The Qualitative Researcher's Companion

Understanding and Validity in Qualitative Research

This PDF has been generated from SAGE Research Methods. Please note that the pagination of the online version will vary from the pagination of the print book.

[p. 36 ↓]

[p. 37 ↓]

Understanding and Validity in Qualitative Research

> Within the last few years, the issue of validity in qualitative research has come to the fore. (Kvale, 1989, p. 7)

> All field work done by a single field-worker invites the question, Why should we believe it? (Bosk, 1979, p. 193)

Validity has long been a key issue in debates over the legitimacy of qualitative research; if qualitative studies cannot consistently produce valid results, then policies, programs, or predictions based on these studies cannot be relied on. Proponents of quantitative and experimental approaches have frequently criticized the absence of "standard" means of assuring validity, such as quantitative measurement, explicit controls for various validity threats, and the formal testing of prior hypotheses. Their critique has been bolstered by the fact that existing **[p. 38 ↓]** categories of validity (for example, concurrent validity, predictive validity, convergent validity, criterion-related validity, internal/external validity) are based on positivist assumptions that underlie quantitative and experimental research designs (Salner, 1989). Qualitative researchers have generally responded either by denying the relevance of the quantitative or scientific paradigm for what they do (for example, Guba & Lincoln, 1989), or by arguing that qualitative research has its own procedures for attaining validity that are simply different from those of quantitative approaches (for example, Kirk & Miller, 1986).

However, explicit attention to *how* qualitative researchers conceptualize validity issues in their research has been slow to develop. Phillips (1987) and Kvale (1989) have argued that the concept is legitimate and useful in qualitative research; Goetz and LeCompte (1984), Kirk and Miller (1986), and Erickson (1989) have proposed various definitions of validity and of different types of validity. Eisenhart and Howe (1992) also accept the legitimacy of the concept, but argue for a unitary conception of validity rather than a typology; they see research studies as arguments and propose various standards for valid arguments in educational research. In contrast, Guba and Lincoln (1989) view validity as a positivist notion and propose to substitute for this the concept of "authenticity" in qualitative research. Finally, Wolcott (1990a) is skeptical that validity or any analogous concept is legitimate or useful in qualitative inquiry.

Mishler (1990) has recently argued that, while the concept of validity is applicable to what he calls "inquiry-guided" research, the attempt to extend the dominant experimental/quantitative model of validity to such research is misguided, since the dominant model's categories of validity are themselves fundamentally flawed. He asserts that the demonstration by Campbell and Stanley (1963), and later by Cook and Campbell (1979), "that validity assessments are not assured by following procedures but depend on investigators' judgments" has proved "to be a death blow for the typology approach" (Mishler, 1990, p. 418), and argues that for this reason, issues of meaning and interpretation have become central. Mishler proposes a model of validity that relies on exemplars of scientific practice, rather than on abstract rules or categories, as the grounds for validating the trustworthiness of observations, interpretations, and generalizations.[1]

I agree with many of Mishler's arguments and conclusions, in particular with his emphasis on the importance of exemplars (see Maxwell, 1990b; Pitman & Maxwell, 1992). However, I think that his **[p. 39 ↓]** report of the demise of validity typologies is greatly exaggerated. His argument is directed mainly at typologies based on the procedures used for determining validity, which certainly comprise the majority of such typologies. Phillips states what seems to be a consensus: "In general it must be recognized that there are *no* procedures that will regularly (or always) yield either sound data or true conclusions" (1987, p. 21). Brinberg and McGrath make the same point: "Validity is not a commodity that can be purchased with techniques. … Rather,

validity is like integrity, character, and quality, to be assessed relative to purposes and circumstances" (1985, p. 13).

But defining types of validity in terms of procedures, an approach generally labeled instrumentalist or positivist, is not the only approach available. The most prevalent alternative is a realist conception of validity that sees the validity of an account as inherent, not in the procedures used to produce and validate it, but in its relationship to those things that it is intended to be an account *of* (Hammersley, 1992; House, 1991; Maxwell, 1990a,b; Norris, 1983). This article is not a response to or critique of Mishler's approach, but an alternative, complementary view: it presents a realist typology of the kinds of validity that I see as relevant to qualitative research.[2]

In adopting a realist approach to validity, I am in basic agreement with the main point of Wolcott's critique—that is, that *understanding* is a more fundamental concept for qualitative research than validity (1990a, p. 146). I see the types of validity that I present here as derivative from the kinds of understanding gained from qualitative inquiry; my typology of validity categories is also a typology of the kinds of understanding at which qualitative research aims (see Runciman, 1983).

However, in explicating the concept of validity in qualitative research, I want to avoid applying or adapting the typologies developed for experimental and quantitative research, for reasons quite separate from Mishler's disapproval of procedure-based typologies. These typologies cannot be applied directly to qualitative research without distorting what qualitative researchers actually do in addressing validity issues, and tautologically confirming quantitative researchers' critiques.

An illustration of this is Campbell and Stanley's (1963) attack on what they disparagingly refer to as the "one-shot case study." They argue that this design is "well-nigh unethical" on the grounds that a single observation of one group, following an intervention, with no control groups or prior measures, provides no way of discriminating **[p. 40 ↓]** among numerous possible alternative explanations for the outcome, explanations that any valid design must be able to rule out. From an experimentalist's perspective, this argument is perfectly logical, but it completely ignores the ways that qualitative researchers actually rule out validity threats to their conclusions. Campbell later recognized the fallacy in his earlier critique and retracted it, stating that "the intensive … case study has a

The Qualitative Researcher's Companion:
Understanding and Validity in Qualitative Research

discipline and a capacity to reject theories which are neglected in my caricature of the method" (1975, p. 184).

This situation, I believe, is similar to one in the history of the study of kinship terminologies in anthropology. Early investigators of kinship often assumed an equivalent conceptual structure between English and the language of the society they were studying (indeed, they often took the English terms to refer to real, natural categories), and simply sought the equivalents for the English kinship terms in the language being investigated. The major contribution of Lewis Henry Morgan (1871) to the study of kinship, and the basis for nearly all subsequent work on kinship terminology, was the recognition that societies have different classification systems for relatives, systems that can differ markedly from those of our own society and that cannot be represented adequately by a simple translation or correlation of their terminology with that of the English language (Trautmann, 1987, p. 57).

This article is thus intended to be a Morgan-like reformulation of the categorization of validity in qualitative research—an account from "the native's point of view" (Geertz, 1974) of the way qualitative researchers think about and deal with validity in their actual practice. Any account of validity in qualitative research, in order to be productive, should begin with an understanding of how qualitative researchers actually think about validity. I am not assuming that qualitative methods for assessing validity are infallible, but a critique of these methods is beyond the scope of this paper. However, if my account of the categories in which validity is conceived is valid, it obviously has implications for the latter task.

In developing these categories, I will work not only with qualitative researchers' explicit statements about validity—their "espoused theory" (Argyris & Schoen, 1974), or "reconstructed logic" (Kaplan, 1964)—but also with the ideas about validity that seem to me to be implicit in what they actually do—their "theory-in-use" (Argyris & Schoen, 1974) or "logic-in-use" (Kaplan, 1964). I am here following Einstein's advice that

[p. 41 ↓]

> if you want to find out anything from the theoretical physicists about the methods they use, I advise you to stick closely to one principle: Don't

> listen to their words, fix your attention on their deeds. (cited by Manicas, 1987, p. 242)

In this, however, I have the additional advantage that I am myself a qualitative researcher and can draw on my own practice and my understanding of that practice, in the same way that a linguist is able to draw on his or her own "intuitions" about his or her native language in constructing an analysis of that language. (I am not claiming infallibility for such intuitions, but simply acknowledging them as a source of data.)

I do not think that qualitative and quantitative approaches to validity are incompatible. I see important similarities between the two, and think that the analysis I present here has implications for the concept of validity in quantitative and experimental research. I am, however, arguing that a fruitful comparison of the two approaches depends on a prior understanding of each of the approaches in its own terms.

THE NATURE OF VALIDITY IN QUALITATIVE RESEARCH

All qualitative researchers agree that not all possible accounts of some individual, situation, phenomenon, activity, text, institution, or program are equally useful, credible, or legitimate. Furthermore, the ways in which researchers make these discriminations do not pertain entirely to the internal coherence, elegance, or plausibility of the account itself, but often refer to the relationship between the account and something external to it—that is, the phenomena that the account is *about.* Validity, in a broad sense, pertains to this relationship between an account and something outside of that account, whether this something is construed as objective reality, the constructions of actors, or a variety of other possible interpretations.

However, I am not assuming that there is only one correct, "objective" account—what Putnam (1990) refers to as the "God's eye view"—of this realm outside of the account itself. As observers and interpreters of the world, we are inextricably part of it; we cannot step outside our own experience to obtain some observer-independent account

of what we experience. Thus, it is always possible for there to be different, equally valid accounts from different perspectives.[3]

[p. 42 ↓]

My approach, therefore, does not depend on a correspondence theory of truth, at least not in the usual sense of a mirroring or isomorphism between account and reality, a sense that has been criticized by Rorty (1979). My analysis employs a critical realism (Bhaskar, 1989; Campbell, 1988; Hammersley, 1992; Manicas, 1987; Putnam, 1990) that assumes that we can have no direct knowledge of the objects of our accounts and thus no independent entity to which to compare these accounts (see Maxwell, 1990a,b). The applicability of the concept of validity presented here does not depend on the existence of some absolute truth or reality to which an account can be compared, but only on the fact that there exist ways of assessing accounts that do not depend entirely on features of the account itself, but in some way relate to those things that the account claims to be about.[4] This concept specifically differs from positivism and instrumentalism in that it does not take these tests to be critical for validity, but only as fallible means for generating evidence about the relationship between the account and its object (see Cook & Campbell, 1979).

An important point about this approach to validity is that it refers primarily to accounts, not to data or methods. This is consistent with the point made previously that validity is relative to purposes and circumstances. Hammersley and Atkinson (1983, p. 191) state that "data in themselves cannot be valid or invalid; what is at issue are the inferences drawn from them." And a classic work on survey research, after critiquing one study, asks, "May one, therefore, conclude that the Gluecks's measurements of these variables are invalid? In order to answer this question, it is necessary to ask what the Gluecks wish to learn from their data" (Hirschi & Selvin, 1967, p. 195).

It is possible to construe data as a kind of account—a description at a very low level of inference and abstraction. In this sense, it is sometimes legitimate to speak of the validity of data, but this use is derived from the primary meaning of validity as a property of accounts. In contrast, a method by itself is neither valid nor invalid; methods can produce valid data or accounts in some circumstances and invalid ones in others.

Validity is not an inherent property of a particular method, but pertains to the data, accounts, or conclusions reached by using that method in a particular context for a particular purpose. To speak of the validity of a method is simply a shorthand way of referring to the validity of the data or accounts derived from that method.

[p. 43 ↓]

I agree with Mishler that validity is always relative to, and dependent on, some community of inquirers on whose perspective the account is based. Validity is relative in this sense because understanding is relative; as argued above, it is not possible for an account to be independent of any particular perspective. It is always possible to challenge an account from outside that community and perspective, but such a challenge amounts to expanding the community that is concerned with the account and may change the nature of the validity issues in ways to be discussed below.

However, the grounding of all accounts in some particular community and perspective does not entail that all accounts are incommensurable in the sense of not being comparable. Bernstein (1983) argues in detail that the incommensurability thesis has been widely misinterpreted in this way, and that the rejection of objectivism does not require the adoption of extreme relativism in this sense. "What is sound in the incommensurability thesis has *nothing to do* with relativism, or at least that form of relativism which wants to claim that there can be no rational comparison among the plurality of theories" (Bernstein, 1983, p. 92; emphasis in the original). He quotes Winch, one of the authors most often cited in support of incommensurability:

> We should not lose sight of the fact that the idea that men's ideas and beliefs must be checkable by reference to something independent— some reality—is an important one. To abandon it is to plunge straight into a Protagorean relativism, with all the paradoxes that involves. (Winch, 1958, p. 11, cited in Bernstein, 1983, p. 98)

Bernstein claims that incommensurability, properly understood, is not a rejection of comparability, nor an abandonment of any attempt to assess the validity of accounts, but instead offers a way to compare or assess accounts that goes beyond the sterile opposition between objectivism and relativism (see also Bernstein, 1991, pp. 57–78).

I argued above that validity pertains to the kinds of understanding that accounts can embody. I see five broad categories of understanding that are relevant to qualitative research, and five corresponding types of validity that concern qualitative researchers. I will refer to these categories, respectively, as descriptive validity, interpretive validity, theoretical validity, generalizability, and evaluative validity.

The typology presented here has been influenced by others' work, particularly Cook and Campbell (1979), Kirk and Miller (1986), and **[p. 44 ↓]** Erickson (1989). However, my primary debt is to the detailed analysis by Runciman (1983) of the types of understanding involved in social theory (though I depart significantly from his definitions of these), and to the discussion of description, interpretation, and explanation by Kaplan (1964). I believe that the distinctions made by Runciman, Kaplan, and others are simply explicit codifications and elaborations of a widespread commonsense conceptual structure, and that this structure is implicit in the work of many qualitative researchers. My account of validity is an attempt, in part, to explicate this implicit theory-in-use. Specific connections between my categories and those of these authors cited will be discussed below.

Because my analysis of validity is based on this taken-for-granted structure, the distinctions I propose may seem unoriginal or even old-fashioned, and lacking in philosophical sophistication. But philosophical sophistication is of value only when it engages with our ordinary ways of seeing and thinking. Although in this article I draw on a number of philosophical concepts and arguments, my primary purpose is not to advance the philosophical understanding of qualitative research, but to explicate how qualitative researchers think about validity. For this reason, I have not attempted to provide a detailed philosophical justification for the positions I have taken (for some of these justifications, see Bernstein, 1983, 1991; Bohman, 1991; Lakoff, 1987; Manicas, 1987; Putnam, 1990), nor to cite all of the relevant philosophical literature. Clearly, my assumptions and arguments are open to philosophical critique. However, as I argue for accounts in general, the validity of the account I provide here should be evaluated not simply on its internal logic and coherence, but also on its relationship to what qualitative researchers actually do in their research.

In addition, I am not arguing that the categories I define are clearly and explicitly demarcated or that every instance of a validity concern falls neatly into one and only

one category. The entire approach to categorization that depends on precise and uniform criteria for determining category boundaries and assigning membership has been undermined by recent research (Lakoff, 1987), and it has become increasingly apparent that ambiguity and fuzzy boundaries are the rule rather than the exception in categorization. I will discuss specific instances of such fuzziness below; here I want to emphasize that I do not accept such cases as evidence *per se* for the inadequacy of the typology. Instances that do not fit my categories *may*, of course, **[p. 45 ↓]** challenge the validity of these categories, but they do so as a result of their implications, not simply because they cannot be assigned unambiguously to a single category.

DESCRIPTIVE VALIDITY

The first concern of most qualitative researchers is with the factual accuracy of their account—that is, that they are not making up or distorting the things they saw and heard. If you report that an informant made a particular statement in an interview, is this correct? Did he or she really make that statement, or did you mis-hear, mis-transcribe, or mis-remember his or her words? Did a particular student in a classroom throw an eraser on a specific occasion? These matters of descriptive accuracy are emphasized by almost every introductory qualitative methods textbook in its discussion of the recording of field notes and interviews.

All of the subsequent validity categories I will discuss are dependent on this primary aspect of validity. As Geertz puts it, "behavior must be attended to, and with some exactness, because it is through the flow of behavior—or, more precisely, social action—that cultural forms find articulation" (1973, p. 17). Wolcott, similarly, states that "description is the foundation upon which qualitative research is built" (1990b, p. 27) and that "whenever I engage in fieldwork, I try to record as accurately as possible, and in precisely their words, what I judge to be important of what people do and say" (Wolcott, 1990a, p. 128).

I will refer to this first type of validity as descriptive validity; it corresponds, to some extent, to the category of understanding that Runciman (1983) calls "reportage" or "primary understanding." Insofar as this category pertains to humans, it refers to what Kaplan (1964, p. 358) calls "acts" rather than "actions"—activities seen as physical and

SAGE researchmethods

behavioral events rather than in terms of the meanings that these have for the actor or others involved in the activity.

The above quotes refer mainly to what I will call primary descriptive validity: the descriptive validity of what the researcher reports having seen or heard (or touched, smelled, and so on). There is also the issue of what I will call secondary descriptive validity: the validity of accounts of things that could in principle be observed, but that were inferred from other data—for example, things that happened in the **[p. 46 ↓]** classroom when the researcher was not present. (This secondary description is also included in Runciman's concept of "reportage.") Secondary descriptive validity can pertain to accounts for which the inference is highly complex and problematic: for example, the claim that the person known as William Shakespeare actually wrote Hamlet, or that a particular stone object was used as a cutting tool by a member of an early human population. These issues concern descriptive validity because they pertain to physical and behavioral events that are, in principle, observable.

There are several characteristics of these sorts of descriptive concerns that I want to emphasize. First, they all refer to specific events and situations. No issue of generalizability or representativeness is involved. Second, they are all matters on which, in principle, intersubjective agreement could easily be achieved, given the appropriate data. For example, a tape recording of adequate quality could be used to determine if the informant made a particular statement during the interview, a videotape could be used to decide if the student threw the eraser, and so on.

Put another way, the *terms* of the description (for example, "throw" in the example above) are not problematic for the community involved in the discussion of the event; their meaning—how they ought to be applied to events and actions—is not in dispute, only the accuracy of the application. This situation is quite different, for example, from the case of an account claiming that a student *assaulted* another student. In this case, it is possible that no amount of videotape or other data could resolve disagreements about the applicability of the term "assault" to the action that took place. This latter dispute is not about descriptive validity, but about the interpretive, theoretical, and/or evaluative validity of the account.

Descriptive validity is by no means independent of theory; all observation and description are based on theory, even if this theory is implicit or common sense. However, descriptive validity is free from *disagreement* about the theory in question. This assertion does not mean that there cannot be disagreement about the descriptive validity of an account, only that such disagreement could in principle be resolved by the appropriate data. Of course, the theory could be *made* problematic by one of the participants in the discussion—for example, by challenging the applicability of "throwing" to what the student did with the eraser. However, this challenge would change the nature of the validity **[p. 47 ↓]** questions involved and make them no longer an issue of descriptive validity for participants in that discussion.

In framing descriptive validity in this way, I am not seeking to revive the positivist view that all disagreements in science ought to be resolvable in principle by means of the appropriate evidence. In my opinion, this view has been convincingly criticized by Kuhn (1970), Rorty (1979), Bernstein (1983), and others. Instead, I am attempting to incorporate into my typology one of Kuhn's fundamental insights: that in normal practice, many scientific disagreements *are* resolved in this way, and that incommensurability becomes crucial mainly in times of scientific crisis. Descriptive understanding pertains, by definition, to matters for which we have a framework for resolving such disagreements, a framework provided in large part by taken-for-granted ideas about time, space, physical objects, behavior, and our perception of these. Raising questions about the definition or applicability of these categories changes the type of validity at issue from descriptive to theoretical, in particular to that aspect of theoretical validity generally known as construct validity.

Descriptive validity can refer to issues of omission as well as of commission; no account can include everything, and "accuracy is a criterion relative to the purposes for which it is sought" (Runciman, 1983, p. 97). For example, a verbatim interview transcript might be descriptively invalid in omitting features of the informant's speech, such as stress and pitch, that are essential to the understanding of the interview. The omission of things that participants in the discussion feel are significant to the account (for the purposes at issue) threatens the descriptive validity of that account.

Descriptive validity can also pertain to statistically descriptive aspects of accounts. A claim that a certain phenomenon was frequent, typical, or rare in a specific situation

at the time it was observed—for example, that few students raised their hands in response to the teacher's question—is also subject to threats to descriptive validity. This is an issue for which Becker (1970) has advocated the use of what he calls "quasi-statistics"—simple counts of things to support claims that are implicitly quantitative. What makes this a matter of descriptive validity is that it does not involve statistical *inference* to some larger universe than the phenomenon directly studied, but only the numerical description of the specific object of study. This is different from what Cook and Campbell (1979) call "statistical conclusion validity," **[p. 48 ↓]** which refers to the validity of inferences from the data to some population. I treat the latter issue below, as one type of generalizability.

Reliability, in my view, refers not to an aspect of validity or to a separate issue from validity, but to a particular type of threat to validity. If different observers or methods produce descriptively different data or accounts of the same events or situations, this puts into question the descriptive validity (and other types of validity as well) of the accounts. This problem could be resolved either by modification of the accounts, so that different observers come to agree on their descriptive accuracy, or by ascertaining that the differences were due to differences in the perspective and purposes of the observers and were both descriptively valid, given those perspectives and purposes.

INTERPRETIVE VALIDITY

However, qualitative researchers are not concerned solely, or even primarily, with providing a valid description of the physical objects, events, and behaviors in the settings they study; they are also concerned with what these objects, events, and behaviors *mean* to the people engaged in and with them. In this use of the term meaning, I include intention, cognition, affect, belief, evaluation, and anything else that could be encompassed by what is broadly termed the "participants' perspective," as well as communicative meaning in a narrower sense. This construction is inherently ideational or mental, rather than physical, and the nature of the understanding, validity, and threats to validity that pertain to it are significantly different from those involved in descriptive validity.

I will call this sort of understanding interpretive, and the type of validity associated with it interpretive validity, following Erickson (1989).[5] The term "interpretive" is appropriate primarily because this aspect of understanding is most central to interpretive research, which seeks to comprehend phenomena not on the basis of the researcher's perspective and categories, but from those of the participants in the situations studied —that is, from an "emic" rather than an "etic" perspective (Bohman, 1991; Headland, Pike, & Harris, 1990). In contrast to descriptive validity, which could apply equally well to quantitative and qualitative research, interpretive validity has no real counterpart in quantitative-experimental validity typologies.

[p. 49 ↓]

Thus, while the terms involved in descriptive validity can be either etic or emic, interpretive validity necessarily pertains to aspects of an account for which the terms are emic. This is because, while accounts of physical and behavioral phenomena can be constructed from a variety of perspectives, accounts of meaning must be based initially on the conceptual framework of the people whose meaning is in question. These terms are often derived to a substantial extent from the participants' own language. The terms are also necessarily, to use Geertz's phrase (1974), "experience-near"—based on the immediate concepts employed by participants (for example, "love"), rather than on theoretical abstractions (for example, "object cathexis").

Like descriptive validity, then, interpretive validity, while not at he-oretical, refers to aspects of accounts for which the terms of the account are not themselves problematic. Interpretive accounts are grounded in the language of the people studied and rely as much as possible on their own words and concepts.[6] The issue, again, is not the appropriateness of these concepts for the account, but their accuracy as applied to the perspective of the individuals included in the account. For example, was the teacher, in yelling at the student for throwing the eraser, really "mad" at the student, or just trying to "get control" of the class? While the relevant consensus about the categories used in description rests in the research community, the relevant consensus for the terms used in interpretation rests to a substantial extent in the community studied.

Unlike descriptive validity, however, for interpretive validity there is no in-principle access to data that would unequivocally address threats to validity. Interpretive validity is inherently a matter of inference from the words and actions of participants in the situations studied. The development of accounts of these participants' meanings is usually based to a large extent on the participants' own accounts, but it is essential not to treat these latter accounts as incorrigible; participants may be unaware of their own feelings or views, may recall these inaccurately, and may consciously or unconsciously distort or conceal their views. Accounts of participants' meanings are never a matter of direct access, but are always *constructed* by the researcher(s) on the basis of participants' accounts and other evidence.[7]

The realist approach to validity that I am adopting here has been held by some interpretive researchers to be incompatible with a concern for interpretive understanding. For example, Lincoln has argued that **[p. 50 ↓]** "critical realism's assumption that there is a singular reality 'out there' … ignores the issue of whether that reality is recognized or rejected by those who may be disadvantaged by that construction" (1990, p. 510).

This critique misses the point that the meanings and constructions of actors are part of the reality that an account must be tested against in order to be interpretively as well as descriptively valid. Social theorists generally agree that any valid account or explanation of a social situation must respect the perspectives of the actors in that situation, although it need not be centered on that perspective (Bohman, 1991; Harre, 1978; Menzel, 1978). My inclusion of interpretive validity in this typology is a recognition of this consensus: that a key part of the realm external to an account is the perspective of those actors whom the account is about (see House, 1991).

Interpretive validity does not apply only to the conscious concepts of participants; it can also pertain to the unconscious intentions, beliefs, concepts, and values of these participants, and to what Argyris and Schoen (1974) call "theory-in-use," as opposed to "espoused theory." However, this aspect of interpretive validity also raises another category of understanding and validity, which, following Kirk and Miller (1986), I will call "theoretical validity."

THEORETICAL VALIDITY

The two previous types of understanding have a number of similarities. First, they depend on a consensus within the relevant community about how to apply the concepts and terms used in the account; any disagreements refer to their accuracy, not their meaning. Second, and closely connected to the first, the concepts and terms employed are "experience-near," in Geertz's sense (1974).

There are two major differences between theoretical understanding and the two types discussed previously. The first is the degree of abstraction of the account in question from the immediate physical and mental phenomena studied. The reason for calling this sort of understanding theoretical is that it goes beyond concrete description and interpretation and explicitly addresses the theoretical constructions that the researcher brings to, or develops during, the study.

This theory can refer to either physical events or mental constructions. It can also incorporate participants' concepts and theories, but its **[p. 51 ↓]** purpose goes beyond simply describing these participants' perspectives. This distinction comprises the second major difference between the theoretical validity of an account and the descriptive or interpretive validity of the same account: theoretical understanding refers to an account's function as an *explanation*, as well as a description or interpretation, of the phenomena.

Theoretical validity thus refers to an account's validity as a *theory* of some phenomenon. Any theory has two components: the concepts or categories that the theory employs, and the relationships that are thought to exist among these concepts. Corresponding to these two aspects of a theory are two aspects of theoretical validity: the validity of the concepts themselves as they are applied to the phenomena, and the validity of the postulated relationships among the concepts. The first refers to the validity of the blocks from which the researcher builds a model, as these are applied to the setting or phenomenon being studied; the second refers to the validity of the way the blocks are put together, as a theory of this setting or phenomenon.

For example, one could label the student's throwing of the eraser as an act of resistance, and connect this act to the repressive behavior or values of the teacher, the social structure of the school, and class relationships in U.S. society. The identification of the throwing as "resistance" constitutes the application of a theoretical construct to the descriptive and interpretive understanding of the action; the connection of this to other aspects of the participants, the school, or the community constitutes the postulation of theoretical relationships among these constructs.

The first of these aspects of theoretical validity closely matches what is generally known as construct validity, and is primarily what Kirk and Miller (1986) mean by theoretical validity. The second aspect includes, but is not limited to, what is commonly called internal or causal validity (Cook & Campbell, 1979); it corresponds to what Runciman calls "explanation" and in part to what Erickson calls "critical validity."[8] This second aspect is not limited to causal validity because theories or models can be developed for other things besides causal explanation—for example, for semantic relationships, narrative structure, and so on—that nevertheless go beyond description and interpretation. Theories can, and usually do, incorporate both descriptive and interpretive understanding, but in combining these they necessarily transcend either of them.

[p. 52 ↓]

What counts as theoretical validity, rather than descriptive or interpretive validity, depends on whether there is consensus within the community concerned with the research about the terms used to characterize the phenomena. Issues of descriptive and interpretive validity focus on the accuracy of the application of these terms (Did the student really throw the eraser? Was the teacher really angry?) rather than their appropriateness (Does what the student did count as resistance?). Theoretical validity, in contrast, is concerned with problems that do not disappear with agreement on the "facts" of the situation; the issue is the legitimacy of the application of a given concept or theory to established facts, or indeed whether any agreement can be reached about what the facts are.

The distinction between descriptive or interpretive and theoretical validity is not an absolute, because (contrary to the assumptions of positivism) objective "sense data"

The Qualitative Researcher's Companion:
Understanding and Validity in Qualitative Research

SAGE research methods

SAGE Research Methods

that are independent of the researcher's perspective, purposes, and theoretical framework do not exist. My distinction between the two types is not based on any such assumption, but on the presence or absence of agreement within the community of inquirers about the descriptive or interpretive terms used. Any challenge to the meaning of the terms, or the appropriateness of their application to a given phenomenon, shifts the validity issue from descriptive or interpretive to theoretical.

These three types of understanding and validity are the ones most directly involved in assessing a qualitative account as it pertains to the actual situation on which the account is based. There are, however, two additional categories of validity issues that I want to raise. The first of these deals with the generalizability of an account, or what is often labeled external validity; the second pertains to the evaluative validity of an account.

GENERALIZABILITY

Generalizability refers to the extent to which one can extend the account of a particular situation or population to other persons, times, or settings than those directly studied. This issue plays a different role in qualitative research than it does in quantitative and experimental research, because qualitative studies are usually not designed to allow systematic generalizations to some wider population. Generalization **[p. 53 ↓]** in qualitative research usually takes place through the development of a theory that not only makes sense of the particular persons or situations studied, but also shows how the same process, in different situations, can lead to different results (Becker, 1990, p. 240). Generalizability is normally based on the assumption that this theory may be useful in making sense of similar persons or situations, rather than on an explicit sampling process and the drawing of conclusions about a specified population through statistical inference (Yin, 1984).

This is not to argue that issues of sampling, representativeness, and generalizability are unimportant in qualitative research. They are crucial whenever one wants to draw inferences from the actual persons, events, or activities observed to other persons, events, or situations, or to these at other times than when the observation was done. (The particular problems of interviewing will be dealt with below.) Qualitative research

almost always involves some of this sort of inference because it is impossible to observe everything, even in one small setting. The sort of sampling done in qualitative research is usually "purposeful" (Patton, 1990) or "theoretical" (Strauss, 1987) sampling, rather than random sampling or some other method of attaining statistical representativeness. The goal of the former types of sampling is twofold: to make sure one has adequately understood the variation in the phenomena of interest in the setting, and to test developing ideas about that setting by selecting phenomena that are crucial to the validity of those ideas.

In qualitative research, there are two aspects of generalizability: generalizing within the community, group, or institution studied to persons, events, and settings that were not directly observed or interviewed; and generalizing to other communities, groups, or institutions. I will refer to the former as internal generalizability, and to the latter as external generalizability. The distinction is analogous to Cook and Campbell's (1979) distinction in quasi-experimental research between statistical conclusion validity and external validity. This distinction is not clear-cut or absolute in qualitative research. A researcher studying a school, for example, can rarely visit every classroom, or even gain information about these classrooms by other means, and the issue of whether to consider the generalizability of the account for those unstudied classrooms internal or external is moot. However, it is important to be aware of the extent to which the times and places observed may differ from those that were not observed, either because of sampling or because of the effect of the observation itself.[9]

[p. 54 ↓]

Internal generalizability in this sense is far more important for most qualitative researchers than is external generalizability because qualitative researchers rarely make explicit claims about the external generalizability of their accounts. Indeed, the value of a qualitative study may depend on its lack of external generalizability in a statistical sense; it may provide an account of a setting or population that is illuminating as an extreme case or "ideal type." Freidson, discussing his qualitative study of a medical group practice, notes that

> there is more to truth or validity than statistical representativeness. …
> In this study I am less concerned with describing the range of variation

SSAGE researchmethods

> than I am with describing in the detail what survey questionnaire
> methods do not permit to be described—the assumptions, behavior,
> and attitudes of a very special set of physicians. They are interesting
> *because* they were special. (1975, pp. 272–273)

He argues that his study makes an important contribution to theory and policy precisely because this was a group for whom social controls on practice should have been most likely to be effective. The failure of such controls in this case not only elucidates a social process that is likely to exist in other groups, but also provides a more persuasive argument for the unworkability of such controls than would a study of a "representative" group.

Interviewing poses some special problems for internal generalizability because the researcher usually is in the presence of the person interviewed only briefly, and must necessarily draw inferences from what happened during that brief period to the rest of the informant's life, including his or her actions and perspectives. An account based on interviews may be descriptively, interpretively, and theoretically valid as an account of the person's actions and perspective in that interview, but may miss other aspects of the person's perspectives that were not expressed in the interview, and can easily lead to false inferences about his or her actions outside the interview situation. Thus, internal generalizability is a crucial issue in interpreting interviews, as is widely recognized,

for example, by Dean and Whyte (1958) and Dexter (1970).[10] The interview is a social situation and inherently involves a relationship between the interviewer and the informant. Understanding the nature of that situation and relationship, how it affects what goes on in the interview, and how the informant's actions and views could differ in other situations is **[p. 55 ↓]** crucial to the validity of accounts based on interviews (Briggs, 1986; Mishler, 1986).

EVALUATIVE VALIDITY

Beyond all of the validity issues discussed above are validity questions about such statements as, "The student was wrong to throw the eraser at the teacher," or, "The teacher was illegitimately failing to recognize minority students' perspectives." This

The Qualitative Researcher's Companion:
Understanding and Validity in Qualitative Research

aspect of validity differs from the types discussed previously in that it involves the application of an evaluative framework to the objects of study, rather than a descriptive, interpretive, or explanatory one. It corresponds to Runciman's "evaluation" as a category of understanding (1983), and is an important component of what Erickson (1989) terms "critical validity."

I have little to say about evaluative validity that has not been said more cogently by others. In raising it here, my purpose is twofold: to acknowledge evaluative validity as a legitimate category of understanding and validity in qualitative research, and to suggest how it relates to the other types of validity discussed. Like external generalizability, evaluative validity is not as central to qualitative research as are descriptive, interpretive, and theoretical validity: many researchers make no claim to evaluate the things they study. Furthermore, issues of evaluative understanding and evaluative validity in qualitative research do not seem to me to be intrinsically different from those in any other approach to research; debates about whether the student's throwing of the eraser was legitimate or justified do not depend on the methods used to ascertain that this happened or to decide what interpretive or theoretical sense to make of it, although they do depend on the particular description, interpretation, or theory one constructs. To raise questions about the evaluative framework implicit in an account, however, as many critical theorists do, *creates* issues of an account's evaluative validity, and no account is immune to such questions.

IMPLICATIONS

I have presented a model of the types of validity that I believe are relevant to, and often implicit in, qualitative research. I have approached **[p. 56 ↓]** this task from a realist perspective, and have argued that this realist approach, which bases validity on the kinds of understanding we have of the phenomena we study, is more consistent and productive than prevailing positivist typologies based on research procedures. A realist view of validity both avoids the philosophical and practical difficulties associated with positivist approaches and seems to me to better represent what qualitative researchers actually do in assessing the validity of their accounts.

Page 22 of 35

The Qualitative Researcher's Companion:
Understanding and Validity in Qualitative Research
$SAGE research methods

SAGE Research Methods

However, having presented this typology, I must add that validity categories are of much less direct use in qualitative research than they are (or are assumed to be) in quantitative and experimental research. In the latter, threats to validity are addressed in an anonymous, generic fashion by prior design features (such as randomization and controls) that can deal with both anticipated and unanticipated threats to validity. In qualitative research, however, such prior elimination of threats is less possible, both because qualitative research is more inductive and because it focuses primarily on understanding particulars rather than generalizing to universale (Erickson, 1986). Qualitative researchers deal primarily with specific threats to the validity of particular features of their accounts, and they generally address such threats by seeking evidence that would allow them to be ruled out. In doing this, they are using a logic similar to that of quasi-experimental researchers such as Cook and Campbell (1979).

This strategy of addressing particular threats to validity, or alternative hypotheses, *after* a tentative account has been developed, rather than by attempting to eliminate such threats through prior features of the research design, is in fact more fundamental to scientific method than is the latter approach (Campbell, 1988; Platt, 1964). This method is accepted by qualitative researchers from a wide range of philosophical positions (for example, Eisner, 1991; Hammersley & Atkinson, 1983; Miles & Huberman, 1984; Patton, 1990). Its application to causal inference has been labeled the "modus operandi" approach by Scriven (1974), but the method has received little formal development in the qualitative research literature, although it is implicit in many substantive qualitative studies.

Thus, researchers cannot use the typology presented here to eliminate, directly and mechanically, particular threats to the validity of their accounts. Qualitative researchers already have many methods for addressing validity threats, and, although there are ways that the state **[p. 57 ↓]** of the art could be improved (see Eisenhart & Howe, 1992; Miles & Huberman, 1984; Wolcott, 1990a), that is not my main goal here. Instead, I am trying to clarify the validity concepts that many qualitative researchers are using —explicitly or implicitly—in their work, to tie these concepts into a systematic model, and to reduce the discrepancy between qualitative researchers' "logic-in-use" and their "reconstructed logic" (Kaplan 1964, pp. 3–11)—a discrepancy that I think has caused both substantial misunderstanding of qualitative research and some shortcomings in its validation practices. I see this typology as being useful both as a checklist of the

$SAGE research methods

kinds of threats to validity that one needs to consider and as a framework for thinking about the nature of these threats and the possible ways that specific threats might be addressed.

I do not see the typological framework presented in this article as antithetical to the exemplar-based approach that Mishler has advocated. In fact, one of my main assumptions is that category-based and context-based approaches to qualitative research in general, and to validity in particular, are both legitimate, and that these are compatible and complementary, rather than competing, alternatives (Maxwell & Miller, n.d.). The ways in which these two approaches could be used in combination is a topic beyond the scope of this article, but I hope that the analysis I have presented is helpful in facilitating this rapprochement.

NOTES

1. The difference between approaches to validity based on exemplars and those based on typologies can be seen as one example of the distinction between syntagmatic (contextualizing or contiguity-based) and paradigmatic (categorizing or similarity-based) strategies (Bruner, 1986; Jakobson, 1956; Maxwell & Miller, n.d.). It is understandable that Mishler prefers a syntagmatic model for validity, since his overall orientation to research is primarily syntagmatic rather than paradigmatic in its emphasis on contextual and narrative analysis rather than categorization and comparison (Mishler, 1986). My approach, in contrast, emphasizes the complementarity of syntagmatic and paradigmatic strategies.

2. The instrumentalist approach to validity is simply one instance of the positivist or logical empiricist program of substituting logical constructions, based on research operations and the sense data that they generate, for inferred (that is, theoretical) entities (see Hunt, 1991; Manicas, 1987, p. 271; Phillips, 1990). As Norris (1983) points out, most approaches to construct validation have combined positivist and realist assumptions without recognizing the problems that this inconsistency creates. One influential validity typology that **[p. 58 ↓]** is predominantly (and explicitly) realist is that of Cook and Campbell (1979), but even this has some residual positivist features (Maxwell, 1990a).

3. For a more general critique of objectivist views, see Bernstein (1983), Bohman (1991), Hammersley and Atkinson (1983), and Lakoff (1987).

4. To apply the distinction introduced previously, between paradigmatic (similarity-based) and syntagmatic (contiguity-based) relationships, I am conceptualizing the relationship between an account and its object as based not on similarity or resemblance (the traditional correspondence theory), but on contiguity—on the implications and consequences of adopting and acting on a particular account. This approach obviously resembles "pragmatist" approaches in philosophy (Kaplan, 1964, pp. 42–46; Rorty, 1979), and is analogous to Helmholtz's (1971) view of experience as a sign rather than an image or reflection of the world (Manicas, 1987, pp. 176ff.); however, a discussion of these connections is beyond the scope of this article. I have attempted to explore some of the implications of such a "non-reflectionist" (McKinley, 1971) view elsewhere (Maxwell, 1979, 1986).

5. As Eisner (1991, p. 35) notes, the term "interpretive" has two meanings in qualitative research. One is the meaning I am adopting here; the other refers to studies that attempt to explain, as well as describe, the things that they study. The latter use is similar to that of Merriam (1988, pp. 27–28) and Patton (1990, p. 423), who use "descriptive" for studies that do not attempt to develop or apply explicit theory and reserve "interpretive" for studies that generate theory or interpret the data from some theoretical perspective. This second use of "interpretive" corresponds to the type of understanding that I term "theoretical." Kaplan's distinction between "semantic explanation," or interpretation, and "scientific explanation" is similar in some ways to my distinction between interpretation and theory, but my use of "theory" includes semantic as well as explicitly explanatory theories.
My use of "interpretation" corresponds, confusingly, to what Runciman (1983) calls "description." The latter term may be related to Geertz's (1973) phrase "thick description," which has been widely employed in discussions of interpretive research. Thick description, for Geertz, is *meaningful* description—that is, the description embedded in the cultural framework of the actors; the term does not refer to the richness or detail of the account. This point is often misunderstood, in part because Geertz uses the phrase precisely to avoid making the distinction between descriptive and interpretive understanding that I have drawn here: between physical/behavioral description, on the one hand, and inference to meaning as a mental phenomenon on

the other. Thus, "thick description" pertains to interpretive as well as to descriptive understanding, as I have defined these terms.

It is ironic that Geertz adopted this term, and its associated philosophical argument, as a characterization of interpretive research, because Gilbert Ryle, who coined the term in his work *The Concept of Mind* (1949), used it as part of an explicit attempt to eliminate mental concepts (what he referred to as "the ghost in the machine") from philosophy and to replace them with dispositional statements referring to an individual's propensity to behave in particular ways. This approach, which came to be known as "logical behaviorism," was a classically positivist strategy of replacing theoretical entities with logical constructions based on observables. (As noted above, this strategy was the Achilles' heel of positivism.)

[p. 59 ↓] This positivist view that mental constructs are theoretical abstractions that ultimately *refer to* behavior and behavioral dispositions is quite different from the realist position I take here, that such mental constructs refer to unobservable but real entities whose existence is *inferred* from observations of behavior (see Manicas, 1987).

Within the category of interpretive understanding, it is possible to make a distinction, similar to that between primary and secondary descriptive understanding, between the *communicative* meaning of speech or actions (which is nonetheless always meaning for some actor or interpreter) and the actor's subjective intentions, beliefs, values, and perspective (see Gellner, 1962; Hannerz, 1992, pp. 3–4; Keesing, 1987, pp. 174–175; Ricoeur, 1981). Both of these types of understanding are ultimately based on inferences from the descriptive evidence, but the validity of inferences to the actor's subjective states depends on the validity of the researcher's account of the meaning of the actor's words and actions. My own alternative to Geertz's attempt to get meaning out of the "secret grotto in the skull" and into the public world rests on this distinction between meaning as a property of discourse, on the one hand, and the actor's subjective states, on the other.

6. In providing a valid account of individuals who lack such an accessible language, such as preverbal children, interpretive validity merges with the following category, theoretical validity.

7. I cannot deal systematically here with one challenge to this approach embodied in the poststructuralist slogan that "there is nothing outside the text." I agree with Manicas that this view represents not just a repudiation of realist conceptions of validity, but

"an epistemological nihilism in which truth is an illusion" (1987, p. 269). I would also argue that, ironically, this approach exemplifies the same goal of eliminating inferred entities that characterized positivism and is vulnerable to the same critiques that led to positivism's demise.

I want to emphasize that: my distinction between descriptive and interpretive validity is not between the "real world" and actors' constructions of that world. First, both descriptive and interpretive understanding pertain to the researcher's *accounts* of the world—that is, to accounts of its physical/behavioral and mental aspects or components, respectively. Both types of accounts are the researcher's constructions. Second, the physical and mental components refer to entities that are equally real, rather than one being a reflection of the other. (For a more detailed analysis of the relationship between the mental and physical frameworks, see Maxwell, 1986.)

8. Runciman, however, places explanation *before* interpretation (his "description") in his typology of kinds of understanding, and argues more generally that "there is no special problem of explanation in the human sciences, but only a special problem of description" (1983, p. 1). This is my main disagreement with Runciman's account of the different types of understanding involved in social theory. In attempting to justify his claim that explanation can be based solely on reportage, without inquiring what he calls descriptive understanding (what I call "interpretation"), Runciman is forced to include a substantial amount of interpretation in reportage, such as inferences from behavior to the mental states of actors. This view, that valid explanations of behavior can generally be formulated prior to, and independently of, an adequate understanding of the perspective of the actors, has been criticized by Bohman (1991), MacIntyre (1967), and Menzel (1978).

[p. 60 ↓] My distinction between interpretation and explanation is quite different from that of Ricoeur (1981), who contrasts the explanation of a text in terms of its internal structure (an approach based on linguistics rather than the natural sciences) with its interpretation in terms of its connection to the world outside of the text (including the reader). I cannot deal here with this alternative, nor with the complex and much-debated issue of what constitutes an explanation (see Bohman, 1991; Kitcher & Salmon, 1989; Salmon, 1984). This current debate over explanation is one of the reasons I have chosen to call this type of understanding "theoretical" rather than "explanatory."

9. The difference between secondary descriptive validity and internal generalizability deserves clarification, because the two seem superficially similar. Both refer to the extension of one's account to things that were not directly observed but that remain within the setting or group studied. The difference between the two is based on the kind of relationship postulated between the immediate data or account and the claim whose validity is in question. In internal (as well as external) generalizability, the claim is that those things not directly observed are *similar* to those described in the account; that the account can be *generalized* to some wider context. For secondary descriptive validity, the issue is not similarity, but the validity of the chain of *inference* from one's primary data to things that were not directly observed—for example, whether one can infer, from an eraser on the floor, a chalk mark on the wall above it, and a student's sullen silence about what had happened when the researcher briefly left the room, that an eraser was thrown. There is no assumption that the primary data resemble or generalize to the secondary conclusion in any way, only that the inferential connection between the two is valid. In addition, internal (and external) generalizability can pertain to interpretive, theoretical, or evaluative conclusions, as well as descriptive ones.

10. I have indicated earlier my disagreement with the approach that deals with this problem by denying it—that is, by treating the interview (or even the interview transcript) as a "text" and asserting that it is illegitimate to attempt to make inferences to some "real" actor.

Joseph A.Maxwell

REFERENCES

Argyris, C., & Schoen, D. A. (1974). Theory in practice: Increasing professional effectiveness. San Francisco: Jossey-Bass.

Becker, H. S. (1970). Problems of inference and proof in participant observation. In H. S. Becker (Ed.), Sociological work: Method and substance (pp. 25–38). New Brunswick, NJ: Transaction Books.

Becker, H. S. (1990). Generalizing from case studies. In E. W. Eisner, ed. & A. Peshkin (Eds.), Qualitative research in education: The continuing debate (pp. 233–242). New York: Teachers College Press.

Bernstein, R. J. (1983). Beyond objectivism and relativism. Philadelphia: University of Pennsylvania Press.

Bernstein, R. J. (1991). The new constellation: The ethical-political horizons of modernity/post-modernity. Cambridge, MA: MIT Press.

Bhaskar, R. (1989). Reclaiming reality. London: Verso.

Bohman, J. (1991). New philosophy of social science. Cambridge, MA: MIT Press.

Bosk, C. (1979). Forgive and remember: Managing medical failure. Chicago: University of Chicago Press.

Briggs, C. (1986). Learning how to ask: A sociolinguistic appraisal of the role of the interview in social science research. Cambridge, Eng.: Cambridge University Press.

Brinberg, D., & McGrath, J. E. (1985). Validity and the research process. Beverly Hills, CA: Sage.

Bruner, J. (1986). Two modes of thought. In J. Bruner (Ed.), Actual minds, possible worlds (pp. 11–43). Cambridge, MA: Harvard University Press.

Campbell, D. T. "Degrees of freedom" and the case study. Comparative Political Studies vol. 8 (2) (1975). pp. 178–193. (Reprinted in Campbell, 1988)

Campbell, D. T. (1988). Methdology and epistemology for social science: Selected papers. Chicago: University of Chicago Press.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), Handbook of research on teaching (pp. 171–246). Chicago: Rand McNally.

SSAGE researchmethods

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

Dean, J. and Whyte, W. F. How do you know if the informant is telling the truth? Human Organization vol. 17 (2) (1958). pp. 34–38. (Reprinted in L. A. Dexter, 1970, pp. 119–131)

Dexter, L. A. (1970). Elite and specialized interviewing. Evanston, IL: Northwestern University Press.

Eisenhart, M., & Howe, K. (1992). Validity in educational research In M. D. LeCompte, ed. , W. L. Millroy,, ed. & J. Preissle (Eds.), The handbook of qualitative research in education (pp. 643–680). San Diego: Academic Press.

Eisner, E. W. (1991). The enlightened eye: Qualitative inquiry and the enhancement of educational practice. New York: Macmillan.

Erickson, F. (1986). Qualitative methods in research on teaching. In M. C. Wittrock (Ed.), Handbook of research on teaching (3rd ed.). New York: Macmillan.

Erickson, F. (1989, March). The meaning of validity in qualitative research. Unpublished paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Freidson, E. (1975). Doctoring together: A study of professional social control. Chicago: University of Chicago Press.

Geertz, C. (1973). Thick description: Toward an interpretive theory of culture. In C. Geertz, ed. , The interpretation of cultures (pp. 3–30). New York: Basic Books.

Geertz, C. "From the native's point of view": On the nature of anthropological understanding. Bulletin of the American Academy of Arts and Sciences vol. 28 (1) (1974). pp. 26–45. (Reprinted in K.H. Basso & H.A. Selby (Eds.) (1976), Meaning in anthropology, pp. 221–237. Albuquerque: University of New Mexico Press)

SAGE researchmethods

Gellner, E. (1962). Concepts and society. Transactions of the Fifth World Congress of Sociology. Vol. 1 (pp. 153–183). Louvain, Belg.: International Sociological Association. (Reprinted in B. R. Wilson, Ed., 1970, Rationality, pp. 18–49. London: Harper & Row)

Goetz, J. P., & LeCompte, M. D. (1984). Ethnography and qualitative design in educational research. San Diego: Academic Press.

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

Hammersley, M. (1992). What's wrong with ethnography? London: Routledge.

Hammersley, M., & Atkinson, P. (1983). Ethnography: Principles in practice. London: Tavistock.

Hannerz, U. (1992). Cultural complexity: Studies in the social organization of meaning. New York: Columbia University Press.

Harre, R. (1978). Accounts, actions, and meanings—The practice of participatory psychology. In M. Brenner, ed. , P. Marsh,, ed. & M. Brenner (Eds.), The social contexts of method. New York: St. Martin's Press.

Headland, T. N., ed. , Pike, K. L., ed. , & Harris, M. (Eds.). (1990). Emics and etics: The insider/outsider debate. Newbury Park, CA: Sage.

Helmholtz, H. (1971). Selected writings (R. Kahn, Ed.). Middletown, CT: Wesleyan University Press.

Hirschi, T., & Selvin, H. C. (1967). Principles of survey analysis. New York: Free Press.

House, E. Realism in research. Educational Researcher vol. 20 (6) (1991). pp. 2–9.

Hunt, S. D. Positivism and paradigm dominance in consumer research: Toward critical pluralism and rapprochement. Journal of Consumer Research vol. 18 (1991). pp. 32–44.

SAGE researchmethods

Jakobson, R. (1956). Two aspects of language and two types of aphasic disturbance. In R. Jakobson, ed. & M. Halle (Eds.), Fundamentals of language (pp. 55–82). The Hague: Mouton. (Reprinted in R. Jakobson, 1987, Language in literature. Cambridge, MA: Harvard University Press)

Kaplan, A. (1964). The conduct of inquiry. San Francisco: Chandler.

Keesing, R. Anthropology as interpretive quest. Current Anthropology vol. 28 (1987). pp. 161–176.

Kirk, J., & Miller, M. L. (1986). Reliability and validity in qualitative research. Beverly Hills, CA: Sage.

Kitcher, P., ed. , & Salmon, W. C. (Eds.). (1989). Scientific explanation. Minneapolis: University of Minnesota Press.

Kuhn, T. S. (1970). The structure of scientific revolutions (2nd ed.). Chicago: University of Chicago Press.

Kvale, S. (1989). Introduction. In S. Kvale (Ed.), Issues of validity in qualitative research. Lund, Sweden: Studentlitteratur.

Lakoff, G. (1987). Women, fire, and dangerous things: What categories reveal about the mind. Chicago: University of Chicago Press.

Lincoln, Y. S. Campbell's retrospective and a constructivist's perspective. Harvard Educational Review vol. 60 (1990). pp. 501–540.

MacIntyre, A. The idea of a social science. Aristotelian Society Supplement vol. 41 (1967). pp. 93–114.

Manicas, P. T. (1987). A history and philosophy of the social sciences. Oxford, Eng.: Blackwell.

Maxwell, J. A. The evolution of Plains Indian kin terminologies: A non-reflectionist account. Plains Anthropologist vol. 23 (79) (1979). pp. 13–29.

The Qualitative Researcher's Companion:
Understanding and Validity in Qualitative Research
SAGE research methods

Maxwell, J. A. (1986). The conceptualization of kinship in an Inuit community: A cultural account. Unpublished doctoral dissertation, University of Chicago.

Maxwell, J. A. Up from positivism. Harvard Educational Review vol. 60 (1990a). pp. 497–501.

Maxwell, J. A. Response to "Campbell's retrospective and a constructivist's perspective." Harvard Educational Review vol. 60 (1990b). pp. 504–508.

Maxwell, J. A., & Miller, B. A. (n.d.). Two aspects of thought and two components of qualitative data analysis. Unpublished manuscript.

McKinley, R. A critique of the reflectionist theory of kinship terminology: The Crow/ Omaha case. Man vol. 6 (1971). pp. 228–247.

Menzel, H. (1978). Meaning: Who needs it? In M. Brenner, ed. , P. Marsh,, ed. & M. Brenner (Eds.), The social contexts of method (pp. 140–171). New York: St. Martin's Press.

Merriam, S. (1988). Case study research in education: A qualitative approach. San Francisco: Jossey-Bass.

Miles, M. B., & Huberman, A. M. (1984). Qualitative data analysis: A sourcebook of new methods. Beverly Hills, CA: Sage.

Mishler, E. G. (1986). Research interviewing: Context and narrative. Cambridge, MA: Harvard University Press.

Mishler, E. G. Validation in inquiry-guided research: The role of exemplars in narrative studies. Harvard Educational Review vol. 60 (1990). pp. 415–442.

Morgan, L. H. (1871). Systems of consanguinity and affinity of the human family. Smithsonian Contributions to Knowledge. Vol. 17. Washington, DC: Smithsonian Institution.

The Qualitative Researcher's Companion:
Understanding and Validity in Qualitative Research

Norris, S. P. (1983). The inconsistencies at the foundation of construct validation theory. In E. R. House (Ed.), Philosophy of evaluation (pp. 53–74). San Francisco: Jossey-Bass.

Patton, M. Q. (1990). Qualitative evaluation and research methods (2nd ed.). Newbury Park, CA: Sage.

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

Phillips, D. C. (1990). Positivistic science: Myths and realities. In E. G. Guba (Ed.), The paradigm dialog (pp. 31–45). Newbury Park, CA: Sage.

Pitman, M. A., & Maxwell, J. A. (1992). Qualitative approaches to evaluation. In M. D. LeCompte, ed. , W. L. Millroy,, ed. & J. Preissle (Eds.), The handbook of qualitative research in education (pp. 729–770). San Diego: Academic Press.

Platt, J. R. Strong inference. Science vol. 146 (1964). pp. 347–353.

Putnam, H. (1990). Realism with a human face. In H. Putnam, ed. , Realism with a human face (pp. 3–29). Cambridge, MA: Harvard University Press.

Ricoeur, P. (1981). Hermeneutics and the human sciences (J. B. Thompson, Ed. and Trans.). Cambridge, Eng.: Cambridge University Press.

Rorty, R. (1979). Philosophy and the mirror of nature. Princeton, NJ: Princeton University Press.

Runciman, W. G. (1983). A treatise on social theory. Vol. 1. The methodology of social theory. Cambridge, Eng.: Cambridge University Press.

Ryle, G. (1949). The concept of mind. London: Hutchinson.

Salmon, W. C. (1984). Scientific explanation and the causal structure of the world. Princeton, NJ: Princeton University Press.

Salner, M. (1989). Validity in human science research. In S. Kvale (Ed.), Issues of validity in qualitative research (pp. 47–71). Lund, Sweden: Studentlitteratur.

Scriven, M. (1974). Maximizing the power of causal investigations: The modus operandi method. In W. J. Popham (Ed.), Evaluation in education—Current applications (pp. 68–84). Berkeley, CA: McCutchan. (Reprinted in G. V. Glass, Ed., 1976, Evaluation Studies Review Annual, 1, pp. 101–118. Beverly Hills, CA: Sage)

Strauss, A. (1987). Qualitative analysis for social scientists. Cambridge, Eng.: Cambridge University Press.

Trautmann, T. R. (1987). Lewis Henry Morgan and the invention of kinship. Berkeley: University of California Press.

Winch, P. (1958). The idea of a social science and its relation to philosophy. London: Routledge & Kegan Paul.

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

Wolcott, H. F. (1990b). Writing up qualitative research. Newbury Park, CA: Sage.

Yin, R. K. (1984). Case study research: Design and methods. Beverly Hills, CA: Sage.

SAGE research methods