

Standards for Qualitative (and Quantitative) Research: A Prolegomenon

KENNETH HOWE

MARGARET EISENHART

The proliferation of qualitative methods in educational research has led to considerable controversy about standards for the design and conduct of research. This controversy has been playing itself out over the last several decades largely in terms of the quantitative-qualitative debate. In this paper we argue that framing the issue of standards in terms of quantitative-qualitative debate is misguided. We argue instead that the problem of standards—for qualitative and quantitative research—is best framed in terms of the “logics in use” associated with various research methodologies. In particular, rather than being judged in terms of qualitative versus quantitative paradigms, logics in use, which are often drawn from other academic disciplines and adapted for the purposes of educational research, are judged in terms of their success in investigating educational problems deemed important. Finally, we proffer five general standards that can apply to educational research of all kinds.

Educational Researcher, Vol. 19, No. 4, pp 2-9

Until 20 years ago, qualitative methodology lacked a history within educational research; not surprisingly, its appearance on the scene prompted concerns about its legitimacy. These concerns have continued, both inside and outside of qualitative research practice. For instance, Rist (1980), himself a qualitative researcher, admonished the educational research community to beware of “blitzkrieg ethnography”; Phillips (1987a), a philosopher, insisted that the “concern with warrant won’t wane” (p. 9). Like Rist and Phillips, we worry that in their eagerness to embrace qualitative methods, many educational researchers do not provide adequate and clear justifications for their methods, findings, or conclusions. In our estimation, that justifications are often inadequate or unclear is due in no small measure to confusion about how best to think about standards for qualitative research design and analysis. We do not mean to suggest that no one has provided useful discussions of standards for qualitative research, much less that no such standards exist. We mean to suggest instead that various aspects of the discussion of standards, particularly its epistemological aspects, stand in need of clarification if the discussion is to proceed in the most fruitful way.

Initially, the debate about the legitimacy of qualitative

research in education was cast in terms of a stark but vaguely characterized choice between an entrenched quantitative methodology and a highly suspect newcomer. The debate has since evolved into two distinct strands. The first strand, exemplified by thinkers such as Lincoln and Guba (1985) and Smith and Heshusius (1986), refined the early debate by distinguishing between research methods on one hand and epistemologies on the other. According to this view, the stark choice exists now as much as ever; however, it is not a choice between quantitative and qualitative methods but between the positivist and alternative (nonpositivist) epistemologies that putatively underlie alternative research paradigms. For these thinkers, justifying qualitative research consists of rigidly distinguishing positivist and alternative epistemologies and then defending the legitimacy of the latter.

The second strand of argument, exemplified by thinkers such as Denzin (1989), Erickson (1982, 1986), and Goetz and LeCompte (1984), largely left the original debate behind, as well its more recent refinements. For these thinkers, attention is focused on the particulars of various research methodologies rather than on abstract epistemology. Accordingly, justifying qualitative research largely consists of developing and articulating methodological design and analysis standards rather than of fending off positivistic challenges.

Because of these two very different ways of approaching the problem, the current discussion of standards for qualitative design and analysis is at an impasse. In particular, thinkers working on the development and articulation of standards pay relatively little attention to the demand to frame their recommendations in terms of the positivist-alternative paradigm split; as a consequence, they are characterized as glossing over deeper epistemological issues, of capitulating merely to what works, and even of embracing positivism (Smith & Heshusius, 1986).

Our major aim in this paper is to offer some preliminary observations—a prolegomenon—with an eye toward mov-

KENNETH HOWE and MARGARET EISENHART are at the School of Education, University of Colorado, Boulder, Colorado 80309-0249.

ing the discussion beyond its current impasse. Our arguments will proceed in three parts. First, we will defend the second strand by arguing that a self-conscious epistemological justification for this approach is readily available: Its adherents merely need to explicitly dismiss the demand contained in the first strand of argument that the epistemological debate must be couched in terms of the positivist-alternative paradigm split.

Second, we will examine various issues concerning standards for qualitative research within a nonpositivist framework. We will claim that a variety of specific standards are legitimate, because standards must be linked to the different—and legitimate—disciplines, interests, purposes, and expertise that fall under the rubric of qualitative research. We will use examples from educational ethnography in order to illustrate the nature of disputes within a particular qualitative tradition and to show that such disputes, though important, should be disentangled from disputes about the general value of a piece of research for education.

Finally, we will advance five very general standards that might be applied to the design and analysis of qualitative educational research. Given the shape that our arguments take, we will suggest that these standards need not be confined to qualitative research in particular, but can and should apply to quantitative research as well (thus the reference to quantitative in the title). We emphasize qualitative standards because this is where the debate about standards has been, and will likely continue to be, focused.

Strand 1—The Positivist Alternative Paradigm Split: A Procrustean Bed

Procrustes was a legendary robber of ancient Greece who had the habit of cutting off or stretching his victims legs, depending on their height, to make them fit his bed. Thinkers attempting to work out standards for research methods make a procrustean bed of their own when they construe explicating standards as an exercise in epistemological foundations. The basic approach is to characterize positivist and alternative paradigms respectively in terms of various dichotomies—facts versus values, objectivity versus subjectivity, fixed categories versus emergent categories, the outsider's perspective versus the insider's perspective, a static reality versus a fluid reality, causal explanation versus understanding—and to identify qualitative research with the characteristics associated with this alternative paradigm (e.g., Guba, 1987; Lincoln & Guba, 1985, 1988; Smith, 1983a, 1983b; Smith & Heshusius, 1986). Standards for qualitative research are then stretched or cut down, as the case may be, to fit the alternative paradigm. Insofar as qualitative research gets associated with an epistemological paradigm that rejects things like facts and objectivity, it becomes vulnerable to the familiar charges that it is hopelessly subjective, unscientific, relativistic, and is virtually without any standards at all.

This problem regarding standards of warrant, supposedly peculiar to qualitative research, can largely be avoided by recognizing that the strategy of articulating standards against a positivistic epistemological backdrop is by no means an obviously legitimate one. In this vein, several educational philosophers (e.g., Garrison, 1986; Howe, 1985, 1988; Phillips, 1983, 1987b) have argued that philosophy of science has moved into a post- or nonpositivistic¹ era in which

positivism is no longer a tenable epistemological position. Given this new philosophy of science, no social research (nor even physics for that matter) is accurately portrayed by positivism, and thus positivism should not serve as the foil against which standards for qualitative research should be developed. This point deserves some further elaboration.

Positivism was initially conceived as partly a description of, and partly a prescription for, the conduct of natural sciences. In Kaplan's (1964) phraseology, it was "reconstructed logic." In the arena of natural science, practicing scientists largely ignored both aspects of positivism, and were guided instead by what Kaplan refers to as "logic in use." That positivism so badly failed as a reconstruction of the logic of natural science and that it was largely ignored by natural scientists makes it somewhat ironic that it has been taken so seriously by social scientists; but take it seriously they did, especially in psychology, where positivism was embraced as an accurate portrayal of the scientific method, and then was cashed out in the form of methodological behaviorism (e.g., MacKenzie, 1977).

John Passmore observed (in 1967 no less) that "logical...positivism is dead, or as dead as a philosophical movement ever becomes" (In Phillips, 1987b, p. 37). Although positivism no doubt still has a powerful influence on social and educational research, it can no longer claim to be based on a viable epistemology. For, the core tenet of positivism, verificationism, has been thoroughly repudiated. In general, the picture of empirical science envisioned by positivism, in which observation could be strictly separated from and remain untainted by the purposes that animate the conduct and evaluation of scientific investigation, has been replaced by the notion that all scientific investigation is inherently theory-laden. Consequently, because all scientific investigation is inherently laden with theory, inherently an outgrowth of human purposes and theoretical constructions, it is, broadly speaking, inherently interpretive² (e.g., Bernstein, 1983; Rorty, 1982). Thus, there is no good reason for educational researchers to attempt to legitimate an alternative paradigm so that it might peacefully coexist with positivism. Indeed, there are good reasons for not doing this, inasmuch as it merely serves to encourage the view that positivism is a worthy competitor.

Abandoning positivism does not entail abandoning standards of objectivity and rationality in empirical research; it entails instead that such standards be understood in a non-positivist way. In particular, the question of standards must be viewed wholly within an interpretive perspective, broadly construed. Furthermore, insofar as no standards completely divorced from human judgments, purposes, and values can exist and insofar as there can, accordingly, be no monolithic unity of scientific method—those were the pipe dreams of positivism—standards must be anchored wholly within the process of inquiry. As Kaplan remarks in the opening paragraph of the Conduct of Inquiry (1964), "the pursuit of truth is accountable to nothing and no one not a part of the pursuit itself" (p. 3). As he adds a few pages later, "standards governing the conduct of inquiry in any of its phases emerge from the inquiry and are themselves subject-to-further inquiry" (p. 5). The question of standards for qualitative research—indeed, for research of any kind—is, then, a fluid one, and one that must be answered in terms of the successes and failures of inquiry. In turn, successes and failures can only be judged relative to given purposes.

Strand 2—Reformulating the Problem: Logics in Use

Kaplan's focus on the standards that are actually employed in social research, standards that he associates with logic in use, is consistent with other thinkers like Bernstein (1983), who urges researchers to overcome the tyranny of method, and Rorty (1979), who urges them to give up the notion of an Archemedian point which might serve as a fail-safe criterion against which to evaluate standards. What happens

Positivism was conceived as a description of, and a prescription for, the conduct of natural sciences.

when the tyranny of method and the quest for an Archemedian point are abandoned in favor of working out logics in use?

The Proliferation of Standards

One consequence of this general nonpositivist stance is that, insofar as methodology ultimately must be tied to research purposes, it must accordingly respond to the variety of purposes that exist. Thus, legitimate research methodologies may and should proliferate. That social research methodologies may legitimately proliferate is especially pertinent to educational research. For, as Shulman (1988) has observed, education is a field of study rather than a discipline. That is, it must bring to bear other disciplines—psychology, sociology, and anthropology, to name a few—on educational problems. Consequently, the ways of thinking about methodologies that exist in any one of these disciplines multiply and overlap when it comes to educational research.

For example, qualitative researchers who draw on Denzin's work, *The research act: A theoretical Introduction to Sociological Methods* (1989), are advised to consider the methods of participant observation, naturalistic inquiry, sociological interviewing, and biography in light of the research purposes of symbolic interactionism. Those who draw on the work of Goetz and LeCompte, *Ethnography and Qualitative Design in Educational Research* (1984), are asked to consider most of the same methods in light of the research purposes of ethnography. Smith and Glass, in *Research and Evaluation in Education and the Social Sciences* (1987), discuss a few of the same methods with respect to the purposes of naturalistic inquiry. Although many of the procedures these authors describe are identical, their use in conjunction with theory varies, and their strengths and weaknesses, given certain purposes, are different.

In defining and illustrating their particular methodologies, all of these authors have written about standards for assess-

ing quality and rigor. These standards, like the selection of the methods themselves, are related to the theoretical orientation of the authors. This state of affairs illuminates the problem in our search for a way to think and talk about standards for qualitative research: Except at a very high level of abstraction, it is fruitless to try to set standards for qualitative research per-se. Even when the focus within qualitative research is significantly restricted, the issues associated with standards are quite complex and extensive. We will use educational ethnography as a case in point. (We emphasize that educational ethnography is but one research tradition, and that we are using it only to illustrate more general features of the nature of standards in educational research.)

Standards in Educational Ethnography: An Illustration

When Rist (1980) expressed his concern that the growing use of ethnography by educational researchers was becoming a mutating movement of an undisciplined mob, he used the term *blitzkrieg ethnography* to refer to the work of self-styled ethnographers who were not trained in or had not studied the method, who did not appreciate the emphasis on exploring the cultural framework of the group or organization in question, and who used various means to reduce the time and uncertainty of traditional fieldwork. Rist worried that the blitzkrieg ethnographer, by not "accepting the domain and underlying assumptions that have heretofore guided the method [is] essentially...free to improvise and relabel [almost anything] as a new form of ethnography" (p. 9). At the same time, Wolcott (1980) voiced a related concern.

Although we understand the worry about labeling anything *ethnography*, we think it is important to observe that neither the domain and underlying assumptions nor the preferred techniques of ethnography have remained fixed and uncontested through time. The recent writings of Marcus and Fischer (1986), Geertz (1988), and Clifford (1988) illustrate the profound effects of social history and theoretical development on definitions and standards for ethnography. Clifford points out that ethnography has variously been a methodology to describe, to explain, and to interpret; that it has been at some points in time fundamentally historical, at other points ahistorical; that it has sometimes emphasized the natural scientist's external observation and at other times emphasized the cultural insider's interpretations and practices. Geertz describes how different ways of "being there," that is, the ethnographer's warrant for the authenticity of his or her account, have been enacted in the works of Benedict, Levi-Strauss, Evans-Pritchard and Malinowski. Geertz (1988) does not present ethnography as a fixed set of rules and procedures but as a series of challenges:

Finding somewhere to stand in a text that is supposed to be at one and the same time an intimate view and a cool assessment is almost as much of a challenge as gaining the view and making the assessment in the first place. (p. 10)

Methodological priorities and concerns held over from the recent past are being challenged and in some cases giving way to new ones. This drift—this evolution of logic in use—is occurring for both theoretical and practical reasons, and may be illustrated with two examples of recent ethnographic research. These examples also make two additional points about the general nature of research standards. First, a methodology must be judged by how well it informs research

purposes, at least as much as by how well it matches a set of conventions. Second, what counts as good educational research will not necessarily match what counts, at any given point in time, as orthodox methodology; for methodology must respond to the different purposes and contexts of research.

Example 1. The first example comes from the work of Roman (1989). Roman set out to conduct an ethnographic study of punk rockers. She began the study as a project in a class taught by an anthropologist of education. She conscientiously intended to use conventional ethnographic methods as presented in the class: gradually participating in the lives of the punk rockers ("going native") by unobtrusively observing them ("like a fly on the wall") and talking with them. She soon realized, however, that these conventional ethnographic methods were inconsistent with her theoretical and political commitment as a feminist materialist. She writes,

I discovered in the course of doing the field work that these...conventions for describing or conveying the appropriate role of an ethnographer actually had the undesirable effect of reproducing my relation to the young women Punks through forms of class privilege and gendered viewing (the distant but fascinated researcher), which I call respectively, "intellectual tourism" and "voyeurism" (p. 7). [This situation] required me to confront a new set of ethical and political dilemmas regarding the level and nature of my involvement in the daily gender and class issues facing the young women. On many occasions the young women themselves demanded that I respond to the particular conditions and situations facing them in their gender relations with the male Punks or in class relations among themselves. I found on these occasions that it was simply impossible and politically untenable to remain a silent or passive observer. (p. 13)

After considerable thought about the bases of conventional ethnographic standards, her own methodological and ethical requirements, and the everyday reality of the punk women (including their experiences of assault and subordination in their families and subculture), Roman changed her research design. She dropped her stance as a neutral or passive researcher and as a thoroughly immersed participant observer. She began to search for ways to meet with the punk women without the men present (unusual in this group) and to participate with the women in defining their gender and class oppression (this participation was unnatural, i.e., a disruption of the ordinary course of events or interpretation). As she formed tentative ideas about the women as a group, she shared these ideas with the women, elicited their responses, and further tried to alter their ways of thinking about their lives. Roman has called her methodology "feminist materialist ethnography," to distinguish it from conventional or naturalistic ethnography.

Some have questioned whether Roman's work is truly ethnographic.³ The issue seems to be that although her study was about the culture of a small group of punks, her methodology was not recognizable as traditional ethnography. We understand what provokes such questions. There are good reasons to exercise some control over what is to count as an ethnography: If the term is to have any meaning, it has to rule certain things out. Insofar as no Platonic form for ethnography exists, however, what is to

count as ethnography is neither given once and for all nor impervious to challenge. Instead, it must be determined by what ethnographers, presumably with good reason and after some debate, decide. Moreover, challenges to the research status quo are not only inevitable; when thoughtfully advanced, such challenges are also healthy. Roman's challenge, for instance, is based on careful consideration of methodological issues in light of her political, ethical, and practical purposes for the research. Her purposes demand a rethinking of conventional procedures, and why not?

Example 2. The second example concerns a dissertation study called "an ethnographic study" by its author (Naff, 1987; Naff Cain, 1989). The study was not designed to investigate culture per se; rather it was designed as a comparative study of the classroom implementation of two distinctly different teacher planning models.

The research focused on two student teachers, matched on many background characteristics and student teaching experience. The only (known) relevant differences between the two was the planning model they were taught to use. To capture and account for any differences in the teachers' thinking and actions, Naff Cain used 14 sources and methods to collect data about the student teachers' training in planning and about their experiences as they planned and implemented a 1-month 12th-grade unit on the play King Henry IV. To collect these data, Naff Cain chose tools often used by ethnographers, such as repeated, open-ended interviews with the student teachers, their cooperating teachers, and their students; participant observation in each classroom; videotapes of class sessions in each room; cooperating teacher logs; stimulated recalls and heuristic elicitations with the student teachers; and collection of artifacts such as university and school documents, unit plans, daily plans, class handouts, and daily journals from the student teachers and cooperating teachers.

Her conclusions were based on findings triangulated from these data sources and methods, and were analyzed in two ways also borrowed from ethnographers—semantic domain analysis (Spradley, 1980) and vignette analysis (suggested by Erickson, 1986, and VanMaanen, 1988). Her conclusions address the issue with which she began: the differences in teachers and their classrooms that were associated with different planning models.

Like Roman's work, Naff Cain's has been criticized by anthropologists for not being ethnography, but for different reasons. In Roman's case, the focus of research—culture—is appropriate to study using ethnographic methods, and it is the appropriateness of the alternative methods she used that engenders the controversy about whether the study is truly ethnography. For Naff Cain, just the reverse is true: The methods are technically correct, and it is the focus of the research—a comparison of planning models—that engenders the controversy about whether the study is truly ethnography.

Is Naff Cain's study truly an ethnography? Perhaps not. However, it doesn't seem as though answering this question in one way or another is, or should be, crucial for education (presuming, of course, that her methods otherwise yield warranted conclusions). In this connection, and in contrast to the reception by anthropologists, Naff Cain's study has received critical acclaim among educators and educational researchers. It was, for example, the basis for naming her a

National Council of Teachers of English (NCTE) Promising Young Researcher in 1988. It has been used as a model for ethnographic research in English education and was the impetus for revisions in several teacher education programs.

The general point we wish to make with the Roman and Naff Cain examples is just this: Failing to follow a given

**Neither the domain nor
the preferred techniques
of ethnography
have remained fixed.**

theoretical perspective or methodological convention does not necessarily diminish the warrant of the conclusions drawn. Although this point might seem altogether obvious—a sociologist can hardly be criticized for failing to observe the methodological canons of physics—it is too easily obscured when researchers work in a recognizable area but, like Naff Cain, with theoretical orientations or, like Roman, with research methods that are just far enough removed from convention to raise questions about how they should be classified. In these kinds of fuzzy situations (which typically attend innovation), questions of definition and questions of warrant easily become entangled, such that if it is not ethnography (or ethnographic or what have you), then it is not good research. Yet settling the definitional question—Is this really ethnography? Is this really ethnographic?—cannot settle the question of whether a given piece of educational research is worthwhile. The question that needs to be answered instead is more general: Are warranted conclusions obtained about some important educational question(s)? This is the question that ought to frame the pursuit of standards for any educational research.

Standards for Educational Research

We began this paper by expressing our worries about the failure to develop adequate and clear standards in the rush to embrace qualitative methods in educational research. So far, our attention has been focused on just how the question of standards should be framed, for, in our estimation, this has been a major obstacle to progress. Before turning to the task of proposing several general standards, we will briefly explain three features of them that grow out of our discussion so far.

First, any general standards for evaluating educational research will have to be very abstract. Because educational research cuts across many disciplines and their associated methodologies, and because no one can be expected to be a master of all of the relevant disciplines, general standards will have to incorporate deference to scholars of the various disciplines on issues of methodology and substance peculiar to the disciplines in question. For example, whether Naff Cain's study is ethnography, or whether it effectively applies ethnography's research methods, is up to ethnographers to decide.

Second, notwithstanding the deference to expertise just described, there must be some feature(s) of educational research that justify the term educational and therefore make it of interest and value to educators. The most obvious requirement is that it must focus on educational issues. This rather vague requirement admittedly leaves certain important questions unanswered. For example, must educational research also incorporate a disciplinary (or theoretical) perspective? Is mere description enough, or must improvement be the direct aim? These are just the sorts of questions that Scriven (1986) raises in his call for educational research to adopt an evaluation paradigm and to emulate the medical research model (a model which brackets theoretical understanding in the quest for relatively immediate remedies). Despite the importance and complexity of these questions, here we must set them aside. The five standards we suggest do not presuppose that they be answered in one way rather than another.

Third, as we stated in the introduction, despite our focus on qualitative research (educational ethnography in particular), our arguments are general enough to apply to quantitative research. This is because our position is staunchly anti- or nonpositivist: it requires all educational research to be grounded in a nonpositivist epistemological perspective. Although positivism helped spawn a set of methods (typically quantitative) such that a vestigial methodological positivism is still alive and well, we deny that such methods can be justified by an appeal to positivist epistemology. Instead, such methods must, like qualitative methods, satisfy the kinds of nonpositivist standards we are about to advance.⁴

The Fit Between Research Questions and Data Collection and Analysis Techniques

Hilary Putnam remarks, "If you want to know why a square peg doesn't fit into a round hole, you had better *not* describe the peg in terms of its constituent elementary particles (cited in Rorty, 1982, p. 201). Although Putnam's target is reductionism in scientific explanation, his remark also has a more prosaic meaning: the data collection techniques employed ought to fit, be suitable for answering, and the research question entertained. Thus, in Naff Cain's planning study the research questions called for descriptive data permitting comparisons of the two classrooms. Likewise, Roman's emergent, critically oriented research questions demanded a reciprocal, change-oriented design.

A corollary of this standard is that research questions should drive data collection techniques and analysis rather than vice versa—and this is the form in which it is most often violated. Consider what might be termed the *quiet desperation syndrome*, a disease that attacks the nervous systems of doctoral students. Students who are afflicted begin with a method, "I want to do a qualitative study," "I want to do a MANOVA," and then cast about for a question. Thus, the question of the fit between research questions and data collection techniques gets turned on its head.

Correctly ordering research questions and methods is, of course, a complex issue. We do not mean to suggest that researchers can proceed as if they are blank slates—free of prior interests, commitments, and methodological expertise—nor as if they have super intellects—capable of competently choosing from all of the relevant questions and methodologies—nor, finally, as if they had available infinite

time and resources. In some sense, then, research methodology will indeed drive research. In contrast, the degree to which this occurs should be minimized. In our estimation, it is incumbent upon educational researchers to give careful attention to the value their research questions have for informing educational practice, whether it be at the level of pedagogy, policy, or social theory, and then to ground their methodology in the nature of these questions.

The Effective Application of Specific Data Collection and Analysis Techniques

In addition to coherently linking up with research questions, data collection and analysis techniques also must be competently applied, in a more or less technical sense. Various principles guide how interviews should be conducted, how instruments should be designed, how sampling should proceed, how data should be reduced, and so forth, such that rather immediate low inference conclusions are rendered credible. If credibility, warrant, is not achieved at this level, then the more general (and interesting) conclusions that ultimately rest on these low inference conclusions will be suspect.

As with the first standard, we do not mean to imply that there are hard and fast rules that must be followed; indeed, such a stance would run counter to much of what we have had to say. We agree with Strauss (1987) that methodology is best characterized as rules of thumb, that is, "guidelines that should help most researchers in their enterprises" (p. 7), and we could not articulate it better than Strauss himself does:

...researchers need to be alive not only to the constraints and challenges of research settings and research aims, but to the nature of their data. They must also be alert to the temporal aspects or phasing of their researches, the open-ended character of the "best research" in any discipline, the immense significance of their own experiences as researchers, and the local contexts in which the researches are conducted....Methods, after all, are developed and changed in response to changing work contexts (pp. 7-8).

Alertness to and Coherence of Background Assumptions

Linking research questions with data collection techniques and competently applying the latter do not insure that a study will render warranted conclusions, for studies must be judged against a background of existent knowledge. For example, if the results of one study contradict those of another (or several others), then some sort of explanation of why this occurred is in order. This is where the familiar review of the literature comes into play.

Whether some grand social scientific theoretical orientation is employed (e.g., Roman's feminist materialist orientation), or whether research is more specifically focused on pedagogy (e.g., Naff Cain's focus on teacher planning), background assumptions should guide the research questions and methods in a coherent and consistent fashion. Perhaps less obvious, and especially relevant to qualitative research, is the researcher's own subjectivity (Peshkin, 1988). Peshkin has argued that subjectivity is the basis for the researcher's distinctive contribution, which comes from joining personal interpretations with the data that have been collected and analyzed. As with assumptions derived from the literature, subjectivities must be made explicit if they are to clarify, rather than obscure, research design and findings.

Overall Warrant

As we are using the term, *overall warrant* encompasses responding to and balancing the first three standards discussed as well as going beyond them, to include such things as being alert to and being able to employ knowledge from outside the particular perspective and tradition within which one is working, and being able to apply general principles for evaluating arguments.

Although it is difficult (indeed wrong-headed) to try to nail down the notion of overall warrant in a much more precise way, some additional articulation is nonetheless possible. For instance, theories, whether derived from the literature or personal experience, are themselves up for grabs. For this reason, it seems that the most warranted conclusions of which we are capable at any given point in time are those that are drawn after robust and respected theoretical explanations have been tentatively applied to the data—what Denzin (1989) and Shulman (1988) call "triangulation by theory"—and the most plausible one, or some modified version of it, is used to explain the research results. Of course, the warrant of such conclusions also rests on the warrant of the research results, which can be assessed, we have argued, by using the three standards already proposed.

Another way to discuss the issue of overall warrant is to call for discussion of disconfirmed theoretical explanations and disconfirming evidence (Erickson, 1986). When researchers explain the arguments by which some theories are rejected and by which disconfirming data are handled, their conclusions are more warranted than when they do not.

Value Constraints

Given the untenability of the positivist fact-value dogma, there is little reason to suffer from what Scriven (1983) calls "value phobia." The conduct of educational research is subject to both external and internal value constraints (Howe, 1985).

External. External value constraints have to do with the worth of research for informing and improving educational practice—the "so what?" question; that research might possess internal validity is insufficient. Although such judgments of educational worth can be very difficult to make, and have the potential to be exceedingly biased (anyone who has served on a human subjects committee can attest to both of these problems), they are not judgments from which researchers can, or do, forever run and hide—witness the recent exchange in *Educational Researcher* between Finn (1988), and Shavelson and Berliner (1988). It is best to get questions of the worth of research out on the table, lest implicit judgments operate behind the scenes, as a kind of hidden agenda. Clearly, even if others might be puzzled, educational researchers themselves should be able to, and be prepared to, communicate what value their research has (if only potentially) for educational practice.

Related to this, the conclusions of educational research ought to be generally accessible to the education community. That is, the language of the results and implications must be in a form that is understandable to, and debatable by, various actors in a particular setting—teachers, administrators, parents, and also educational researchers with varying perspectives and expertise. Accordingly, the research process itself must give attention to the nature of the contexts

and individuals it investigates and to which its results might be applied, that is, to their social, political, and cultural features.

Internal. Internal value constraints have to do with research ethics. We call research ethics "internal" because they have to do with the way research is conducted vis-a-vis research subjects, not with the (external) value of results. For example, Milgram's research on obedience to authority rendered valuable insights regarding the power of researchers to elicit compliance from subjects to perform ethically objectionable actions. The way Milgram treated his subjects was highly objectionable, however—so much so that he would not be permitted to do his research today. (Ironically, Milgram's findings, at least indirectly, underpin current requirements for informed consent, especially to clearly communicate to subjects that they are free to withdraw from research at any time and without penalty.)

Internal value constraints are distinguishable from standards of warrant insofar as observing them sometimes requires reducing warrant. For instance, randomized double-blind experiments are notorious for the kind of trade-off they engender between the risk-benefit ratio that applies to the subjects of such research and the value of the knowledge that can be obtained for guiding future action. Especially relevant to qualitative research, researchers must weigh the quality of the data they can gather (and whether they can gather any data at all) against principles such as confidentiality, privacy, and truth-telling. Although internal value constraints, research ethics, can be distinguished from more conventional issues of warrant, they are nonetheless clearly relevant to evaluating the goodness, that is, the acceptability or legitimacy of research designs and procedures.

Conclusion

As we stated at the outset, our aim in this paper has been to offer some preliminary observations that might serve to guide future discussion of standards in qualitative educational research. We did not set out to end the discussion of standards, but to redirect it. We will briefly recapitulate our arguments and then offer several general observations about where they leave us.

The common strategy of grounding qualitative research in an alternative paradigm creates a procrustean bed for itself by assuming that it must coexist with positivism. For, once it makes this assumption, it must then define itself as positivism's polar opposite, which entails relinquishing to positivism objectivity, facts, the outsider's perspective, and a host of other concepts that go into making up various dualisms. Refusing to entertain positivism as a viable epistemological doctrine—a refusal that is now univocal within the philosophy of science—is how to avoid this procrustean bed.

Escaping Procrustes' clutches sets the stage for reformulating the problem of standards. Once positivism is removed from the scene, the positivist-alternative paradigm split, along with its various dualisms, collapses; the upshot is that standards must be anchored wholly within a non-positivist perspective, which is to say they must be anchored nowhere other than in logics in use, in the judgments, purposes, and values that make up research activities themselves. Furthermore, within educational research there are various traditions, each with its own logic in use and its own peculiar disagreements about how methodology should evolve.

Because education is a field of study, which cuts across different logics in use, it presents special problems regarding standards for research. In particular, general standards for educational research—standards applicable to any research that can be called "educational"—will have to be relatively abstract and will have to turn certain questions of standards over to individuals possessing various kinds of expertise. Thus, by its very nature, educational research requires a division of labor. Such a division can take two forms: give-and-take collaboration or insular fragmentation. We surmise that the latter form has too often been the norm. Our five general standards, tentative and inchoate as they may be, are designed to promote the former.

Acknowledgments

An earlier version of this paper was presented at the Annual AERA meeting in San Francisco, March 30, 1989.

Thanks to our colleague Ernie House and to the *Educational Researcher* reviewers for their helpful comments on earlier versions of this paper.

Notes

¹Philosophers typically use *postpositivism* in a literal sense, and this is what we mean by the term. However, because it seems to mean something much closer to neopositivism in the education literature, we will use the term *nonpositivism* in its place. By this we mean any view that embraces the heart of the new philosophy of science, that all observation is theory-laden. As we use the term, then, it includes views as diverse as critical theory, pragmatism, and Popperianism.

²We recognize that *interpretivism* is often used in a specialized sense, to indicate an exclusive focus on understanding the insider's perspective. We will use the term in a more expansive sense, to mean roughly the same thing as nonpositivism. We use it in several places instead of nonpositivism to highlight the important implication that all scientific observation, analysis, and theorizing unavoidably involve acts of interpretation by researchers.

³Eisenhart was present at the 1989 Conference on Qualitative Research in Education, held at the University of Georgia, when Roman's work, as well as Naff Cain's, were questioned. Both Roman and Naff Cain have reviewed our interpretation of the responses to their work at that conference.

⁴Even thinkers who insist that the positivist-alternative paradigm debate is a live one deny that the purported incompatibility at the paradigm level carries over to the level of quantitative versus qualitative methods (e.g., Guba, 1987; Smith & Heshusius, 1986).

References

- Bernstein, R. (1983). *Beyond objectivism and relativism*. Philadelphia: University of Pennsylvania Press.
- Clifford, J. (1988). *The predicament of culture: Twentieth-century ethnography, literature, and art*. Cambridge, MA: Harvard University Press.
- Denzin, N. (1989). *The research act: A theoretical introduction to sociological methods* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Erickson, F. (1982). Taught cognitive learning in its immediate environments: A neglected topic in the anthropology of education. *Anthropology and Education Quarterly*, 13(2), 149-180.
- Erickson, F. (1986). Qualitative methods of research on teaching. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 119-161). New York: Macmillan.
- Finn, C. (1988). What ails educational research. *Educational Researcher*, 17(1), 5-8.
- Garnson, J. (1986). Some principles of postpositivistic philosophy of science. *Educational Researcher*, 15(9), 12-18.
- Geertz, C. (1988). *Works and lives: The anthropologist as author*. Stanford: Stanford University Press.
- Goetz, J., & LeCompte, M. (1984). *Ethnography and qualitative design in educational research*. New York: Academic Press.

- Guba, E. (1987) What have we learned about naturalistic evaluation? *Educational Practice*, 8(1), 23-43
- Howe, K. (1985) Two dogmas of educational research *Educational Researcher*, 14(8), 10-18
- Howe, K. (1988). Against the quantitative-qualitative incompatibility thesis (or, dogmas die hard) *Educational Researcher*, 17(8), 10-16.
- Kaplan, A. (1964) *The conduct of inquiry* San Francisco: Chandler
- Lincoln, Y., & Guba, E. (1985) *Naturalistic inquiry*. Beverly Hills, CA Sage Publications
- Lincoln, Y., & Guba, E. (1988, April). *Criteria for assessing naturalistic inquiries as reports*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.
- MacKenzie, B. (1977) *Behaviorism and the limits of scientific method*. Atlantic Highlands, NJ, Humanities Press
- Marcus, G., & Fischer, M. (1986) *Anthropology as cultural critique: An experimental moment in the human sciences*. Chicago: University of Chicago Press
- Naff, B. (1987). *The impact of prescriptive planning models on preservice English teachers' thought and on the classroom environments they create. An ethnographic study* Unpublished Ed D dissertation, Virginia Polytechnic Institute and State University.
- Naff Can, B. (1989, January) *Using low inference data to create ethnographic vignettes* Paper presented at the Qualitative Research in Education Conference, Athens, GA.
- Peshkin, A. (1988) In search of subjectivity—one's own. *Educational Researcher*, 17(7), 17-22
- Phillips, D. (1983). Postpositivistic educational thought *Educational Research*, 12(5), 4-12
- Phillips, D. (1987a) Validity in qualitative research: Why the worry with warrant will not wane *Education and Urban Society*, 20(1), 9-24
- Phillips, D. (1987b). *Philosophy, science, and social inquiry*. New York Pergamon Press
- Rist, R. (1980). Blitzkrieg ethnography On the transformation of a method into a movement *Educational Researcher*, 9(2), 8-10.
- Roman, L. (1989, January) *Double exposure. Politics of feminist research* Paper presented at the Qualitative Research in Education Conference, Athens, GA
- Rorty, R. (1979) *Philosophy and the mirror of nature*. Princeton, NJ: Princeton University Press
- Rorty, R. (1982). Method, social science, social hope. In R. Rorty (Ed.), *Consequences of pragmatism* (pp. 191-210). Minneapolis. University of Minnesota Press
- Scriven, M. (1983). The evaluation taboo. In E. House (Ed.), *Philosophy of evaluation* (pp 75-82) San Francisco. Jossey-Bass
- Scriven, M. (1986). Evaluation as a paradigm for educational research In E. House (Ed.), *New directions in educational evaluation* (pp 53-67) Philadelphia: The Falmer Press.
- Shavelson, R., & Berliner, D. (1988). Erosion of education research infrastructure *Educational Researcher*, 17(1), 9-12.
- Shulman, L. (1988) Disciplines of inquiry in education. An overview. In R. Jaeger (Ed.), *Complementary methods for research in education* (pp. 3-17) Washington, DC. American Educational Research Association.
- Smuth, J. K. (1983a). Quantitative versus qualitative research: An attempt to clarify the issue. *Educational Researcher*, 12(3), 6-13.
- Smuth, J. K. (1983b). Qualitative versus interpretive: An attempt to clarify the issue. In E. House (Ed.), *Philosophy of evaluation* (pp. 27-52). San Francisco: Jossey-Bass.
- Smith, J. K., & Heshusius, L. (1986). Closing down the conversation: The end of the quantitative-qualitative debate among educational researchers *Educational Researcher*, 15(1), 4-12
- Smith, M. L., & Glass, G. V. (1987). *Research and evaluation in education and the social sciences*. Englewood Cliffs, NJ: Prentice-Hall.
- Spradley, J. P. (1980) *Participant Observation*. New York. Holt, Rinehart, and Winston
- Strauss, A. (1987) *Qualitative analysis for social scientists*. Cambridge, MA: Cambridge University Press.
- VanMaanen, J. (1988) *Tales of the field* Chicago: University of Chicago Press
- Wolcott, H. (1980) How to look like an anthropologist without really being one *Practicing Anthropology*, 3(1), 6-7, 56-60

MAKE THE WINNING MOVE . . .

(. . . in item analysis and test scoring)

Move from classical to IRT methods —

- for more effective use of your item pool
- for more reliable scoring
- for easier item replacement and test updating
- for simpler equating of test forms

Move from special-case programs to

BILOG 3®

The IRT program that includes—

- all of the binary logistic models
- all of the important IRT scoring procedures
- information analysis
- the group-level model
- test of fit and residual analysis
- 175-page User's Guide with introduction to Item Response Theory
- User's Hot-Line for BILOG and other item analysis programs
- PC and MVS/CMS mainframe versions

For details, write, phone, or FAX

SCIENTIFIC SOFTWARE, Inc.
 1369 Neitzel Road • Mooresville, IN 46158
 Phone (800)247-6113 • FAX (317)831-8998



(White to mate
in two moves)



"No testing program can afford to be without BILOG."
 —Dr. Richard Hill, President
 Advanced Systems in Measurement and Evaluation