the time for the participants in a study work for one particular teacher in one particular class dealing with one particular subject? Tutoring generally helps students to learn how to read, but for a student who is acting out against authority, and who views the tutor as an authority figure, tutoring might prevent the student from learning to read.

As a "reductionistic fallacy" (Lincoln & Guba, 1985, p. 117), generalization simplifies decision making and simultaneously reduces the understanding of the particular. Teachers operate in particular environments, and the findings from a "scientific" study with a high degree of generalizability do not ensure a program's utility in a given classroom. The purpose of scientific research is to eliminate uncertainty so that the operator can predict and control the future. Applied to education, this goal is not a research norm or a teaching norm but rather a political norm.

The Shadow of Research Design

Research is seductive because it promises to give the participants, as producers or consumers, those things that they imagine they want. We are seduced by research, viewing it as a beatific process by which we can glimpse the bright light of pure knowledge. Scholars would have no agenda other than furthering knowledge, a value shared by those who fund research, publish it, and base policy on it. It would be a collegial environment where differences exist only about approach, all participants share the ultimate goals of research, and no ethical problems exist. These utopian goals include a greater understanding of individual and group processes in a given discipline, with the potential to apply these findings to improve both individual lives and society collectively. Researchers would design their studies for the sole purpose of sharing information to better understand the issues at hand and distribute the knowledge widely so that it can improve practice.

The shadow of research design comes as a series of dispositions and paradoxes when the person designing research must make decisions for which the search for disciplinary truth

provides no direction. A researcher has some control, but not complete control, over deciding what research to conduct. A researcher has limited control, or no control, over how research is funded, how it is evaluated, and how it is used. The shadow of research appears when one confronts the lack of creativity in research and psychological barriers to the free flow of ideas. It occurs when a researcher encounters difficulties related to the disciplinary research environment and the primary and secondary social environments associated with the research.

The Loss of Creativity

If the world were static, then creativity would not be necessary; what worked in the past would continue to work in the future. In a dynamic social world existing in a turbulent ecology, the generation of new ideas is necessary for survival. In the natural world, mutation is a random process, and selection occurs where the fit to the natural environment of the new form has advantages over existing forms. In the social world, creativity is the source of variation and must be present before selection can take place. Without creativity in identifying problems to be addressed or methods to be used, a field of study would atrophy.

If research has a core more important than anything else, it is creativity. Without creativity, researchers would only repeat themselves. Without creativity, the questions we pose, the starting place for research design, would be endlessly repetitive. Creativity allows the clientele of researchers—be they the public, practitioners, or other researchers—to bring new ideas into their intellectual or practical lives. They can agree or disagree with each other's findings. They can find fault in their methodologies. But the new ideas remain as work to be examined, understood, enacted, selected, and retained for use (Weick, 1979).

Getzels and Csikszentmihalyi (1976) describe problem finding as being at the heart of the creative process. Educational researchers who invent the best problems have the greatest chance of contributing to their fields. A good problem implies the way in which it should be studied. A superb methodology will not make up for a poor research problem. Structured processes for becoming more creative that emphasize steps to be followed have been identified (Parnes, 1992). However, the *content* of the steps is not well understood. If it were, then everyone would be creative and have plenty of excellent problems around which to design research.

To be creative, as opposed to simply novel, a researcher should be well-versed in substantive knowledge of the topic and the limitations of the chosen methodology. Creativity has at its foundation a sense of play—of suspending normal constraints so as to see new patterns, possibilities, or connections. Play is usually an "idea non grata" in a workaholic environment, although the hermeneutical philosopher Godamer considered play to be an expression of great seriousness (Neill & Ridley, 1995). Play is characterized by just those things that are likely to lead a researcher into creative work, including taking risks, testing new ideas in safety, avoiding rigidity, and suspending judgment (Schwartzman, 1978).

A risk-averse, judgmental, assessment-oriented environment focused on short-term gains will have a negative effect on creativity. If proposals are assessed by published criteria, then how can new projects that do not fit established criteria be funded? We live in a judgment-rich environment, where we have been socialized for years into viewing work as something that will be graded. Peer reviews, editorial reviews, administrative reviews, and granting agency reviews occur regularly. Faculty work can be assessed on an annual basis, with the expectation of products in hand making the time frame for completing work within a year or less. In graduate schools, students are steered out of creative projects because such projects are too risky. It is unfortunate when research is designed not out of the possibility of success but rather out of the fear of failure. In the process, creativity—researchers' best friend and asset—is shunted to the rear. Academic reproduction (Bourdieu, 1984/1988; Bourdieu & Passeron, 1990) ensures reproduction, not evolution. Creativity suffers in the current context of conducting research, and producing valuable new understandings is ever more difficult.

Fear and the Researcher's Ego

There are a number of personal factors that affect research design. Morgan (1997) describes "psychic prisons" as a metaphor for the ways in which our imaginations become trapped. Whatever neuroses we have can emerge in the

task of developing research. Researchers can fixate on certain ideas, repress others, idealize states, or project their own views on the data. A frequently occurring form of projection occurs when the conclusions of a research study are not connected to the data. Researchers project their beliefs onto the data, concluding what they wanted to conclude before they began conducting the study.

Fear has an unacknowledged influence on research design that manifests itself in a variety of ways. The first is internal censorship. Certain topics and methods are never given serious consideration because to do so would be to invite trouble, at least in the minds of the researchers. For example, during the 1970s, many people did not consider qualitative research to be an appropriate form of educational research. Fearing rejection by colleagues, granting agencies, advisers, or editors, researchers steered themselves away from the use of qualitative research. It was not surprising that much of the emphasis of Lincoln and Guba's (1985) *Naturalistic Inquiry* is a justification for, and not an explanation of, this kind of study. Researchers engaged in self-censorship by avoiding Black studies, women's studies, GLBT (gay, lesbian, bisexual, transgendered) studies, and the study of emotional aspects of organizational behavior (Fineman, 2000).

Fear also underlies what has been called the "imposter syndrome" (Harvey & Katz, 1985), where researchers might fear that they are fakes. This problem can show up in an obsessive need to review the literature because a researcher "doesn't know enough yet." A researcher might fear not being professional enough or not being thorough enough and might analyze data in an endless pattern of trivial changes. This fear is a pathology, not a motivator.

Research can also be conducted in service to the ego, not the discipline, where the researcher is driven by the extrinsic value of research. This drive results in designing research for maximum visibility, regardless of substance. Finding the smallest publishable unit in a data set inflates one's résumé but clutters journals. The ego thrives on high levels of productivity. The discipline thrives on high levels of quality. The current research market defines what is acceptable, and clever marketing may be more important to one's ego than a quiet but long-term contribution to the field.

There is a competitive aspect to designing research. Instead of a "best knowledge for the discipline" model, it involves "I got there first," "I'm right and you're wrong," "I win the argument," "My theory is right," "I got the grant and you didn't," "My university is ranked higher than your university," and the like. These are the concerns of the ego, extrinsic to the creation of knowledge, casting a shadow on research. From the point of view of designing research to discover knowledge, it is bizarre that information is not shared. From the point of view that research is not about improving knowledge but rather is about supporting the ego, making a name for oneself, and providing research overhead to one's institution, it makes perfect sense. The impulse is to design research in order to win some imaginary (or real) competition, not because it is vital to the field.

Disciplinary Norms and Groupthink

The kind of study a researcher can conduct depends on the development of the field. Mature fields, such as the arts and sciences, medicine, and engineering (Parsons & Platt, 1973), have a long tradition of theories and methods that are thought to be appropriate to use when conducting research. Cultural studies, critical theory, and other postmodern approaches have preferred methods that keep other disciplines vital by challenging their traditions. Research norms become institutionalized through accepting papers for professional meetings and publication. A disciplinary language develops, and a kind of parochialism develops in citation patterns: Cite from journals in the field only. Disciplines within education and in the professions have become ever more specialized. Research followed suit and led the way to disciplinary specialization.

Research design reflects this specialization in topic and method. Specialization can have the advantage of accuracy and the disadvantage of triviality. Researchers who venture outside the norms can be transformational if they are lucky or can be ignored or ridiculed if they are not. New ideas are sometimes blocked by the disciplinary equivalent of groupthink. *Groupthink,* first described by Janis (1972), includes many factors that limit creativity and risk taking, including sharing stereotypes that guide the decision, exerting direct pressure on others,

maintaining the illusion of unanimity and invulnerability, and using mind guards to protect the group from negative information.

Groupthink is more than a norm; it is an exclusionary process designed to protect the group from outside influence. Groupthink in research can limit the topics studied and the methodology used. The long period during which editors silenced the voices of women; African Americans; and the gay, lesbian, bisexual, and transgendered community in education is one example. Another currently exists among those in education who support only "scientific research" (National Research Council, 2002). Berliner (2002) suggests that the problem is not one of science but rather one of politics and money. Those who label good research in education as "scientific" are stuck in groupthink, as are those who consider the research method as essential and all else as trivial. When methodology precedes identifying the problem to be studied, groupthink wins and research suffers.

Methodology and Methodological Correctness

At the extreme, the result is "methodological correctness," a term I coin as a play on "political correctness." It is associated with taking oneself very seriously and is related to academic fundamentalism, where skepticism is replaced by dogma. Methodological correctness means that the purpose of research is to optimize methodology. It is an example of goal displacement, where the purpose of research is no longer to find out something important but rather to use method flawlessly. The hegemony of methodologists in determining the value of research has a chilling effect on exploring new approaches to research, on studying topics not studied previously, and on studying topics that do not lend themselves to study using preferred methods.

Institutionalized methodological correctness takes the form of guidelines, where if the guidelines are not followed, the result is funding not being given or results not being taken seriously. The U.S. Department of Education has provided *A User-Friendly Guide,* one that is not "friendly" at all, that can be summarized as follows: The only rigorous evidence that can be used to evaluate an educational intervention comes from research using randomized controlled trials (Institute of Education Sciences, 2003).

Simple solutions are often wrong. Randomization means that individual student differences will not be a factor in the research and that all kinds of students can expect to benefit equally from the program. The results are designed to mask individual differences to see whether the program worked for the majority. It works if the mean of the criterion variable for the treatment group is significantly higher than the mean of the control group. Like randomization, means are designed to mask individual differences. Berliner (2002) makes the point that there is a "ubiquity of interactions" and that a program could have remarkable positive effects on a small segment of the treated population, none of which would be discovered by this research design. A program could benefit gifted students, African American students, girls, athletes, or special needs students in a manner invisible to scientific methods.

Much of the contentiousness about educational research design centers on whether the research is scientific, a desiderata identified by the National Research Council's (NRC) publication of *Scientific Research in Education* (2002). The debate revolves around using scientific methodologies to examine educational issues. The NRC's position is generally supported by some (Feuer, Towne, & Shavelson, 2002; Slavin, 2002) and cautioned against or rejected by others (Berliner, 2002; Erickson & Gutierrez, 2002; Olson, 2004; St.Pierre, 2002). Research design, research funding, and politics are interconnected (Burkhardt & Schoenfeld, 2003). The Obama administration has done much to restore the importance of scientific knowledge in policy making, but one can never assume that such a change is permanent.

A research article, like the tip of an iceberg, contains only a small percentage of the information that the author encountered in the study. Given this situation, research becomes an enactment of the map–territory relationship, that is, the relationship between the object studied and the symbol for that object—the research report (Bateson, 2000). How complete does the symbol need to be to represent some objective reality? Borges (1998), in his story "On Exactitude in Science," provides a fictional example of an empire that was so enamored of mapmaking that the cartographers were encouraged to make maps that were larger and more accurate. In the end, they made a map that was so detailed, it

needed to be exactly the same size as the land it described. As a map, it was perfectly useless.

In this case, greater accuracy and greater methodological correctness diminished utility. Bateson (2000) argues that maps are useful not because they are literal representations but rather because they are in some way analogous to reality. Research provides a map, an analog of reality. If Bateson is right, then it might be more appropriate to design and evaluate research not on the basis of how correct the methodology is or how literally it represents reality, but rather on how useful it is for understanding and acting in our environments.

The Primary Research Audience

Research design is affected by the primary research audience for the study. For doctoral students, the primary audience is their advisors and other members of their research committees. For faculty, the primary audience is journal and publishing house editors and grantors. Refereed journal editors are the gatekeepers of much of the research that is published, which in turn influences what is taught and who is tenured and promoted at schools that value research.

Recognizing this power, a researcher responds to the real or imagined preferences for topic or method of this audience. The obvious way of finding editorial preferences is to read the journal, see what is being published, and use a similar approach to one's own study. Doctoral students would be prudent to read dissertations directed by a prospective dissertation advisor to see what these preferences actually are. This situation begs the question, should these gatekeepers set the research agenda? Editors of research journals usually have been successful researchers in their fields and have published widely. The advisory board that hires an editor increases a journal's prestige by hiring the most prestigious editor it can find. The editor then seeks out other successful researchers in the field and brings them on board. This selection procedure produces a conservative bias: It rewards what has worked in the past.

One model for the editorial process is that reviewers have had long experience in the field and make prudent judgments about what studies will advance educational practice or knowledge. Another model views editorial decisions as being on show because what editors approve is published. The imposter syndrome is ever-present: "How do I, as an editor, make decisions that will make me look like I know what I'm doing?" The ordinary response is risk aversion: "If I don't take chances, I'm least likely to look like an imposter." Editors are likely to reject methodologically flawed research in favor of methodologically correct research. Imaginatively flawed research, research whose consequences are trivial for the discipline or practitioners, can be published if the methods are correct but with predictable disdain from the public (Kezar, 2000). I have heard of no cases where an editor has written an author saying, "The ideas in this article are so compelling that I'm going to publish it even though it contains obvious methodological flaws." Editorial referees work at the pleasure of the editor, and if they are to be retained, they work in line with the editorial vision. Reviewers are often shown the comments of other referees so that they can compare their responses. Feedback provides an implicit pressure to conform.

The upward drift in methodology can be considered paradoxical. To get published, authors use sophisticated methodologies. The newer and more complex the method is, the fewer the people who will be able to evaluate the article, and the fewer the practitioners who will be able to understand the research and judge whether using the results would be beneficial. In attempting to gain the approval of other researchers, a researcher might not care whether an article advances practice in the field. Good research can do both; some publications do neither.

The Secondary Research Audience

It is a desirable state when secondary research audiences—other researchers, practitioners, and the public—are more important than the primary ones. From an altruistic perspective, it is for these audiences that the research is conducted. Research should be designed to be useful to the discipline and to advance theoretical or empirical understanding of what is happening in some area of education. Does this research provide new ideas, new understandings, and new practices that advance the ways in which professionals and practitioners in the field can serve the public good? An affirmative answer would justify the use of public and philanthropic resources in pursuit of educational knowledge. Good research should benefit everybody.

A measure of the value of research is not just acceptability to the editorial process, that is, the merit indicated by its publication; rather, the impact of a piece of research on theory or practice becomes the ultimate measure of its value or worth. Research that does not meet at least minimal methodological acceptability is not published and does not become available to its potential audience. Assuming that it does reach a larger audience, does it affect future research in the field?

A well-designed study should include, at the end of the article, recommendations for future research, but in practice, these recommendations tend to focus on narrow methodological concerns, such as improving a questionnaire and using a different sample. The implicit recommendation for future researchers is that they continue to advance the theoretical orientation of the line of research. A second concluding section should deal with the practical applications of the study. The form of the application is along these lines: "If your educational world is similar to the one in which this study was conducted, here are the things you should do, based on my findings, that would improve educational practice and understanding in your world."

Educational research can be influential not because of its quality but rather because the findings confirm what policy makers already believe. This situation is distressing because it means that an excellent study will affect policy only if policy makers *want* it to affect policy. When two studies are excellent but lead to opposite conclusions, policy makers lose confidence in the research and return to intuition to set policy. The politics of educational research seems to be one of its salient features (Cooper & Randall, 1999).

CONCLUSION

The reporting of research can be viewed as storytelling, as part of a mythic process of identifying who we are. In storytelling, we seek to remember the past, invent the present, and envision the future (Keen & Valley-Fox, 1989). Research can be viewed as a similar process in remembering the past by examining the literature; inventing the present by conducting the study and describing the findings; and envisioning the future where this research influences thought, policy, and practice.

To design research is to make a map, an analogy of what happens in the world. Research design depends on what is being studied and what the researcher wants to find out. The double entendre of "wants to find out" is intentional. The researcher wants to find out something about, say, how to improve literacy rates in rural areas. The researcher also wants to find out that his or her hypothesis is true, for example, that tutoring improves literacy.

The choice of the topic focuses the endeavor. The choice of method limits what can be discovered, emphasizing some possibilities and eliminating others. Each choice involves trade-offs. Each methodology chosen should, if done well, supply some beneficial information. There is one best way in which to find out something extremely simple, such as the mean length of time it takes students to memorize a list of spelling words. As the question addressed becomes broader and more complex, it can be studied using a variety of designs. There is no best way in which to study education; each approach emphasizes some things and is silent on others. Political and research ideologies can drive research or be ignored.

Research could be designed purely on the basis of curiosity if the researcher wants to know something. The methodology is likely to be emergent as the researcher plays with the topic, thinking of it without preconception; delighting in possibility; and creating an ongoing dialogue with the topic, methods, other researchers in the field, the persons being studied, and so on.

Research can also be designed around extrinsic reasons: "How can I make myself famous, promoted, tenured, or rich on the basis of my research?" For that research, the researcher should let money and disciplinary popularity lead the endeavor. For research to affect policy, one should follow the money out of governmental or other granting agencies and heed their guidelines for topics and methods. Research should be designed to meet their expectations using methods they prefer. An effective presentation of the results might demand that they be presented in the most simple or most mystifying forms.

Designing research for altruistic purposes, to benefit humanity, is more complicated, because what benefits one group might not benefit another. Any discovery can have wonderful unanticipated consequences. Basic research has

grand possibilities, but the environment must thrive on patience and failure—on trying many new things that do not work to find the few that do. Such environments are rare. Research designed to solve well-defined problems—applied research—can also benefit humanity. Other applied research is intended to profit the patent holder. Research designed to provide an educational environment that will save humanity should get top billing, but who could agree on what that research would be?

In the near future, methodological correctness will likely maintain its salience. I would expect that humanistic and aesthetic values will be neglected in research in the face of issues of social justice and pragmatism. Capitalistic elements related to the costs of education and the ways in which the education system provides a suitable labor force for the nation's economy will likely be emphasized. Whatever work we do or we neglect, our research must refresh our intellect.

References

Bateson, G. (2000). *Steps to an ecology of mind: Collected essays in anthropology, psychiatry, evolution, and epistemology.* Chicago: University of Chicago Press.

Baumrind, D. (1964). Some thoughts on ethics of research: After reading Milgram's behavioral study of obedience. *American Psychologist, 19,* 421–423.

Bean, J. P. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in Higher Education, 12,* 155–187.

Bean, J. P. (1997, March). *How painting can inform qualitative inquiry.* Paper presented at the meeting of the American Educational Research Association, Chicago.

Berliner, D. (2002). Educational research: The hardest science of all. *Educational Researcher, 31*(8), 18–20.

Biglan, A. (1973). The characteristics of subject matter in different academic areas. *Journal of Applied Psychology, 57,* 195–203.

Borges, J. L. (1998). *Collected fictions* (A. Hurley, Trans.). New York: Penguin Books.

Bourdieu, P. (1988). *Homo academicus* (P. Collier, Trans.). Stanford, CA: Stanford University Press. (Original work published 1984)

Bourdieu, P., & Passeron, J. C. (1990). *Reproduction in education, society, and culture* (R. Nice, Trans.). London: Sage.

Bowen, H. R. (1977). *Investment in learning.* San Francisco: Jossey-Bass.

Braithwaite, R. (1955). *Scientific explanation.* Cambridge, UK: Cambridge University Press.

Burkhardt, H., & Schoenfeld, A. H. (2003). Improving educational research: Toward a more useful, more influential, and better-funded enterprise. *Educational Researcher, 32*(9), 3–14.

Cooper, B., & Randall, E. (1999). *Accuracy or advocacy: The politics of research in education.* Thousand Oaks, CA: Corwin.

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

Creswell, J. W. (2007). *Educational research* (3rd ed.). Thousand Oaks, CA: Sage.

Daniel, L. G. (1996). Kerlinger's research myths. *Practical Assessment, Research, & Evaluation, 5*(4). Available at http://pareonline.net/getvn.asp?v=5&n=4

Erickson, F., & Gutierrez, K. (2002). Culture, rigor, and science in educational research. *Educational Researcher, 31*(8), 21–24.

Feuer, M. J., Towne, L., & Shavelson, R. J. (2002). Scientific culture and educational research. *Educational Researcher, 31*(8), 4–14.

Feyerabend, P. (1993). *Against method* (3rd ed.). New York: Verso.

Fineman, S. (2000). *Emotions in organizations.* Thousand Oaks, CA: Sage.

Gall, M. D., Borg, W., & Gall, J. P. (2002). *Educational research: An introduction* (7th ed.). Boston: Allyn & Bacon.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences.* New York: Basic Books.

Geertz, C. (1973). Thick description: Toward an interpretive theory of culture. In C. Geertz, *The interpretation of cultures* (pp. 3–32). New York: Basic Books.

Getzels, J. W., & Csikszentmihalyi, M. (1976). *The creative vision: A longitudinal study of problem finding in art.* New York: Wiley.

Harvey, J., & Katz, C. (1985). *If I'm so successful, why do I feel like a fake? The imposter phenomenon.* New York: St. Martin's.

Institute of Education Sciences. (2003). *Identifying and implementing educational practices supported by rigorous evidence.* Washington, DC: U.S. Department of Education.

Janis, I. (1972). *Victims of groupthink: A psychological study of foreign-policy decisions and fiascoes.* Boston: Houghton Mifflin.

Jones, J. (1993). *Bad blood: The Tuskegee syphilis experiment* (Rev. ed.). New York: Free Press.

Keen, S., & Valley-Fox, A. (1989). *Your mythic journey.* New York: Putnam.

Kerlinger, F. (1973). *Foundations of behavioral research* (2nd ed.). New York: Holt, Rinehart & Winston.

Kezar, A. (2000). Higher education research at the millennium: Still trees without fruit? *Review of Higher Education, 23,* 443–468.

Krathwohl, D. R. (1988). *How to prepare a research proposal: Guidelines for funding and dissertations in the social and behavioral sciences* (3rd ed.). Syracuse, NY: Syracuse University Press.

Lawrence-Lightfoot, S. (1995). *I've known rivers: Lives of loss and liberation.* New York: Penguin Books.

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

Merton, R. K. (1973). The normative structure of science. In N. W. Storer (Ed.), *The sociology of science* (pp. 267–278). Chicago: University of Chicago Press. (Original work published 1942)

Milgram, S. (1974). *Obedience to authority.* New York: Harper & Row.

Mitroff, I. I., & Kilmann, R. H. (1978). *Methodological approaches to social science.* San Francisco: Jossey-Bass.

Morgan, B. (1997). *Images of organization* (2nd ed.). Thousand Oaks, CA: Sage.

Myers-Briggs, I. (1962). *Manual for the Myers-Briggs Type Indicator.* Princeton, NJ: Educational Testing Service.

National Research Council. (2002). *Scientific research in education* (R. J. Shavelson & L. Towne, Eds.). Committee on Scientific Principles for Education Research. Washington, DC: National Academies Press.

Neill, A., & Ridley, A. (Eds.). (1995). *The philosophy of art.* New York: McGraw-Hill.

Olson, D. R. (2004). The triumph of hope over experience in the search for "what works": A response to Slavin. *Educational Researcher, 33*(1), 24–26.

Parnes, S. J. (1992). *Source book for creative problem solving.* Buffalo, NY: Creative Foundation Press.

Parsons, T., & Platt, G. M. (1973). *The American university.* Cambridge, MA: Harvard University Press.

Schwartzman, H. (1978). *Transformations: The anthropology of children's play.* New York: Plenum.

Schwartzman, H. (1993). *Ethnography in organizations* (Qualitative Research Methods Series, No. 27). Newbury Park, CA: Sage.

Shils, E. (1984). *The academic ethic.* Chicago: University of Chicago Press.

Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher, 31*(7), 15–21.

St.Pierre, E. A. (2002). "Science" rejects postmodernism. *Educational Researcher, 31*(8), 25–27.

Thorngate, W. (1976). In general vs. it all depends: Some comments on the Gergen–Schlenker debate. *Academy of Management Journal, 25,* 185–192.

Weick, K. (1979). *The social psychology of organizing* (2nd ed.). Reading, MA: Addison-Wesley.