Chapter 2
Qualitative Research and Public Policy: The Challenges of Relevance and Trustworthiness

William G. Tierney and Randall F. Clemens

Over the last generation, education scholars have seen a great deal of ferment in methodological circles about the strength of one or another method. Although disagreements often occur within a particular framework, the largest point of contention has been between those who identify as quantitative researchers and those who are qualitative. Oftentimes the arguments have resulted in an ideological standoff where one group has claimed supremacy over another, or researchers have pointed out the strengths of their approach and of consequence, the terminal weakness of the other (Altheide & Johnson, 1994; Rolfe, 2004; Sandelowski, 1986). These arguments have fluctuated in intensity for many years, but of late, scholars have entered into an odd Cold War of sorts.

The disagreements frequently revolved around ontological and epistemological issues. Both sides, qualitative and quantitative, have for the most part moved on with their methodological lives and taken very different routes. Although neither group ultimately could claim victory, each succeeded in gaining adherents to their cause. The supporters of a particular approach, however, have adopted different personas and had varying degrees of influence on differing audiences, such as funders, policymakers, and methodologists.

The federal government's focus on science-based research and standard-based accountability has been at the core of methodological debates and policy decisions (Feuer, Towne, & Shavelson, 2002). The quest for clear principles to guide educational research has permeated nearly every aspect of research, including the discourse of politicians, the criteria for funding, and even classes offered and topics discussed in researcher and practitioner preparation programs. The advocates of both methodologies have publicly and privately engaged in debates at professional conferences and in academic journals (see Feuer, 2006; Moss et al., 2009; St. Pierre, 2006). Among the criticisms voiced by qualitative methodologists is the lack of

W.G. Tierney (✉)
Center for Higher Education Policy Analysis, University of Southern California,
Los Angeles, CA 90089, USA
e-mail: wgtiern@usc.edu

We appreciate the feedback and comments of Ronn Hallett, Karri Holley, Yvonna Lincoln, Laura Perna, Susan Twombly, and Ed St. John

J.C. Smart, M.B. Paulsen (eds.), *Higher Education: Handbook of Theory and Research* 26, DOI 10.1007/978-94-007-0702-3_2,

representation of beliefs from an array of paradigms. For instance, the National Research Council (Towne, Shavelson, & Feuer, 2002) presents a series of scientific, positivist-inspired principles, including the ability of research to "replicate and generalize across settings" (p. 4). The principles are certainly well intentioned; however, the authors prefaced the report with the statement that all scientific research is based on their six guiding principles, and as a result, some qualitative researchers felt slighted when their beliefs were not recognized (Lather, 2004). The repeated amendment of standards by the American Educational Research Association (2006, 2008, 2009) reveals the tension among researchers to develop a more inclusive set of standards.

Quantitative researchers now largely are able to lay claim to the policy and practice arena and make use of hundreds of millions of dollars in governmental and foundation funding. The majority of federal funding with regard to educational research, for example, has a quantitative focus. The structural changes and additions that occurred in the federal government have mainly supported quantitative research. The Institute of Education Science (IES) allows that some qualitative studies might be supported as supplementary research, but what they really desire is quantitative studies, in general, and, in particular, approaches that utilize controlled experiments and random assignment. On the IES website, for instance, they describe their own research: "We collect and analyze statistics on the condition of education, conduct long-term longitudinal studies and surveys, support international assessments, and carry out the National Assessment of Educational Progress, also known as the Nation's Report Card" (US Department of Education, 2010). The entire page displays no mention of qualitative research or words that might be associated with qualitative research. Even the name change from Office of Educational Research and Improvement (OERI) to Institute of Education Science was meant to signal a more rigorous and scientific approach to the study of education ("Education Sciences," 2002). The creation of the What Works Clearinghouse suggests that through an amalgamation of rigorous research, practitioners would be able to make better informed decisions and create policies and practices that would improve educational reform. The assumption of many policymakers and quantitative researchers has been that educational work should more closely follow the medical model of research and be able to produce valid findings that are generalizable (Feuer et al., 2002; Towne et al., 2002; Shavelson, Phillips, & Feuer, 2003; Slavin, 2002).

Insofar as qualitative researchers acknowledged that their work did not generalize or align with traditional definitions of validity, they did not find a place at the policy table in federal discussions of educational reform. The result is that qualitative researchers did not so much move into another pre-existing arena (such as policy) but instead created new environments for their studies. They sought to open up the discourse and create an intellectual space for their work. Journals such as the *International Journal of Qualitative Studies in Education* and *Qualitative Inquiry* and meetings such as the *Congress of Qualitative Inquiry* enabled qualitative scholars to work in emerging domains of qualitative research and receive standard academic rewards—publications, presentations, and with it, tenure and promotion. Indeed, the *Handbook of Qualitative Research* (Denzin &

Lincoln, 1994) has not only gone through several editions but also consistently been one of SAGE Publications' largest sellers. Qualitative researchers, who have created unique spaces for their work, have not attempted to argue using a quantitative perspective but rather have investigated issues such as auto-ethnography, queer theory, experimental writing, performance ethnography, decolonizing and indigenous methodologies, and a host of other creative, progressive methodologies (Altheide & Johnson, 2010). We do not wish to overdraw the distinction. To be sure, some journals have utilized qualitative work that struggles to better define or even resolve a particular problem. However, the majority of efforts have gone toward more emergent domains of inquiry that has less interest in specific problem or policy resolution, while quantitative researchers have moved aggressively into that area.

Occasionally both groups share the same geographic space such as in a school of education and at large conferences such as the American Educational Research Association (AERA); yet, even in these areas we have seen spaces carved out for both groups rather than an inter-relationship with one another. Some quantitative and qualitative researchers, in large part because of their frustration with AERA's perceived bias toward either one or the other approach, have started separate educational research conferences. Educational psychology departments and divisions largely rest with quantitative researchers; the largest Special Interest Group (SIG) in AERA is the qualitative SIG with over 800 members. Teacher education has those who do narrative and others who focus on outcomes research; their work rarely overlaps. Even those who claim to call upon mixed methodologies largely do quantitative studies; as graduate students they focus intensively on quantitative research methods and also learn how to conduct a few interviews and focus groups. Seldom does one see someone undertake a mixed methods study where a qualitative method is the primary design of choice.

Our purpose here is not to contest methodologies of either point of view or convince readers to adopt one stance rather than the other. First, we acknowledge that collapsing many different ideas into two groups is problematic and risks missing the nuances of varied beliefs within each methodology. For our purposes, however, we are writing as if qualitative and quantitative research resembles ideal types. Second, both groups have constructed logically consistent, compelling epistemological standpoints; to enter into yet another discourse that assumes we are likely to persuade individuals of either point of view that they have, in many cases, been intellectually mistaken for the last generation is a fool's errand. We are unlikely to convince those who have been advocates for their position, and we would not move the discussion further in any particularly new direction. We also do not envision that détente is likely any day soon. As noted, both groups have their own conferences, journals, departments, and intellectually consistent viewpoints. There is much that can be said of a well-designed quantitative study, just as a well-crafted auto-ethnography can push the reader in a direction that he or she may not have previously considered.

We consider, instead, the role of qualitative research in policy studies. Such a concern has largely been pushed to the periphery by both groups. On the one

hand, quantitative analysts have discounted the role of qualitative research in terms of providing data that might be generalizable across settings. On the other hand, qualitative researchers have been largely unconcerned with issues of policy reform, or when they have, such work has focused on critiquing processes or outcomes rather than suggesting alternatives or recommendations about various issues. We do not quarrel with either stance. Given the epistemological position of quantitative researchers, we understand their concern with validity and generalization. Similarly, the turn to a critique of research and the development of alternative voices that qualitative researchers have created has been of enormous benefit.

We are nevertheless troubled when critical educational issues are left only to quantitative research. However many strengths an approach to a problem may have, our assumption is that many of today's most pressing policy issues are extraordinarily complex and will benefit from carefully conceived and analyzed studies utilizing multiple methodological approaches. Thus, we view the preponderance of IES-funded quantitative studies as shortsighted; in similar fashion that qualitative researchers have largely vacated the policy field is also problematic. Christensen, Johnson, and Horn (2008) have usefully pointed out:

> Just as researchers in medicine are working to understand disorders by their causes as opposed to their symptoms in order to move toward precision medicine, education research must move toward understanding what works from the perspective of individual students in different circumstances as opposed to what works best on average for groups of students or groups of schools (p. 162).

Such an observation necessitates thoughtful qualitative research that will be useful for informed policy making.

What might be done to stem the tide of dropouts from high school? How might we improve access to college for low-income youth? Why are transfer rates from community colleges to 4-year institutions generally abysmal? What barriers exist to increasing the percentage of women in the sciences and engineering? How do health care and retirement policies impact the working environment at colleges and universities? Answers to such questions are of critical import for the health and well-being not merely of educational organizations but also of society. Quantitative research is necessary but insufficient to answer these compelling and complex questions. The more full-bodied and nuanced an analysis that can be provided the better.

Accordingly, in the following sections, we offer an overview of the two stances and identify the criteria each has developed to ensure that their findings are usable. The purpose of the summary is to outline the landscapes of both methodologies; to undertake qualitative research in the policy arena, one needs to understand the frameworks of others in order to take into consideration how one's own work will be reviewed and critiqued. We then elaborate on pressing public policy issues and consider how qualitative research might be useful. We next turn to a consideration of what criteria might be employed to ensure the trustworthiness of data for those of us who intend to undertake policy-focused qualitative research. We conclude with a discussion of the challenges that exist and conceivable next steps toward such work.

Validity in Quantitative Research

Validity is both a simple term commonly used by men and women every day, and a complex statement that has been debated throughout history by scholars and philosophers. To state that a proposition is valid is to imply that it is logically correct, sound, and understandable. If we undertake a study and claim that we are confident in the validity of our findings, we are stating confidence in our knowledge claims.

When someone says "You have a valid argument," they are harkening back to school day syllogisms: "All men are mortal; Socrates is a man. Therefore, Socrates is mortal." Such an argument has a logic to it that is based on the truth of the premises that lead to the conclusion. The problem, of course, is if one or both statements are false or misinterpreted, then the argument is not valid: "All men are white. Barack Obama is a man. Therefore, Barack Obama is white." False premises lead to a false conclusion. Similarly, the listener may interpret a word or a phrase in different fashion from the speaker: "All men are flawed. Socrates and Barack Obama are men. Socrates and Barack Obama are flawed." The speaker may have intended the comment to mean that all men are mortal and not gods. The listener may have heard the comments as a statement about character flaws. Our understanding of the meaning of words and our interpretation of the ideas in the sentences lead to confusion.

In terms of research, scholars are asking if the study accurately reflects what they are attempting to investigate, which involves issues of external and internal validity. The challenge of quantitative research in part has been to create valid findings that are then generalizable to other situations. By making this statement the assumption is that such an undertaking not only can be attempted but also can be achieved. Its achievement, however, depends in large part upon the strength and elegance of a research design. Researcher bias, unintended effects, temporal matters, and a host of other issues need to be confronted before anyone may claim to have a valid finding.

Internal validity pertains primarily to the rigor of the study and if alternative explanations have been taken into account and discarded with good reason (Campbell, 1957; Campbell & Stanley, 1963). A causal connection between the independent and dependent variable is assumed. Cook and Campbell (1979) defined internal validity as "the approximate validity [the best available approximation of the truth or falsity of a statement] with which we infer that a relationship between two variables is causal or that the absence of a relationship implies the absence of a cause" (p. 37). In other words, researchers seek to understand if the connection between two variables is related in a way that can be proven to such an extent that they can state with certainty that a relationship exists.

External validity largely refers to the extent that the study's findings are generalizable (Campbell, 1957; Campbell & Stanley, 1963). Can the causal relationship that has been found in one study, or set of studies, be generalized across different types of persons, settings, and times? If it is true that all men are mortal, then regardless of location, time, or individual, every man whom we study is mortal. However, if we conduct a study of diabetes, we are likely to find different conditions for different men. Race, class, geography, and time matter. African-American men,

for example, may be susceptible to diabetes in ways that differ from white men. African-American men who are poor may have different threats to diabetes than those African-American men who are wealthy.

The import and assumptions of validity are critical. If the goal of research is to develop findings that offer certainty to the questions under investigation and can be generalized, then particular designs and methods make sense. In other words, if the findings are backed by causal evidence and they accurately represent what has been investigated, then we need to undertake large-scale studies that can enable the research study to be valid. At the most intimate level, for example, a life history of one individual is hardly able to claim any possibility of generalization. Similarly, a series of interviews of assistant professors or early career high-school teachers are unlikely to be convincing insofar as the individuals, their experiences, and their statements about work will vary. The researcher also needs to account for his or her own bias. Optimally, the experiment needs to be designed to the extent that subsequent researchers can perform the same experiment and reach the same results, which leads to a consideration of reliability.

For a project to be valid we also want it to be reliable. Reliability, like validity, has a commonsensical understanding as well as a much more theoretical one. "He's reliable" suggests that the person is dependable, consistent, and in a sense, predictable. If we ask an individual to pick us up at noon and we assume she will, then we are basing our judgment most likely on past practice. The assumption over time is that she will continue to pick us up at noon, day after day. We often make casual assumptions based on past practice: "Graduate students are always nervous about their qualifying exams. You can rely on it." Such a statement presumably derives from the speaker's past involvement with graduate students and makes a prediction that the current graduate student will be nervous when his qualifying exams approach.

The same sort of relationship exists with research. Reliability in this sense is a precondition for validity. Repeated tests that lend the same results create the conditions for reliability. If something is unreliable, then it cannot be valid. The ability for something to be reliable also extends to different researchers conducting the same tests; if only one researcher can produce the conditions for reliability, then they are not replicable. Validity concerns whether what we have chosen to study is an appropriate measure of the study, whereas reliability is looking at if those measures produce the same results over time and researchers. Whereas not all reliable accounts are valid, the assumption, in principle, is that all valid accounts will be reliable.

Validity and Trustworthiness in Qualitative Research

The manner in which we have framed validity in the previous section has long been of concern to qualitative researchers (Altheide & Johnson, 1994; Cho & Trent, 2006; Creswell & Miller, 2000; Eisenhart & Howe, 1992; Lincoln, 2001). From

a quantitative perspective, using their criteria, even carefully designed qualitative research cannot be valid. Although many quantitative researchers have respect for the findings of qualitative work, that respect has little to do with the validity of such research. That is, a historical study may provide great interpretive insight into a particular period or person, but a historian makes no claim to the validity of the research. Similarly, a case study of one site may lend itself to understanding an issue or an organization, but surely no attempt can be made to generalize the findings of such a study.

Many qualitative researchers have focused primarily on philosophical, epistemological, and ideological concerns (see Hammersley, 2009a; 2009b; Smith & Hodkinson, 2009). One problem pertains to the idea of realism. The majority of qualitative researchers reject the assumption that an individual or individuals can attain a direct, unmediated knowledge of the world (Angen, 2000; Garratt & Hodkinson, 1998; Kvale, 1996). From this perspective, humans are always interpretive beings and the world simply does not "exist" irrespective of individuals. An observer-independent account of experience, from a qualitative perspective, is impossible. And if validity is dependent upon that, then validity, too, is impossible.

How is it possible, qualitative researchers ask, to assume that truth is a preexisting condition that individuals and researchers simply discover, as if it is waiting "out there" like buried treasure? Rather, individuals—participants and researchers—co-create the world (Guba & Lincoln, 1994) and that world is bound by the language we use and how we understand it. Those who study an organization's culture or a tribal culture, for example, do not discover that culture as if it was simply in existence and a researcher stumbled upon it. The researcher's meanings in part create the culture, and its meaning is amorphous, always changing. There can be no ahistorical frameworks that researchers employ to judge the validity of a situation because our understanding of the world is not objective. Some individuals will take this critique further and identify the modernist and Enlightenment ideals to which terms such as validity fit as the intellectual straitjacket that has led to the oppression of women and gays and the colonization of oppressed peoples. Lather (1986, 1993), for instance, has sought to unsettle the concept of validity by invoking multiple meanings; as a result, contradictory claims annul the notion of science and allow for new discoveries and understandings. Such a stance contrasts starkly with Hammersley (2006), who resolutely states that "it is misleading to believe that there can be different types of validity. Validity is singular not multiple; it concerns whether the findings or conclusions of a study are true" (p. 44).

It is important to note that not all qualitative researchers completely reject realism. Several scholars suggest that no absolute, objective truth exists, but also believe scientists, through repeated, systematic tests, can approximate truth (Hammersley, 1992, 2009a; Phillips, 1987). Others adopt a pragmatic approach and, depending on the project's requirements, borrow aspects from numerous paradigms (see Cho & Trent, 2006; Miles & Huberman, 1994; Onwuegbuzie & Johnson, 2006; Seale, 1999). Miles and Huberman (1984), for example, respect the myriad epistemological underpinnings of researchers and in order to nurture understanding among

researchers recommend the documentation and presentation of the research analysis process. Certainly, numerous beliefs about validity exist.

Reliability is a different matter. Many qualitative researchers maintain that with careful research, individuals are able to provide an accurate description of an object under study; observations, for example, of a particular event may well have been seen repeatedly over a long period of time. Assume that an observer spent a year watching teacher behavior in a classroom and counted the number of times the teacher called on girls and called on boys and determined that boys were asked questions more than were girls. Such an observation, with additional data, may be deemed reliable by some qualitative researchers. However, no qualitative researcher would make the claim that different researchers will observe, collect, analyze, and interpret data in precisely the same manner. Thus, if validity is rejected on theoretical grounds, reliability is rejected as impossible based on the manner in which qualitative work is conducted, a context-bound process that includes unique interactions between the researcher and the researched.

The point is less that an accurate account is of an independently existing reality irrespective of a particular researcher and more that the account can be deemed plausible by readers based on the descriptions developed by the writer. We will expand on this point in detail in a later section. Such an assumption, however, still rejects the traditional notions of external validity. The interpretation that a researcher develops is dependent upon a finite number of interviews, observations, or cases. Understanding is relative, rather than universal. A more radical interpretation will hold that understanding cannot even exist across interviews or sites—that meaning is entirely a mediated social construction and an individual's understanding of another's statement is always partial and unclear.

Simply because qualitative researchers have varying concerns about validity (and reliability), however, does not mean that criteria for goodness in undertaking studies have been irrelevant. A great deal of work has been done that has tried to think through what constitutes a good qualitative study (e.g., Altheide & Johnson, 2010; Hammersley, 1992; Lincoln, 2001; Rolfe, 2004; Seale, 1999; Smith, 1993). Trustworthiness has become a qualitative way to speak about the rigor of one's research (Guba, 1981; Lincoln, 1995; Lincoln & Guba, 1985). If one does not employ traditional quantitative measures for determining quality, then how might one employ the idea of trustworthiness to determine the worth of a qualitative study? We will elaborate on these issues below, but in large part they turn on four criteria: *credibility*, *transferability*, *dependability*, and *confirmability* (Guba, 1981; Lincoln & Guba, 1985).

Although researchers acknowledge that the quantitative definition of internal validity is impossible, they have developed ideas pertaining to *credibility* (Guba, 1981; Lincoln & Guba, 1985). The question that concerns this criterion is if the researcher has presented data in a manner that is credible to the respondents. A simple example of credibility is if a researcher studying teacher–student behavior claims that a classroom where an observation occurred was crowded, whereas the teacher believed it was not. The researcher's obligation is not only to take into account the

respondent's interpretation but also to provide the reader with enough description to determine whether the room is crowded.

Transferability is the qualitative analog to the quantitative conception of external validity (Guba, 1981; Lincoln & Guba, 1985). Again, the point is not that the research project can be generalized to all similar studies but that some sort of information can be gleaned that will be helpful to subsequent researchers of similar studies. Importantly, transferability occurs as a result of detailed, illustrative description, and the responsibility shifts from the researcher (the sender) to individuals in other settings (the receivers) (Lincoln & Guba, 1985). The reader of the report determines the level of applicability of the research. Such a perspective is opposite to external validity in which the researcher generalizes to other populations. Reliability, although rejected by qualitative researchers as unattainable, has its own criterion for trustworthiness—*dependability*. The concern here is the logic of the process of inquiry—the research design and methods employed. What kinds of data have been collected and over what period of time? Was the research collected over a determined period of time, or was it simply a rushed study where one visit and a few interviews occurred (Ray, 1980)? Finally, *confirmability* asks that the research findings be clearly linked to analysis, data, and the research site. The goal is to enable the reader to see the train of thought of the researcher to determine how he took a piece of data, analyzed it, and then reached a plausible conclusion.

We will discuss trustworthiness in greater detail in a subsequent section, but it is useful to note that these criteria have been criticized (Garratt & Hodkinson, 1998; Mishler, 1990; Rolfe, 2004; Sandelowski, 1993). Quantitative researchers have continued to refine definitions of validity rather than ways to substitute another framework for it. While qualitative researchers have worked to refine the criteria, develop ways to judge it, and also focus more intensively on the idea of rigor, quantitative researchers simply dismissed the ideas. Recall that qualitative critique of quantitatively defined validity is not that it is inappropriate for qualitative research but that it is inappropriate for research as it has been defined. If one begins an argument with a different set of epistemological positions, then there will be little room for agreement on the propositions one develops regardless of how carefully they are constructed. Simply stated, if reality is socially constructed, then issues of validity are difficult, if not impossible. If a pre-existing reality exists prior to a researcher entering a situation, then the potential exists for us to create the conditions for validity. The result has been that quantitative methodologists have spent a great deal of time expanding their repertoire of analyses to make the achievement of validity on a broad range of issues much more possible. Even though many scholars are utilizing new methods that have brought about useful results, the conclusion has been that a researcher's use of trustworthiness may be appropriate for qualitative work but irrelevant for quantitative studies. Further, quantitative scholars argue qualitative work has little utility for the kind of policy-oriented, problem-solving research undertaken by the quantitative community.

Qualitative researchers had a different concern with trustworthiness criteria. The most severe critics saw the idea of trustworthiness as simply a qualitative attempt to adopt quantitative criteria for rigorous research. Those who subscribe to notions of

postmodernism and the like reject the idea that knowledge can entirely be agreed upon or that consensus can be reached about ideas. Criteria such as credibility and dependability may be of use to the researcher; however, the idea that one set of criteria could accurately and consistently result in a final truth is mistaken. Understanding cannot be achieved and to try to do so is a fool's errand at best, and at worst, deceptive. Guba and Lincoln (1989) subsequently developed a different set of criteria based on the idea of authenticity and praxis. They used a different approach compared to trustworthiness criteria. Rather than parallel criteria, which focus on translating quantitative terms to support qualitative research, they founded the criteria on qualitative-specific concerns, such as ontology, epistemology, and axiology. The framework asks how a research project has created action and empowered those who were involved in the study. The respondents' viewpoints become essential and self-understanding is an end in itself. The researcher's role is less that of the disengaged scientist and more that of a community activist trying to foment change. Of consequence, the quantitative community has had little good to say about such a framework in relation to validity. Those qualitative researchers who worked from a critical perspective, in contrast, found much to applaud in the work, albeit acknowledging the criteria had little to do with validity (Morrow, 2005). Postmodern qualitative researchers in general had little interest in analyzing or refining such criteria and instead focused on generating alternative forms of narrative.

The Public Policy Need for Qualitative Research

As a consequence of the movements mentioned above, the qualitative voice in public policy discussions surrounding education has been relatively muted. Those in control of federal, state, and foundation monies have tended to prefer quantitative studies that subscribe to traditional notions of validity; those who conduct and teach qualitative research have largely adopted new purposes and focused more on the critique of traditional notions of research and the development of alternative representational voices. We have a dual concern with this development. First, the role of qualitative research is diminished and researchers end up thinking in a unitary manner about pressing public issues. The point is less that looking at a problem in a particular manner is necessarily mistaken, but that framing an issue from a singular perspective runs the risk of overlooking alternative explanations for why the problem exists or how to think about studying the issue.

Second, public policy is less well informed by the lack of qualitative work. Many of the most pressing issues in education, for example, deal with problems related to children, adolescents, and young adults. Questions pertaining to the efficacy of online learning, effective working conditions for teachers or faculty, and what might be useful ways to deal with students in need of remedial writing and math when they graduate from high school have multiple angles from which they can be studied. To narrow the focus to one particular paradigm runs the risk of reducing complex environments when actually what we need is a fuller understanding of their intricacies. The result is that the voices of those under study are hushed, if not lost. Oftentimes a

compelling argument can be made by the urgency of portrayals that qualitative work can provide. The meaning of research can be heightened by the contextual data that qualitative research can lend.

In subsequent sections, we will argue that qualitative work provides a voice and a face to those individuals whom researchers study; of note, qualitative research has the potential to provide a social urgency to issues that are always voluntary: no one must enact educational reforms. Our focus here is not on issues of implementation, but we are mindful that any research undertaking not only involves identifying a problem, designing a study, analyzing data, and reaching conclusions but also needs to help decision makers reach some sort of conclusion about what actions should be taken. The emphasis on translational research (Bulterman-Bos, 2008; Henig, 2008; Stokes, 1997) over the last decade underscores the rising importance on being able to produce texts that translate academic scholarship into texts that are readable and usable by policymakers and the general public. Qualitative research can help provide understanding on issues in ways that quantitative research cannot, just as quantitative research is able to resolve issues that qualitative research cannot. Accordingly, we first summarize five areas where qualitative work will not be useful, and then raise five areas where the work of a qualitative researcher will be useful.

Qualitative Research Cannot

Be generalized. Although we have suggested that qualitative researchers often take a more philosophical stance with regard to questions about the nature of research whereas quantitative researchers are often more pragmatic, one cannot escape basic notions of research if we are to understand one another's assumptions and epistemological framework. Just as the concept of validity has a simple meaning as well as deeper theoretical notions, so too does generalization. "All faculty are pessimistic" may be a common perception, but it may not have the research strength to claim external validity. One might infer that someone has used inductive logic to reach his or her conclusion and that the conclusion has been reached after numerous interactions with multiple professors over a significant period of time. However, if someone said "Professor Jones is pessimistic, therefore all professors are pessimistic," we would likely assume that the statement is false or based on inadequate data.

The same point could be made between quantitative and qualitative research. Studies that have a small sample size are not generalizable. We concur with such an opinion when the discussion pertains to public policy research. We have been troubled when a qualitative study has a very small sample size and then proceeds to make recommendations. This is the sort of conjecture that has enabled some researchers to reject qualitative research in large part because the study is making a claim that cannot be supported. The focus on only a few subjects or cases, however, makes it possible to investigate in detail the relationship of a variety of variables to one another. Thus, generalization is certainly one possible end in research, but it ought not to be the only end, and it cannot be an end in qualitative work. An alternative possibility might be that the contextuality of knowledge is useful.

We appreciate that some critics (Gomm, Hammersley, & Foster, 2000; Kvale, 1994; Ruddin, 2006) suggest that generalizable knowledge is possible from a single case, or a handful of cases. But they are thinking less in terms of public policy research and more about theoretical advances that might be made. A useful case study can add to our theoretical knowledge, and it is possible that one may attempt, as Ruddin (2006) suggests, "hypothetico-deductive theorizing" (p. 800) as opposed to generalizations based on empirical data and statistical inference. However, policy research generally is after large-scale answers to specific questions: What is, for example, the optimal classroom size for teaching a class? In some instances, we have become less rigorous with what we are researching in our studies which has led to inappropriate inferential generalizations rather than fine-grained findings based on in-depth research. We are in agreement, then, with Gomm et al. (2000), who state that general relevance can be gained from case studies and the like but that such researchers "often are not very clear about the basis on which are claiming general relevance of their findings" (p. 111) and, of consequence, cannot generalize.

Although we sympathize with Janet Schofield's (2000) concern about the importance of generalization, we remain unconvinced that it is possible "to achieve greater generalizability of qualitative research to situations of interest" (p. 88). If one holds to traditional interpretations of generalizability, then adding cases, or doing them more thoroughly, will not enable the researcher to achieve external validity. Schofield, it seems to us, is trying to extend qualitative research's reach beyond its grasp. Suggestions such as conducting multi-site studies or carefully choosing the cases for one's sample, or thinking about trends in the subject under study, impact research design and data collection and analysis, but they do little to enhance the ability of the researcher to achieve generalizations with his or her research. Similarly, reconceptualizing generalization, such as Donmoyer (2000) and others have suggested, is a useful thought experiment but it does not enable qualitative researchers to function particularly well in the policy arena. The point for us is less that generalization is a mistaken concept, although we understand the epistemological concerns, and more with how we might enable qualitative research to be undertaken that is useful, but not tethered to a concept that is unachievable.

Be objective. The assumption of the quantitative paradigm is that the researcher conducting the experiment will be entirely objective. In one sense, human judgment has been removed from the undertaking and the researcher is assured of accurately representing the test conditions such that reality gets represented without interference from the researcher. The point, of course, is not that humans are replaced by automatons, or that multiple decisions are made in the design and analysis of the undertaking. However, when one undertakes an experiment and analyzes data, the interpretation and the conclusion are definite.

The ability to be objective presumes that independence exists between the subject (researcher) and the object (the experiment) (Burawoy, 1998). Objectivity implies that the subject has accurately represented objective reality. Personal knowledge is irrelevant, or even harmful, in the quest for objectivity. Ultimately, two ideas drive

objectivity. The first is that reality is distinct from the individual. Reality as a social construction is rejected. The second is that without objectivity a study is open to charges of researcher bias, and bias is wrong.

Given the strictures of such a framework, pure objectivity is impossible for a qualitative researcher. Many feminist and postmodern researchers will reject even the notion of researcher independence (Garratt & Hodkinson, 1998; Lather, 1993). But any qualitative researcher will acknowledge that perfect objectivity—even if it were to exist—is impossible in qualitative research. The researcher's hand is always in the design of a study—from the inception to the writing and publication. One researcher may develop different questions and therefore have different responses from a research subject. Two individuals may interpret a transcript differently because of different perceptions. Even the meaning of words and phrases is open to multiple interpretations because researchers cannot entirely understand every single word or inflection of a respondent. Individuals also have different writing styles and how they present the data will undoubtedly vary. Some individuals, for example, will use the passive voice, whereas others will utilize the active voice. Even so inconsequential, a matter as narrative voice could play a role in how the text gets read and interpreted; quantitative researchers, in comparison, have far fewer authorial concerns. Qualitative researchers cannot dismiss these challenges based on their epistemological assumptions.

At the same time, the more egregious examples of bias in qualitative research in public policy can be eschewed. A researcher can control for bias in numerous ways that we will discuss below. Questions can be asked in a way that is open-ended rather than in a manner that solicits a desired response. Researchers have developed ways to check for bias in interpretation of transcripts and the like, but ultimately, objectivity is another concept that qualitative researchers are unable to achieve no matter how many procedures get developed.

Be structured akin to experimental conditions. Similarly, very few qualitative researchers think of their work as an experiment, or that the undertaking even mimics experimental conditions. Humans act differently from moment to moment; the observation of teacher behavior or interviews of early career faculty may be impacted by the events of the day, but an assumption of qualitative research is that such impacts are useful data. The result is that any attempt at creating pristine experimental conditions violates the basic assumptions of qualitative work. Good public policy, however, calls for order and regimentation so that the conditions for objectivity and internal validity will be met.

The result is that those who subscribe to the quantitative paradigm often disdain what they see as disorder in qualitative research designs. Or rather, the problem is less that the work is disorganized, but that it holds little perceived utility for solving complex policy issues. One response might be that qualitative work's goal begins with the assumption that one's methods necessarily involve less organization and a greater attempt at understanding an environment or people holistically. While one may understand that attempt, the fact remains that the work will not be experimental and fails to meet typical scientific standards that have been established as good practice.

Be replicated. If experimental conditions cannot be created, then that which is being studied cannot be replicated. Again, replicability is a fundamental concept of experimental science. How can one make a statement about causality if the object of investigation has not been replicated? Simply because one group at one point in time acted in a particular manner or responded to a particular intervention does not provide any certainty that the next time such an act is attempted that it will be successful or meet with the same response. Replication is essential for scientific work. Medicine is based in part on its ability to conduct tasks that have been verified because of replicable experiments. The counter-argument that oftentimes future findings refute previous ones is certainly true. But there is also enough evidence that as a society we base laws and our ability to act in part on replicability. If someone has a certain blood–alcohol level, then he or she should not drive. When a patient has surgery, the instruments should be sterilized. If a person is choking, the Heimlich maneuver should be attempted. Of course there are exceptions to the rule, but we know of few individuals who would disagree with these three examples.

Qualitative researchers cannot generate similar rules because they do not have the conditions for experiments and cannot replicate what they have found time and again. The result is that qualitative research lacks yet another key concept that affords science its utility and worth for public policy. Although conditions may be structured in a manner such that the researchers see similar activities, they will not be precise in a manner akin to what exists in the laboratory. That is, two researchers at different points in time may have identical protocols and may try to study similar students in similar classrooms in order to determine if there is a gender bias in science. A person's experiences differ, and the variability of temporal conditions may differ such that to claim that these tests are replicable would be a mistake. While temporal conditions also change in the laboratory, the ability to control for change is much greater and possible than it is with qualitative research.

Be voiceless. Finally, quantitative researchers have a history of employing the passive voice in their narratives as a technique to demonstrate distance from the text: "The findings were analyzed ...." The tone that an author wants to convey is consistent with the imperative to erase the imprint of a personality. Obviously, an author's name is on a text, but if the research is objective, reliable, generalizable, and experimental, then the author's specific role in the research has been at arm's length of the findings. Of course, the individual developed the design, analyzed the data, and wrote the text, but the complete focus needs to be on the data, not the author.

Qualitative research is almost never voiceless. Authors not only disdain the passive voice but often write in the first person singular: "I interviewed ...." The assumption of the qualitative author is that he or she obviously played a role in the creation of the data and development of the text, so the text should accurately reflect that role by inserting the author more actively in the narrative. A skeptic may argue that the author weakens the significance of the research by unnecessarily shifting focus from the data to the researcher; we argue, instead, that voice can strengthen the presentation of qualitative research and add useful, vivid dimensions

that allow for a better understanding of the research design and data collection and analysis, which are essential to policy decisions. In qualitative research, voice matters. Authors of quantitative texts consciously try not to use adjectives or make the data presentation and analysis read like a novel; qualitative authors struggle to create portraits that are interesting and frequently read like fiction. Style, tempo, pace, tone, and texture are all concepts that qualitative researchers think about as they develop their texts. Because they think of such ideas, they yet again fail to live up to scientific notions of good research.

The concepts we have outlined are inter-related and all fall under the earlier observations we had made about quantitative research. Taken together they form a schema for the conduct of research, and conceivably as an indictment of qualitative work. If qualitative research cannot accomplish these tasks, then what is the worth of qualitative work? We do not wish to state the question too strongly for certainly no one disdains the anthropological and sociological traditions that have promulgated qualitative work for over a century. However, many might ask if research cannot achieve these various tasks, then certainly their utility for policy-oriented research is limited, if at all. We disagree. In the next section, we outline what qualitative research can do and then turn to how such concepts might be verified.

Qualitative Research Can

Provide context. One of the strengths of qualitative research is that well-crafted studies can provide meaning to otherwise ambiguous observations and statements. Statements, for example, that students find homework "boring" may well be true, but the understanding of such a statement depends on the sort of students who are being discussed, and more importantly, how students define homework and boredom. Similarly, although there has been a great deal of research about leadership, its meaning and implications have defied researchers since Thomas Carlyle (1897) first wrote his treatise on "great men." Suggestions, for example, that good managers "walk around" have appeared as a recipe that neophytes have followed and then subsequently have failed as leaders. What went wrong, they ask? They walked around and did other actions that manuals have prescribed only to find that the recommendations were for naught. Qualitative researchers have the potential of providing specific contexts so that a more nuanced understanding might be developed to abstract terms.

What is the utility, one might ask, of a contextual understanding of one site, however well developed and carefully drawn? Why would public policy experts want a portrait of one or a handful of cases and how would they be used? Such questions assume that the kind of knowledge an individual needs to make a policy decision is large-scale databases that outline trends and composite findings. Policy analysts need to know, for example, if advanced placement classes enable students to pass the AP exam and gain college credit. But they also need to know why students do not pass the exam, and that sort of information can be gleaned by way of

interviews, observations, and focus groups. In other words, what occurs between the formulation and evaluation of a policy, between the decision to add AP classes to a master schedule and the performance of the students, is oftentimes complex and nuanced. In such cases, multiple perspectives, i.e., a mix of quantitative and qualitative methodologies, to present social phenomena are not only useful but also essential.

Qualitative research's history in anthropology has allowed individuals who have no understanding of cultures different from their own an understanding of the lives and practices of different peoples (Malinowski, 1922; Mead, 1928). Those readers, consumers, and decision makers of educational articles may only know about a particular practice or group indirectly or through their own experiences. Well-crafted qualitative studies of homeless youth, for example, enable individuals who have no direct experience of the daily lives of such individuals what they experience and how they make sense of these experiences (Tierney & Hallett, 2009). The point is not only that such portraits have the potential of creating a moral urgency to such issues but also that they lend understanding in ways that quantitative research cannot do, or not do as well.

Provide understanding. The assumptions any researcher makes are critical to one's understanding of the findings that he or she puts forward. Obviously biased survey questions—"Do you support the federal deficit spiraling out of control?"—are as inappropriate as similarly framed interview questions: "Why do you think the Democrats can't control spending?" If readers see the questions, then they may have a better understanding of the answers and make use of them or not. With qualitative research, however, the potential exists to frame a text in a manner that ensures that the readers understand not merely how the research design has been framed but also the epistemological understandings of the researchers.

Why should policy analysts care about admitted abstractions such as epistemology? We were critical earlier of qualitative research that obsesses over the standpoint of the author and we certainly do not suggest that sort of navel gazing in policy-related work. However, the data that get developed in qualitative research in part depend upon the background of the researcher. If readers are to have faith in the findings of the researcher, then they need more background in qualitative research than what is provided by quantitative researchers. The ability of the author to lend judicious insight into his or her background is not merely useful, then, but also a way to develop credibility for the research. Such a fundamental observation has often gotten obscured by the assumption that the researcher is simply interested in self-promotion. When a crisis erupts in another country or the legislature debates a particular issue, the media frequently turns to experts. Listeners will have more faith in the comments of someone who has spent years in Afghanistan and speaks Pashtu, as opposed to someone who spent a weekend there and speaks only English. The point is not that listeners and readers must believe the expert, but background about the individual enables the reader a better understanding about how to interpret the data.

The same assertion can be made with regard to qualitative research. The background of the researcher and his or her familiarity with the material is essential in

permitting the reader to make sense of the text. Ultimately, the reader-cum-decision maker is the one who makes a determination about what to do or what not to do. With qualitative research the ability of providing that individual with information about how one thinks about the particular question strengthens the ability of the individual to come to a better informed decision. The possibility exists, for example, that the decision maker may not have thought about the problem in a manner akin to the qualitative researcher—or may disagree with the assumptions of the researcher. Regardless, policy analysts need thoughtful interpretations of the problems that confront them and qualitative work helps frame questions in ways that will be different from those of their quantitative colleagues.

Provide depth. As we noted, quantitative research relies on the ability to have conditions that are controlled and sterile. The use of the word "contamination" is purposeful; one does not want the data to be contaminated by effects that have not been considered. The strength of qualitative research is in its capacity to allow the reader to understand the situation not so that the next study will be precisely like the last but to think about how the particular study might inform future ones or different situations.

Educational organizations are in a constant state of flux; students arrive with different expectations; not only different faculty teach and work in a manner that differs from one another but also the same individuals have different ways of looking at their work from year to year. The capability of qualitative researchers to document what they have done enables a reader to learn and reflect on his or her own situations, even if the case study is an *n* of one or the interviews only amount to 50 or 60. To be sure, the design has to be elegant, thoughtful, and thorough, but the assumption that policy analysts have little to learn from a well-designed case study belies the fact that individuals learn from difference. When researchers combine contextual information with a framework that demonstrates the assumptions of the researcher and the manner in which the project was carried out, the reader will be provided with enough information to see whether the findings are useful in helping him or her reach a decision about how to proceed and what kinds of decisions need to be reached. Depth provides richness to contexts so that readers do not merely read about different contexts but are able to understand how those contexts differ from one another.

In daily life, if individuals only called upon survey data to reach decisions, nothing would ever get done. Qualitative research cannot be whimsical, such that decisions are reached on extraneous data, but to assume that a series of interviews or observations or case studies will not heighten policymakers' understanding of the problem at hand is to overlook the importance of reflection for decision making. Qualitative research, then, affords individuals the opportunity to utilize data that most likely are unavailable from quantitative work regardless of how elegant or thoughtful the instruments that have been used.

Provide comparison. The sine qua non of qualitative research is its ability for reflection and comparison (Corbin & Strauss, 1990; Glaser & Strauss, 1967). Comparison may occur by the research itself when two or more case studies, life histories, or ethnographies are developed. Any number of useful case studies has

been developed, for example, about decision making in higher education or how public boards of trustees operate. The ability to hold the interview protocols constant and to have conducted them for a similar purpose at a similar point in time provides a useful comparative perspective for the reader that cautions against simplistic recommendations. The research is useful in enabling the reader to reflect not on causal relationships between and among variables but on the complexity that exists in educational organizations and how to think about those complexities.

Indeed, even with a single case study the potential for comparison exists. At the center of qualitative research, then, is an ideology of reflexivity. Well-crafted studies provide the reader an opportunity to reflect on his or own situations in regard to similarities and differences. An analysis of how students make use of technology may provoke a sense of how different teenagers are from the reader and help the individual to think of technology in a different manner. Or, the reader may have a similar interpretation of technology but reach a different conclusion from the author.

We pointed out earlier that both research perspectives have different worldviews and we have maintained that both are useful. Well-designed, large-scale generalizable studies that depend upon causality can aid policy analysts in thinking about specific ways to improve learning or enhance educational environments. Similarly, however, qualitative research has the potential to enable decision makers to recognize the complexity of issues that confront them and recognize their own perspectives. Different points of view offer individuals the opportunity to reflect on the issues under study in ways that quantitative work cannot.

Provide voice. A final point that relates to the other claims we have made here pertains to the readability of the text. Just as quantitative research strives to have prose that is cool, disengaged, and stripped down such that the text is written in the most neutral of styles to strengthen its findings, so too does qualitative research have its own particular style. Qualitative research has to be readable, compelling, and well written (Caulley, 2008). Readers need to feel engaged with the text in a way that is unnecessary and/or impossible with quantitative work. The point is not that authors need to utilize a particular voice, or that one particular style should be preferred over another. Rather, our observation is that qualitative work can provide an understanding of the problems and people under investigation frequently through well-crafted work that creates some sort of socio-emotional bond with the reader. Studies of urban youth, for example, afford the reader the chance to see adolescents in ways that might provoke understandings that are impossible to attain through surveys or statistical techniques. The longitudinal study by Bourgois and Schonberg (2009) of homeless men and women who were addicted to drugs was compelling not simply because they had a superb research design but also because they were able to portray the situations of these individuals in ways that provoked in readers a sense that they could see the lives of these individuals, and in a way, cared about them.

Some of the most compelling social science research of the last century has been qualitative texts written by authors who had a deft sense of narrative. Franz Boas (1964) and Bronislaw Malinowski (1922), for example, not only advanced an understanding of culture, symbolism, and ritual but also enabled readers to understand

lives vastly different from their own. Michael Harrington's *The Other America* (1997) created a sense of urgency in the country about those who lived in poverty and helped create a raft of public legislation. Elliott Liebow's *Tell Them Who I Am* (1995) brought a face to those homeless individuals who are often faceless and voiceless and enabled legislators to come to grips with the increasing problem of homelessness. And of course, the wealth of books by Jonathan Kozol (1991, 2005) have demonstrated how inequity works in America's schools and has forced the reader and public policymakers to ask if it is impossible to create a better educational environment for children. All of these texts accomplish a different task from quantitative research and they have been of enormous benefit to policymakers attempting to develop policies that address the problems raised in their books.

Evolving Criteria for Conducting Policy-Oriented Qualitative Research

If one agrees with the purposes of and need for qualitative research, then a fair question is to ask how researchers might ensure that the conduct of the work is of high quality and utility. In what follows we offer a rubric that delineates those criteria based on previous work and recent studies by groups such as the National Science Foundation (Lamont & White, 2005; Ragin, Nagel, & White, 2004).

Description. The strength of qualitative research is that the analysis should be able to provide a text that provides a great deal of data about the topic under investigation. Description is fundamental to understanding social action; it offers the empirical bases for the judgments readers make about a text and situation. One of the shortcomings of some qualitative research is that only a handful of interviews or focus groups occur and the resulting analysis provides the reader with the thinnest of descriptions. Such work ends up as scholarly conjecture rather than a convincing argument. Clifford Geertz's famous dictum to gather "thick description" speaks to one of the strengths of this form of research (1973). If policymakers are struggling to create policies about topics which they have little first-hand experience, the ability to describe various situations as convincingly as possible will contribute to their understanding.

The strength of such an approach is that it has the potential of being complementary to quantitative studies that lack the ability for explanatory detail of specific issues. Obviously, different sorts of descriptions are possible. A thematic description, for example, will focus on a specific theme (such as socialization in academe) and a case description may involve a particular problem (such as strategic planning) across several sites. Group descriptions may look at a specific class or type of people (such as graduate students) and cultural descriptions focus on a particular group's cultural mores to understand what is shared across individuals.

One manner to judge the quality of a description is by its ability to convince the reader of the author's conclusions. The most frequent problem that arises for an author is that the conclusions go well beyond the data that are presented or that

the data are inadequate to arrive at any conclusion. On the one hand, a well-crafted description of a particular problem may enable the reader to understand a particular situation, but the author then jumps to conclusions that may or may not be warranted, but the data do not support the findings. On the other hand, the author has provided such a small amount of data that the reader is unable to make sense of the text. Both problems highlight the importance of retaining a narrow focus on the research project and providing as much detail as possible. The ability to be judicious in the scope of a study and to employ well-crafted, convincing data are standards for such a criteria.

Objectivity. Although the objectivity proposed on the part of the quantitative paradigm is impossible, one key criterion for policy-related qualitative research remains objectivity in both design and presentation. Readers need to understand the standpoint of the author on the particular issue, how the research design has been developed, and how the data were collected and analyzed. Ultimately, the reader—as with all texts—determines whether a text is biased or not. Such an observation is often difficult for the neophyte researcher. To insert one's standpoint into a text can be distracting and of little use. For example, to state that the author is vegetarian in a study about teaching is of no import; however, in a study about the ways chickens are slaughtered, a reader most likely should know of the author's dietary habits. Protocols provide readers with a sense of the manner in which the author approached the study and enable individuals to ascertain the even-handedness of the author. "I bet you have encountered a great deal of bias en-route to tenure" is different from a question that asks "Tell me about the bias you have encountered en-route to tenure" or "Tell me about the experiences you have had en-route to tenure." Each question prompts a respondent in a different way. Although the first two questions might be useful for some sorts of studies, they tend to tilt the responses in a way that would be unhelpful for policy-related work.

Similarly, data collection and analysis have advanced a great deal from simply writing down notes and then making file cards to follow themes. The confidence of a reader in the worthiness of a text will be buttressed by providing background on the manner in which the author collected and analyzed the data. Transparency, in this way, becomes essential. Whether or not the researcher conducted enough interviews to make the interpretation credible is not for the researcher to decide post hoc. Similarly, contradictory data should not be concealed for the purposes of presenting a coherent, believable argument. During the design and execution of a study, the researcher has the ability to interview more participants and explore contradictions. During the presentation stage, the author has the duty to present the research as it was not as it should be. In addition, all researchers/authors have idiosyncratic ways of approaching their research topic and ensuring credibility. Glaser and Strauss (1967) suggest constant comparisons of conceptual categories to limit bias. LeCompte and Goetz (1982) suggest frequent participant reaction to data collection and interpretation. Idiosyncrasies, however, do not mean sloppiness or inevitable bias. Indeed, the background of a researcher may contribute to a reader's understanding of the research findings and recommendations. To the extent that a text offers readers background on the project's scope and purpose, the likelihood rises

in their confidence about the findings. When authors provide no discussion about the research and instead focus entirely on the data, they reduce the likelihood for objectivity and belief in the text. Thus, a significant discussion of research method, design, and researcher standpoint, focus, and purpose are examples of standards for these criteria.

Interpretation. Any intellectual problem has a multitude of interpretations. Although quantitative work often moves the reader toward a generalizable conclusion that forecloses other alternatives, qualitative research does the opposite. A shortcoming of some work is that authors present findings in deterministic fashion as if there is only one interpretation. Authors want readers to believe that the finding they have reached is based on data and then present confirmatory evidence. When one thinks about how the "real world" actually functions, however, there are almost always a variety of possible interpretations. In a study of the organizational culture of a college, for example, a great many interviewees may express confidence in the leadership of the institution; we doubt that everyone will. Similarly, in a study of a college access program, many students may talk about their desire to go to college and their inability to understand how to pay for it; but again, to assume that all humans think and speak alike belies how groups and societies function.

Even when data are presented in a manner that adequately mirrors the social reality of the setting, the findings themselves will be open to interpretation. Far too often research findings get presented as if the data logically lead to a causal outcome, which defies the underpinnings of qualitative research. Instead, the author needs to suggest possible interpretations for the data and present his or her conclusion about why one particular viewpoint is better than others. Consider, for example, an academically underachieving student. A range of interpretations may explain the student's poor grades—i.e., engagement, peer group, home conditions. The author's duty is to present the most salient data to justify his interpretations, but also introduce alternative perspectives. One indicator of the strength of qualitative research is in the author's ability to demonstrate an understanding of multiple interpretations to data; in effect, the researcher needs to be able to anticipate the interpretations the readers might have and then provide an understanding about why one interpretation is the most plausible. If the data get presented in a manner that enables only one interpretation, then the author has failed in helping policymakers see various ways to think about complex issues. A willingness to offer more than one possible interpretation is a standard for this criteria.

Transferability. The strength of qualitative work is in the ability of a text to make connections to other settings. Transferability is not generalizability. However, no qualitative social science research would be very successful if the work had little or no connection to any other setting. Indeed, literature, philosophy, and history have a similar function. William Shakespeare's plays have lasted for these last several centuries not because all of us live lives similar to Hamlet or Othello but because they enable us as readers and playgoers to think about the human condition. And a reader's thinking at times may be empathic about what is read, but it does not need to be. What is necessary is a sense that the text pertains to something larger than the specific context under consideration.

For as long as anthropologists have been writing about other cultures, the assumption has been that readers pick up a text not simply to read about something, or someone, that they are ignorant about but to learn about their situations and their own lives. Such learning does not suggest that the way the subject under study is the way that all people should, or do, live but that we have something to learn when we study situations different from our own. To suggest otherwise closes off learning as if one only reaches conclusions from one's own experiences or data that applies to all across settings and contexts. Transferability, then, is a criterion based on the assumption that policymakers are able to read a qualitative text not so that they might understand the world in a teacup, so to speak, but so that they might think through how a world, fully described, gives them added information to reach an informed decision.

We are not suggesting, then, that transferability is a watered-down version of generalizability. Rather, transferability refers here to the ability of an author to evoke in readers an understanding of the research project in a manner that enhances understanding and presumably provokes questions regarding similarities and differences. Good qualitative research is meant to provoke conversations and debate rather than proffer a conclusion served as a fait accompli. The manner in which one accomplishes transferability has less to do with the data that are collected or the way they are analyzed and more to do with the ability to write a narrative and develop research questions that are specific yet still focus on critical questions. "It *is* [the researcher's] responsibility" states Lincoln and Guba (1985), "to provide the data base that makes transferability judgments possible on the part of potential appliers" (p. 316). Any text that is so narrowly defined that it has no implications for anyone other than the specific situation being studied has failed in a key function for policy-oriented qualitative work. Why ask someone to read a text, however elegantly written or cleverly analyzed, if it affords no opportunities for the individual to learn anything about the problems he or she is facing? The standard for this criterion, then, is the ability of the author to offer lessons that will be of use to the reader in his or her situations.

Authenticity. Textual responsibility pertains not only to the reader who is a policymaker but also to those who were interviewed and involved in the study as research subjects. We referred earlier to the importance of interpretation. In order to understand a situation, one needs not only to provide thick description and undertake long-term research but also to see if the findings appear authentic to those under study. The point is not that the interviewees have veto power over what a researcher will write. However, qualitative data requires that the interviewees are involved in some manner in seeing if the findings that have been developed are in sync with their version of reality.

Minimally, we are suggesting that member checks ensure accuracy. An author should not write that the interviewee was 58 years old when she was 48, and so on. But more importantly, because of the indeterminacy of language, what one says may always be misinterpreted by the listener. Accuracy of the spoken word is not always possible, but one way to ensure that the researcher is closer to the intended meaning of the speaker is to try to gain understanding that is authentic (Wolcott, 1990). Accordingly, the research subjects' different constructions and interpretations of a

situation need to be solicited and portrayed in a manner that is even-handed and judicious. Frequently, qualitative research has been conducted on issues pertaining to people on the margins, and oftentimes those people have been portrayed as powerless or embedded in a culture of poverty. Just as often, certain individuals get portrayed in these situations as oppressors. A study of immigration, for example, may discuss those who are undocumented as culturally impoverished and incapable of change; those who go to the border of the United States to catch individuals trying to come into this country may be seen as culturally ignorant imperialists. Our assumption is that neither group sees themselves in that manner. Undocumented individuals most likely think they face challenges, but they may also see themselves in a quite positive, empowering manner. Similarly, those who try to defend the border may not think of themselves as racist bigots; instead, they may believe they are patriots. Qualitative work that is authentic struggles to ensure that the interpretations people give to their lives are represented rather than falling by the wayside of an omnipotent author. For most qualitative researchers, an objective, knowable world does not exist. Yet, a better understanding of the world is possible, even if it is unstable. Authenticity as a standard improves credibility because it allows for multiple individuals to co-construct a more accurate interpretation of the phenomena studied. Authenticity ensures that those under study will be fairly represented and portrayed, and that steps will be taken to ensure that individuals are able to react to the data that have been collected and the manner in which they have been interpreted and portrayed.

Presentation. All of the previous criteria pertain to this final point. Unlike in quantitative work, the manner in which an author portrays the findings becomes a key part of the strength (or weakness) of the study. The manner in which one describes the situation and what gets included about the methodology, research design, researcher standpoint, and the like become much more complex decisions than simply cutting and pasting one's methodological findings. The elegance of the interpretations one provides and how one delineates whether those under investigation concur with the findings will either keep the reader involved in reading the text or fail because of a wooden or flawed writing style. If a text is written so narrowly that readers can make no larger connections, then transferability will be impossible. Furthermore, presentation is not bound by text alone. Photographs of a dilapidated apartment complex can vividly and quickly inform an audience of the living conditions of urban students.

Presentation is a skill that can be developed like any other. Just as someone becomes trained in regression discontinuity or multiple regression, so too can someone become versatile in portraying situations with style and grace. And just as some researchers will be better than others at one or another method, so too will some writers be better than others. The challenge for scholars is that they frequently need to un-learn or be able to write in a different register from standard academic prose. Qualitative research will be most useful for policy-oriented work if the texts are able to convince by the words, ideas, and images employed. Such a statement suggests that qualitative research has more in common with disciplines such as philosophy or history rather than those who subscribe to quantitative methodologies. Numerous

strategies, including the use of thick description, member checks, and peer feedback, not only improve credibility but also minimize the potential variance of meaning that occurs between the writer and the reader (Wolcott, 1990). The challenge for these criteria is in the ability of the author to convince the reader of the interpretations that have been developed and of what utility they may be for impacting policy.

Conclusion

Methodological arguments have raged for multiple decades (Gage, 1989; Guba & Lincoln, 2000; Lather, 2004). The points of contention among researchers, which have resulted in vehement disagreements, reflect different ontological, epistemological, and axiological beliefs. We acknowledge the methodological fissures caused after such contention. We also acknowledge a void in the current processes of policy formulation, implementation, and evaluation. We have argued here that standard utilization of quantitative criteria for ensuring validity and reliability in a research project is impossible for qualitative work. Unlike many qualitative critiques, however, which challenge the epistemological notions of quantitative work, we have struggled with what sorts of criteria we might utilize if one wants to undertake research that will help inform policy making. In doing so, we have tried to walk a narrow path. On the one hand, we have not wished to embark on yet another philosophical critique of positivism, and on the other hand, we have no desire to disdain the thoughtful experimental work of those who undertake auto-ethnography and the like.

Instead, we have tried to develop criteria that might be employed for those who wish to use qualitative work to inform public policy. An elegantly designed qualitative study has the ability to inform policy just as an elegantly designed quantitative study may do, albeit from a very different epistemological stance. Yet, the ways in which the two studies inform policy differ. This is the point we wish to highlight. Rather than viewing qualitative and quantitative methods as antagonistic in relation to policy, we view them as complimentary for policy-related work. Just as effect sizes can indicate the strength of relationship between two variables, so too can thick description present a rich understanding of the context of those relationships.

References

Altheide, D. L., & Johnson, J. M. (1994). Criteria for assessing interpretive validity in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative methodology* (pp. 485–499). Newbury Park, CA: Sage.

Altheide, D. L., & Johnson, J. M. (2010). Reflections on dimensions of interpretive adequacy in qualitative research. In N. K. Denzin & Y. Lincoln (Eds.), *Handbook for qualitative research*. Thousand Oaks, CA: Sage.

American Educational Research Association (2006). *Standards for reporting on empirical social science research in AERA publications*. Washington, DC: Author.

American Educational Research Association (2008). *Definition of scientifically based research*. Washington, DC: Author.

American Educational Research Association (2009). *Standards for reporting on humanities-oriented research in AERA publications*. Washington, DC: Author.

Angen, M. J. (2000). Evaluating interpretive inquiry: Reviewing the validity debate and opening the dialogue. *Qualitative Health Research*, *10*(3), 378–395.

Boas, F. (1964). *The central Eskimo*. Lincoln, NE: University of Nebraska Press.

Bourgois, P., & Schonberg, J. (2009). *Righteous Dopefiend*. Berkeley, CA: University of California Press.

Bulterman-Bos, J. A. (2008). Will a clinical approach make education research more relevant for practice? *Educational Researcher*, *37*(7), 412–420.

Burawoy, M. (1998). The extended case method. *Sociological Theory*, *16*(1), 4–33.

Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, *54*(4), 297–312.

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

Carlyle, T. (1897). *The works of Thomas Carlyle: Critical and miscellaneous essays*. New York: Peter Fenelon Collier.

Caulley, D. N. (2008). Making qualitative research reports less boring: The techniques of writing creative nonfiction. *Qualitative Inquiry*, *14*(3), 424–449.

Cho, J., & Trent, A. (2006). Validity in qualitative research revisited. *Qualitative Research*, *6*(3), 319–340.

Christensen, C. M., Johnson, C. W., & Horn, M. B. (2008). *Disrupting class: How disruptive innovation will change the way the world learns*. New York: McGraw.

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

Corbin, J., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluation criteria. *Qualitative Sociology*, *13*(1), 3–21.

Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into Practice*, *39*(3), 124–131.

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

Donmoyer, R. (2000). Generalizability and the single-case study. In R. Gomm (Ed.), *Case study method*. London: Sage.

Education Sciences Reform Act of 2002, Pub. L. No. 107–279 (2002).

Eisenhart, M. A., & Howe, K. R. (1992). Validity in educational research. In M. D. LeCompte, W. L. Millroy, & J. Preissle (Eds.), *The handbook of qualitative research in education*. San Diego: Academic Press.

Feuer, M. J. (2006). Response to Bettie St. Pierre's "Scientifically based research in education: Epistemology and ethics". *Adult Education Quarterly*, *56*(4), 267–272.

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

Gage, N. L. (1989). The paradigm wars and their aftermath: A "historical" sketch of research on teaching since 1989. *Educational Researcher*, *18*(7), 4–10.

Garratt, D., & Hodkinson, P. (1998). Can there be criteria for selecting research criteria? A hermeneutical analysis of an inescapable dilemma. *Qualitative Inquiry*, *4*, 515–539.

Geertz, C. (1973). Thick description: Toward an interpretive theory of culture. In *The interpretation of cultures: Selected essays*. New York: Basic.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.

Gomm, R., Hammersley, M., & Foster, P. (2000). Case study and generalization. In R. Gomm (Ed.), *Case study method* (pp. 98–115). London: Sage.

Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communication and Technology*, *29*(2), 75–91.
Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.
Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105–117). Thousand Oaks, CA: Sage.
Guba, E. G., & Lincoln, Y. S. (2000). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 191–215). Thousand Oaks, CA: Sage.
Hammersley, M. (1992). *What's wrong with ethnography?* London: Routledge.
Hammersley, M. (2006). Assessing validity in social research. In P. Alasuutari, L. Bickman, & J. Brannen (Eds.), *Handbook of social research methods* (pp. 42–53). London: Sage.
Hammersley, M. (2009a). Challenging relativism: The problem of assessment criteria. *Qualitative Inquiry*, *15*(1), 3–29.
Hammersley, M. (2009b). Closing down the conversation? A reply to Smith and Hodkinson. *Qualitative Inquiry*, *15*(1), 40–48.
Harrington, M. (1997). *The other America: Poverty in the United States*. New York: Touchstone.
Henig, J. R. (2008). The evolving relationship between researchers and public policy. *The Phi Delta Kappan*, *89*(5), 357–360.
Kozol, J. (1991). *Savage inequalities: Children in America's schools*. New York: Crown.
Kozol, J. (2005). *The shame of the nation: The restoration of apartheid schooling in America*. New York: Crown.
Kvale, S. (1994). Ten standard objections to qualitative research interviews. *Journal of Phenomenology Psychology*, *25*(2), 147–173.
Kvale, S. (1996). The social construction of validity. In *Interviews: An introduction to qualitative research interviewing*. Thousand Oaks, CA: Sage.
Lamont, M., & White, P. (2005). *Workshop on interdisciplinary standards for systematic qualitative research*. Washington, DC: National Science Foundation.
Lather, P. (1986). Issues of validity in openly ideological research: Between a rock and a hard place. *Interchange*, *17*, 63–84.
Lather, P. (1993). Fertile obsession: Validity after post-structuralism. *Sociological Quarterly*, *35*, 673–693.
Lather, P. (2004). This IS your father's paradigm: Government intrusion and the case of qualitative research in education. *Qualitative Inquiry*, *10*(1), 15–34.
LeCompte, M. D., & Goetz, J. P. (1982). Problems of reliability and validity in ethnographic research. *Review of Educational Research*, *52*(1), 31–60.
Liebow, E. (1995). *Tell them who I am: The lives of homeless women*. New York: Penguin.
Lincoln, Y. S. (1995). Emerging criteria for quality in qualitative and interpretive research. *Qualitative Inquiry*, *1*(3), 275–289.
Lincoln, Y. S. (2001). Varieties of validity: Quality in qualitative research. In J. Smart & W. G. Tierney (Eds.), *Higher education: Handbook of theory and research* (pp. 25–72). New York: Agathon Press.
Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
Malinowski, B. (1922). *Argonauts of the Western Pacific*. New York: Dutton.
Mead, M. (1928). *Coming of age in Samoa*. Oxford: Morrow.
Miles, M. B., & Huberman, A. M. (1984). Drawing valid meaning from qualitative data: Toward a shared craft. *Educational Researcher*, *13*(5), 20–30.
Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.
Mishler, E. G. (1990). Validation in inquiry-guided research: The role of exemplar in narrative studies. *Harvard Educational Review*, *60*, 415–422.
Morrow, S. L. (2005). Quality and trustworthiness in qualitative research in counseling psychology. *Journal of Counseling Psychology*, *52*(2), 250–260.

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*(7), 501–517.

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

Phillips, D. C. (1987). Validity in qualitative research: Why the worry about warrant will not wane. *Education and Urban Society*, *20*(1), 9–24.

Ragin, C. C., Nagel, J., & White, P. (2004). *Workshop on scientific foundations of qualitative research*. Washington, DC: National Science Foundation.

Ray, R. C. (1980). Blitzkrieg ethnography: On the transformation of a method into a movement. *Educational Researcher*, *9*(2), 8–10.

Rolfe, G. (2004). Validity, trustworthiness, and rigour: Quality and the idea of qualitative research. *Journal of Advanced Nursing*, *53*, 304–310.

Ruddin, L. P. (2006). You can generalize stupid! Social scientists, Bent Flyvbjerg, and case study methodology. *Qualitative Inquiry*, *12*(4), 797–812.

Sandelowski, M. (1986). The problem of rigor in qualitative research. *Advances in Nursing Science*, *8*(3), 125–130.

Sandelowski, M. (1993). Rigor or rigor mortis: The problem of rigor in qualitative research revisited. *Advances in Nursing Science*, *16*(2), 1–8.

Schofield, J. W. (2000). Increasing the generalizability of qualitative research. In R. Gomm (Ed.), *Case study method* (pp. 69–97). London: Sage.

Seale, C. (1999). Quality in qualitative research. *Qualitative Inquiry*, *5*(4), 465–478.

Shavelson, R. J., Phillips, D. C., & Feuer, M. J. (2003). On the science of education design studies. *Educational Researcher*, *32*(1), 25–28.

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

Smith, J. K. (1993). *After the demise of empiricism: The problem of judging social and educational inquiry*. Norwood, NJ: Ablex.

Smith, J. K., & Hodkinson, P. (2009). Challenging neorealism: A response to Hammersley. *Qualitative Inquiry*, *15*(1), 30–39.

St. Pierre, E. (2006). Scientifically based research in education: Epistemology and ethics. *Adult Education Quarterly*, *56*, 239–266.

Stokes, D. E. (1997). *Pasteur's quadrant: Basic science and technological innovation*. Washington, DC: Brookings Institution Press.

Tierney, W. G., & Hallett, R. E. (2009). Writing on the margins from the center: Homeless youth and politics at the Borders. *Cultural Studies <=>Critical Methodologies*, *10*, 19–27.

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

US Department of Education (2010). About IES: Connecting research, policy and practice. Retrieved April 07, 2010, from http://ies.ed.gov/aboutus/

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