Problems of Reliability and Validity in Ethnographic Research

Margaret D. LeCompte; Judith Preissle Goetz

Review of Educational Research, Vol. 52, No. 1 (Spring, 1982), 31-60.

Stable URL:
http://links.jstor.org/sici?sici=0034-6543%28198221%2952%3A1%3C31%3APORAVI%3E2.0.CO%3B2-5

Review of Educational Research is currently published by American Educational Research Association.

Review of Educational Research
Spring 1982, Vol. 52, No. 1, Pp. 31–60

Problems of Reliability and Validity in Ethnographic Research

Margaret D. LeCompte
Houston Independent School District

and

Judith Preissle Goetz
University of Georgia

Although problems of reliability and validity have been explored thoroughly by experimenters and other quantitative researchers, their treatment by ethnographers has been sporadic and haphazard. This article analyzes these constructs as defined and addressed by ethnographers. Issues of reliability and validity in ethnographic design are compared to their counterparts in experimental design. Threats to the credibility of ethnographic research are summarized and categorized from field study methodology. Strategies intended to enhance credibility are incorporated throughout the investigative process: study design, data collection, data analysis, and presentation of findings. Common approaches to resolving various categories of contamination are illustrated from the current literature in educational ethnography.

The value of scientific research is partially dependent on the ability of individual researchers to demonstrate the credibility of their findings. Regardless of the discipline or the methods used for data collection and analysis, all scientific ways of knowing strive for authentic results. In all fields that engage in scientific inquiry, reliability and validity of findings are important. A common criticism directed at so-called qualitative investigation (e.g., Magoon, 1977; Reichardt & Cook, 1979) is that it fails to adhere to canons of reliability and validity. This discussion applies the tenets of external and internal validity and reliability as they are used in postivistic research traditions to work done by ethnographers and other researchers using qualitative methods. In so doing, these tenets are translated and made relevant for researchers in the qualitative, ethnographic, or phenomenological traditions.

In this paper ethnographic research is used as a shorthand rubric for investigations described variously as qualitative research, case study research, field research, anthropological research, or ethnography (Smith, 1979). Characteristics of ethno-

This article is based on a paper presented at the annual meeting of the American Anthropological Association, Cincinnati, December 1979. Its revision has benefited from the comments of our colleagues: M. Brightman, M. Ginsburg, J. Levin, M. Melville, L. E. Munjanganja, K. Newman, A. D. Pellegrini, J. Pyper, J. Schreiber, R. T. Sieber, and the *Review of Educational Research*'s two anonymous reviewers.

graphic research include participant and nonparticipant observation, focus on natural settings, use of participant constructs to structure the research, and investigator avoidance of purposive manipulation of study variables. Although these approaches are most common in sociology and anthropology, they are used to some extent by all social science disciplines. Wherever they are used, credibility mandates that canons of reliability and validity be addressed, even when ethnographic techniques are adapted within a broader, more positivistic design.

Reliability in ethnographic research is dependent on the resolution of both external and internal design problems (Hansen, 1979). External reliability addresses the issue of whether independent researchers would discover the same phenomena or generate the same constructs in the same or similar settings. Internal reliability refers to the degree to which other researchers, given a set of previously generated constructs, would match them with data in the same way as did the original researcher.

While reliability is concerned with the replicability of scientific findings, validity is concerned with the accuracy of scientific findings. Establishing validity requires determining the extent to which conclusions effectively represent empirical reality and assessing whether constructs devised by researchers represent or measure the categories of human experience that occur (Hansen, 1979; Pelto & Pelto, 1978). Internal validity refers to the extent to which scientific observations and measurements are authentic representations of some reality. External validity addresses the degree to which such representations may be compared legitimately across groups.

Although reliability and validity are problems shared by ethnographers, experimenters, and other researchers, some factors confounding the credibility of findings in experimental designs are inapplicable to ethnographic research; others need to be defined in special ways. In comparing and contrasting threats to validity and reliability recognized by both experimental researchers and ethnographers, we seek to clarify their relevance to other research traditions as well.

The results of ethnographic research often are regarded as unreliable and lacking in validity and generalizability. Some ethnographers ignore such criticisms; others, recognizing potential threats to the credibility of their findings, develop strategies addressing the issues. A few codify their techniques for comprehensibility across research disciplines and traditions (e.g., Cicourel, 1964; Denzin, 1978; Hansen, 1979; Naroll, 1962; Pelto & Pelto, 1978).

Ethnographic research differs from positivistic research, and its contributions to scientific progress lie in such differences. These may involve the data gathering that necessarily precedes hypothesis formulation and revision or may focus on descriptive investigation and analysis. By admitting into the research frame the subjective experiences of both participants and investigator, ethnography may provide a depth of understanding lacking in other approaches to investigation. Ignoring threats to credibility weakens the results of such research, whatever its purpose may be. However, addressing threats to credibility in ethnography requires different techniques from those used in experimental studies. A discussion of reliability and validity problems in ethnographic research properly begins with specification of major differences between the two research traditions.

Differences between Experimentation and Ethnography

Distinctive characteristics of ethnographic research designs (discussed exhaustively elsewhere [e.g., Rist, 1977; Smith, 1979; Wilson, 1977; Wolcott, 1975]) result in variations in the ways problems of reliability and validity are approached in ethnographic and experimental research. Three significant areas are the formulation of research problems, the nature of research goals, and the application of research results.

Formulation of Problems

Formulation of an initial research problem involves both the delineation of a content area and the choice of appropriate design and methods for investigation. Positivistic and ethnographic research differ in approach to these issues.

In research focusing on the examination of effects caused by a specific treatment, credibility of the research design and the power of the treatment effect are established by holding constant or eliminating as many of the extraneous and contextual factors as possible. Ethnography, on the other hand, emphasizes the interplay among variables situated in a natural context. It rarely focuses on treatment unless a treatment or experimental manipulation is part of an overall context. Credibility is established by systematically identifying and examining all causal and consequential factors (Goetz & LeCompte, 1981; LeCompte & Goetz, in press; Scriven, 1974). The process involved differs from the post hoc analysis, which provides contextual information in positivistic traditions. The naturalistic setting in which ethnography normally is conducted both facilitates on-the-spot analysis of causes and processes and precludes precise control of so-called extraneous factors. The interrelationship among such factors generally constitutes the focus of ethnographic concern.

Nature of Goals

A second distinction between the two research traditions lies in the nature of their research goals. This issue relates less to initial formulation of a research question than to the stage of the research at which the use of theory becomes salient, the way theoretical considerations are integrated into the study, and the extent to which the goal of the study is to substantiate existing theory or to generate new theories (Goetz & LeCompte, Note 1).

Ethnographers attempt to describe systematically the characteristics of variables and phenomena, to generate and refine conceptual categories, to discover and validate associations among phenomena, or to compare constructs and postulates generated from phenomena in one setting with comparable phenomena in another setting. Hypotheses, or causal propositions fitting the data and constructs generated, then may be developed and confirmed. Ethnographers commonly avoid assuming a priori constructs or relationships. By contrast, experimental research is oriented to the verification or testing of causal propositions developed externally to the specific research site. Having hypothesized specific causal relationships between variables, experimenters test the strength or power of causes on effects. In a sense, experimental

researchers hope to find data to match a theory; ethnographers hope to find a theory that explains their data.[1]

Application of Results

Most findings from experiments, survey designs, and quasi-experimental studies are intended to be generalized from the subjects sampled to some wider population. Reichardt and Cook (1979) note that such generalization is warranted only where subjects have been sampled randomly from the entire population to which the findings are applied, and they caution that this statistical condition obtains in few cases. Experimenters and survey analysts more commonly depend on design controls, sample size, and assumptions of equivalence to legitimize their generalizations.

Ethnographers rarely have access even to these nonstatistical conditions for generalization. As a consequence, they aim in application for comparability and translatability of findings rather than for outright transference to groups not investigated. Comparability and translatability are factors that could contribute to effective generalization in experimental studies; they are crucial to the application of ethnographic research.

Comparability requires that the ethnographer delineate the characteristics of the group studied or constructs generated so clearly that they can serve as a basis for comparison with other like and unlike groups (Wolcott, 1973). Translatability assumes that research methods, analytic categories, and characteristics of phenomena and groups are identified so explicitly that comparisons can be conducted confidently. Assuring comparability and translatability provides the foundation upon which comparisons are made. For ethnographers, both function as an analog to the goals of more closely controlled research: generalizability of research findings and production of causal statements.

For comparative purposes, ethnographers may choose phenomena to study because they are similar or because they differ systematically along particular dimensions. In either case, the intention is the clarification, refinement, and validation of constructs. This method can be used to compare phenomena identified in a single research site (Glaser & Strauss, 1967; Goetz & LeCompte, 1981), or it can be used by researchers engaged in ethnographic study of special phenomena in a number of research sites (e.g., Cassell, 1978; Herriott, 1977; Herriott & Gross, 1979; Stake, 1978; Tikunoff, Berliner, & Rist, 1975; Wax, in press; Whiting, 1963; Rist, Note 2).

---

[1] A stereotypic distinction labels experimentation as hypothesis verifying and ethnography as hypothesis generating. This simplification has been challenged legitimately by some scholars (e.g., Reichardt & Cook, 1979). Our position is that such dimensions as generation-verification and induction-deduction are continuous rather than discrete processes and that researchers shift along these continua as they proceed through any particular research project and follow some line of investigation. Although ethnographers customarily depend on generative and inductive strategies in the early phases of a research study, they direct later stages of the interactive collection-analysis process to deductive verification of findings. Even where ethnographers begin with an explicit theory to verify (e.g., Erickson, 1943, cited in Campbell, 1979), discrepant data are used first to reject initial explanations and then to generate and verify more adequate explanations. Likewise, experimenters will use unexpected findings as stimuli to generate new theory and will examine its feasibility over a series of studies (Mehan & Griffin, 1980).

Triangulating Research Design

Specifications of differences in overall design between experimental and ethnographic research do not preclude legitimate sharing of data collection strategies (Denzin, 1978). Ethnographic techniques may be supplemental, augmenting reliability or validity of an experimental design. Such strategies enhance the replicability of a treatment by providing a procedural and contextual frame for experimental manipulation.

In contrast, an informal experiment occurs when ethnographers use deliberate manipulations to elicit participant sanctions for the violation of social norms or to provoke other reactions from subjects of a study (e.g., King, 1967; Rosenfeld, 1971). In these cases experimental manipulations are supplemental to ethnography, providing special data for a naturalistic study.

This discussion first addresses problems of reliability and their redress in ethnographic studies. An analysis of problems of validity will follow. In certain respects these issues overlap; what threatens reliability in ethnographic research also may threaten the validity of a study. The two are separated here for heuristic purposes, with indications of overlap where necessary. For both issues, the discussion will refer to the three characteristics of ethnographic design delineated above: contextual focus, eclectic approaches to theory, and comparative applications.

Reliability

Reliability refers to the extent to which studies can be replicated. It requires that a researcher using the same methods can obtain the same results as those of a prior study. This poses a herculean problem for researchers concerned with naturalistic behavior or unique phenomena. Establishing the reliability of ethnographic design is complicated by the nature of the data and the research process, by conventions in the presentation of findings, and by traditional modes of training researchers.

Constraints on Ethnography Reliability

When compared to the stringently controlled designs of laboratory experiments or to the regulated procedures of field experiments, ethnographic design may appear to baffle attempts at replication. The type of data and the research process itself may preclude the use of standardized controls so essential in experimental research. Accommodating the strictures of experimental control requires manipulation of phenomena, which distorts their natural occurrence. Attempts at rigorous measurement may impede construction of powerful analytic categories if the phenomena observed are prematurely or inappropriately reduced or standardized.

Ethnographic research occurs in natural settings and often is undertaken to record processes of change. Because unique situations cannot be reconstructed precisely, even the most exact replication of research methods may fail to produce identical results. For example, Fuch's study (1966) of a racial incident at an urban elementary school cannot be replicated exactly because the event cannot be reproduced. Problems of uniqueness and idiosyncrasy can lead to the claim that no ethnographic study can be replicated. However, generation, refinement, and validation of constructs and postulates may not require replication of situations. Moreover, because human behavior is never static, no study can be replicated exactly, regardless of the methods and designs employed.

Among experimental researchers there is substantial familiarity with the analytic and statistical techniques appropriate to particular kinds of data. These are codified in textbooks and are shared across disciplines. Well-established norms also dictate that research reports and proposals include a description of the population studied as well as methods and instruments used, including established measures of reliability and validity and discussion of analytic techniques.[2]

Reliability in ethnography may be affected by traditions and ideologies in anthropology and field sociology regarding the way a report is presented. A consequence of the debate as to whether anthropology is an art (e.g., Evans-Pritchard, 1962) or a science (e.g., Kaplan & Manners, 1972) is the custom of presenting the results of a study artfully and accessibly. While this style is defended as providing effective communication of cultural knowledge, it could lead neophytes to the unwarranted conclusion that the ethnographic process is facile and simplistic.

The tradition of an artful presentation of results, combined with the strictures imposed by journal-length manuscripts, has resulted in the use of shorthand descriptors for research design and analytic techniques meaningful to research peers but deceptive to the uninitiated. Ethnography uses as its primary data collection technique the writing of field notes, either in situ or as immediately following the event observed as is ethically and logistically possible. However, ethnography is also multimodal; ethnographers use a wide range of techniques to supplement and corroborate their field notes, including the manipulations of the field that would be familiar to an experimental researcher (Wilson, 1977). Describing research merely as ethnography may obscure researcher use of on-site observations, structured and unstructured interviews, projective tests, photographs and videotapes, and survey censuses.

Ethnographers share a common intellectual heritage in which knowledge of all these research techniques is acquired in apprenticeships. This knowledge may be assumed on the part of the reader when results are presented. Ethnographic researchers themselves recognize the necessity for probing beyond journal-length articles to the more complete description of design, data collection, and data analysis located in technical reports and monographs. In some cases, replication may require direct communication with the individual who conducted the original research. Researchers untrained in anthropology or sociology may not exercise such care.

The ethnographic process also is personalistic; no ethnographer works just like another. A researcher's failure to specify precisely what was done may create serious problems of reliability.

Failure among ethnographers to provide sufficient design specificity has led to controversy. Pelto and Pelto (1978) and Kaplan and Manners (1972) identify the

---

[2] Claims for the systematic codification across disciplines of experimental, statistical, and other quantitative research techniques are not intended to imply either single-solution approaches to design problems or agreement among scholars on either significance of problems or effectiveness of solutions (see Cook & Campbell, 1979, for delineation of diverse issues in quantitative design). Our treatment of quantitative methods is simplified for contrastive purposes. We do assert, however, that quantitative strategies have been explicated more widely and systematically than qualitative methodology, a factor contributing to the intensity of debates among experimenters, statisticians, and survey analysts.

highly publicized discrepancy between two ethnographers' studies of the same Mexican village (i.e., Lewis, 1951; Redfield, 1930) as a consequence of the differences in their research designs. Redfield and Lewis addressed different issues, used different methods and time periods, and elicited responses from different segments of the population. Their studies were conducted from different, unexplicated world views and scientific assumptions. The problem was aggravated by presenting their results as representative of the belief system and social structure of the village as a whole rather than as derived from the discrete units actually investigated.

Neither external nor internal reliability, as threats to the credibility of inquiry, are problems unique to ethnographers. However, the discussion below examines these two issues in an ethnographic context and identifies ways that ethnographers address them.

External Reliability

Because of factors such as the uniqueness or complexity of phenomena and the individualistic and personalistic nature of the ethnographic process, ethnographic research may approach rather than attain external reliability (Hansen, 1979; Pelto & Pelto, 1978). Ethnographers enhance the external reliability of their data by recognizing and handling five major problems: researcher status position, informant choices, social situations and conditions, analytic constructs and premises, and methods of data collection and analysis.

Researcher status position. This issue can be phrased, "to what extent are researchers members of the studied groups and what positions do they hold?" In some ways, no ethnographer can replicate the findings of another because the flow of information is dependent on the social role held within the studied group and the knowledge deemed appropriate for incumbents of that role to possess (Wax, 1971). For example, male researchers in tribal societies may find it difficult to obtain information about female rituals and child-rearing practices because these subjects may be unknown to men, known only through an artfully constructed set of myths, or deemed taboo for men even to consider (cf., e.g., Hammond & Jablow, 1976; Paulme, 1963; Reiter, 1975). Similarly, researchers who have friends among student groups and peer cliques (e.g., Cusick, 1973) will obtain different information about student values than those who have little access to students and who must rely on reports from teachers and principals (e.g., Fuchs, 1969).

Ethnographic conclusions are qualified by the investigator's social role within the research site. Other researchers will fail to obtain comparable findings unless they develop corresponding social positions or have research partners who can do so. Although research results generated by ethnographers whose positions were limited in scope may be only narrowly applicable, they are nonetheless legitimate. Such conclusions delineate facets of reality within a group, other aspects of which may be identified by researchers taking other social positions. Glaser and Strauss (1967) refer to these individual facets as slices of data which, taken together, contribute to the total picture of group life. McPherson's analysis (1972) of schooling in a small U.S. town is based on her observations as an elementary schoolteacher. Her description of schoolchildren may represent the relatively narrow perspective of teacher, but can be replicated only by researchers who assume comparable roles. Studies of students in other small U.S. towns, conducted from alternative role positions, must be regarded as supplemental studies rather than replicative studies.

Because ethnographic data depends on the social relationship of researcher with subjects, research reports must clearly identify the researcher's role and status within the group investigated (e.g., Sieber, in press). In addition, some researchers enter settings as nonparticipant observers who develop no personal relationships with members of the groups, while others develop friendships that provide access to some kinds of special knowledge while limiting access to others. Ethnographers customarily label their investigative stance toward participants according to taxonomies such as that developed by Gold (1958) and describe the content and development of the social status and position accorded them by the group participants (e.g., Janes, 1961; Wax, 1971).

Informant choices. Closely related to the role the researcher plays is the problem of identifying the informants who provide data. Different informants represent different groups of constituents; they provide researchers with access to some people, but preclude access to others. For example, in Cusick's ethnographic study (1973) of student culture in a midwestern high school, his initial association with a clique of senior athletes facilitated his entry to groups with whom the athletes associated, but hindered his access to other cliques and to student isolates. In associating with one group, researchers may forfeit information about the life experiences of people in other groups. Berreman's retrospective analysis (1962) of fieldwork in India provides a classic example of the extent to which knowledge gathered is a function of who gives it.

Participants who gravitate toward ethnographers and other field researchers may be atypical of the group under investigation; similarly, those sought by ethnographers as informants and confidants also may be atypical (Dean, Eichhorn, & Dean, 1967). Sometimes this is necessary because people who speak languages comprehensible to researchers, who understand the analytic categories used by ethnographers, and who are introspective and insightful about their own lives are rare in most groups. The qualities that make them valuable as informants and research assistants may mark them as deviant from their own groups.

Threats to reliability posed by informant bias are handled most commonly by careful description of those who provided the data. Such characterization includes personal dimensions relevant to the researcher as well as dimensions significant to the informant and others in the group. External reliability requires both careful delineation of the types of people who served as informants and the decision process invoked in their choice.

Social situations and conditions. A third element influencing the content of ethnographic data is the social context in which they are gathered. What informants feel to be appropriate to reveal in some contexts and circumstances may be inappropriate under other conditions. In Ogbu's study (1974) of education in an ethnic neighborhood of a big city, he distinguishes carefully the information parents reveal when they enter the school context from what they reveal in their home neighborhood. He quotes extensively from his field notes to demonstrate that this discrepancy is recognized and discussed among the parents themselves. Ogbu's experiences highlight the necessity for ethnographers to specify the social settings where data are collected.

Other social circumstances also affect the nature of information revealed. In their analysis of medical school student culture, Becker, Geer, Hughes, and Strauss (1961) differentiate between data gathered from participants alone with the researchers and

information acquired from participants in group contexts. Their study indicates that what people say and do varies according to others present at the time.

Delineation of the physical, social, and interpersonal contexts within which data are gathered enhances the replicability of ethnographic studies. To an extent, these factors are subject to change over time. What may be a center for informal gathering among one group of high school seniors, for example, may be anathema to the succeeding class. Consequently, descriptions of contexts should include function and structure as well as specification of features.

Analytic constructs and premises. Even if a researcher reconstructs the relationships and duplicates the informants and social contexts of a prior study, replication may remain impossible if the constructs, definitions, or units of analysis which informed the original research are idiosyncratic or poorly delineated. Replication requires explicit identification of the assumptions and metatheories that underlie choice of terminology and methods of analysis. For example, the culture concept is defined differently by different researchers. Some use it globally: Linton (1945) identified it as the way of life of a people. Others prefer to define culture more narrowly in terms of observed behavior (e.g., Harris, 1971). Some virtually deny that culture exists independently as an analytic construct, preferring to examine the minute-by-minute interactions by which shared meanings are negotiated among individuals and small groups (e.g., Furlong, 1976; Gearing, 1973, 1975).

If defined idiosyncratically in a study, major organizing constructs such as these can lead to findings that differ widely in their emphasis and interpretation. When underlying assumptions and definitions remain unclarified, the results may be incomprehensible. Researchers may develop their own conceptual schemes in ignorance or disregard of constructs used by other researchers and may fail to provide an analysis of or theory about their implicit structures (Biddle, 1967).

Smith and Brock (1970), for example, note that the work of certain ecological psychologists (i.e., Barker & Wright, 1954) implies the obviation of behavior that appears to have no purpose. In positing both the logical supremacy of the largest unit, the behavioral episode, and a world governed by linear causality, Barker & Wright base their analysis on a simple stimulus-response model of behavior; however, this theoretical underpinning is not made explicit. It may be useful for post hoc analysis of behavior transcripts, but the proposition that behavioral episodes (or any other units of analysis) are natural or intrinsic to the human condition is unverified. Smith and Brock legitimately observe that behavioral episodes may be congruent with common sense, but with common sense as viewed by a given researcher using a specific paradigm. To the extent that invented constructs such as these are mandated by the data, their assumptions, definitions, and limitations should be delineated explicitly, and their relationships to existing concepts should be clarified.

Outlining the theoretical premises and defining constructs that inform and shape the research facilitates replication. However, development of lower level constructs and terms creates problems for internal as well as external reliability. Creating categories for coding is the first step of analysis; it is vital to the process of organizing the naturally occurring stream of behavior into manageable units. Units of analysis should be identified clearly: where they begin and end and, when appropriate, which variables form the framework for data collection and analysis (Goetz & LeCompte, 1981).

Some ethnographers specify clearly their categories of data. They may use standard typologies and checklists (e.g., Henry, 1960; Hilgar, 1966; Whiting, Child, & Lambert, 1966). More problematic are situations in which researchers devise their own schemes. This process may be necessary to provide a valid analytic frame that matches the data collected and the questions posed. However, unless categories are defined carefully and their theoretical antecedents outlined, the dangers of idiosyncrasy and lack of comparability are magnified. Establishing interobserver reliability may be impossible. On the other hand, established classificatory schemes may be used merely because they are well known and easy to administer, even though they may result in premature categorization that misrepresents the data or inadequate standardization and mechanical reduction that trivializes ethnographic findings.

Methods of data collection and analysis. Ideally, ethnographers strive to present their methods so clearly that other researchers can use the original report as an operating manual by which to replicate the study (e.g., Becker, Geer, & Hughes, 1968; Mehan, 1979; Ogbu, 1974; Smith & Geoffrey, 1968; Wolcott, 1973). Failures to specify methods of data collection and analysis may be related to the aforementioned brevity that journals often require in manuscripts. Pelto & Pelto (1978) note the regularity with which journal authors fail to report sufficiently their research designs and methodology. To an extent, this is because of the difficulty of explaining in a few sentences the scope and development of ethnographic research techniques.

Replicability is impossible without precise identification and thorough description of the strategies used to collect data (for compendiums of the range of alternatives, see LeCompte & Goetz, in press; Pelto & Pelto, 1978; Schatzman & Strauss, 1973; Spradley, 1979, 1980; Williams, 1967). Although this admonition may appear elementary to experimental researchers, knowledge of ethnographic technique is apprehended incompletely and shared unevenly across the disciplines now using them (Burns, 1976; Herriott, 1977; Ianni, 1976; Wolcott, 1971). Until commonly understood descriptors for these complex techniques are developed, shorthand designations will continue to obstruct reliability, and researchers seeking to replicate studies will depend on fugitive monographs, technical reports, and personal communications.

A more serious problem for both external and internal reliability is the identification of general strategies for analyzing ethnographic data. The analytic processes from which ethnographies are constructed often are vague, intuitive, and personalistic. Ethnographers disagree on the extent to which such processes can and should be articulated (cf., e.g., Erickson, 1973; Pelto & Pelto, 1978; Wolcott, 1975; Wolcott, Note 3). Recent efforts to codify techniques for data analysis include Pelto and Pelto's system (1978) of deductive, inductive, and abductive strategies; Smith (1974, 1979) and Smith and Brock's (1970) efforts to generate models of the analytic process; and Goetz and LeCompte's comparative examination (1981) of analytic induction (Mehan, 1979; Robinson, 1951; Znaniecki, 1934), constant comparison (Glaser & Strauss, 1967), typological analyses (e.g., Lofland, 1971), enumerative systems (e.g., McCall, 1969), and standardized protocols (e.g., Flanders, 1970). Because reliability depends on the potential for subsequent researchers to reconstruct original analytic strategies, only those ethnographic accounts that specify these in sufficient detail will be replicable.

Internal Reliability

Problems of internal reliability in ethnographic studies raise the question of whether, within a single study, multiple observers will agree. This issue is especially critical when a researcher or research team plans to use ethnographic techniques to study a problem at several research sites (e.g., Cassell, 1978; Herriott, 1977; Herriott & Gross, 1979; Stake, 1978; Tikunoff, Berliner, & Rist, 1975; Whiting, 1963; Rist, Note 2). Crucial to internal reliability is interrater or interobserver reliability, the extent to which the sets of meanings held by multiple observers are sufficiently congruent so that they describe phenomena in the same way and arrive at the same conclusions about them.

Because ethnographers rarely use the standardized protocols for which some types of interrater reliability are crucial, the more pertinent concern is whether multiple observers agree with each other and with the originator of general constructs on their classifications or on a typology with which to begin categorization. Thus, the agreement ethnographers seek is more appropriately designated interobserver reliability. Agreement is sought on the description or composition of events rather than on the frequency of events.

This is a key concern to most ethnographers. Of necessity, a given research site may admit one or few observers. In the absence of other means of corroboration, such investigations may be idiosyncratic, rather than careful and systematic recordings of phenomena. Ethnographers commonly use any of five strategies to reduce threats to internal reliability: low-inference descriptors, multiple researchers, participant researchers, peer examination, and mechanically recorded data.

Low-inference descriptors. The format, structure, and focus of ethnographic field notes vary with the research problem and design and with the skills and styles of individual ethnographers. However, most guides to the construction of field notes distinguish between two categories of notations. Low-inference descriptors, phrased in terms as concrete and precise as possible, are mandated for all ethnographic research. These include verbatim accounts of what people say as well as narratives of behavior and activity (Lofland, 1971; Pelto & Pelto, 1978; Schatzman & Strauss, 1973). The second category of notation may be any combination of high-inference interpretive comments and will vary according to the analytic scheme chosen.

Low-inference narratives provide ethnographers with their basic observational data. Interpretive comments can be added, deleted, or modified, but the record of who did what under which circumstances should be as accurate as possible (Wax, 1971). This material is analyzed and presented in excerpts to substantiate inferred categories of analysis (Wolcott, 1975). Those ethnographies rich in primary data, which provide the reader with multiple examples from the field notes, generally are considered to be most credible (e.g., Bossert, 1979; Leemon, 1972; Modiano, 1973; Smith & Keith, 1971; Ward, 1971; Wolcott, 1977).

Multiple researchers. The optimum guard against threats to internal reliability in ethnographic studies may be the presence of multiple researchers. In some cases, investigations take place within a team whose members discuss the meaning of what has been observed until agreement is achieved (e.g., Becker et al., 1961, 1968; Peshkin, 1978; Spindler, 1973). Tikunoff, Berliner, and Rist (1975) conducted an intensive, 3-week training period for their 12 observers to prepare them to obtain comparable

descriptive protocols from the 40 elementary classrooms examined in a study of effective reading and mathematics instruction.

Ethnographies based on team observation constitute the minority, and most involve only two researchers (e.g., Cicourel & Kitsuse, 1963; Hostetler & Huntington, 1971; Whiting, 1963). The same constraints of time and money that preclude the use of research teams limit the size and scope of teams: ethnographic research often is too time consuming and labor intensive for participation of most lone researchers, let alone multiple investigator teams. Funding is rarely available for more than a single fieldworker. In this case, ethnographers depend on other sources for corroboration and confirmation. Some of the recent, federally funded multiple-site research programs have employed research teams (e.g., Cassell, 1978; Wax, in press); others have used confirmation by short-term observers (e.g., Stake, 1978); more commonly, each field observer is responsible for an independent site (e.g., Herriott, 1977; Herriott & Gross, 1979). Especially under the latter circumstances, problems of establishing internal reliability are much the same as for single-site studies.

Participant researchers. Many researchers enlist the aid of local informants to confirm that what the observer has seen and recorded is being viewed identically and consistently by both subjects and researcher (Magoon, 1977). In some cases, participants serve as arbiters (e.g., Smith & Geoffrey, 1968), reviewing the day's production of field notes to correct researcher misperceptions and misinterpretations. Other researchers (e.g., Carroll, 1977) operate in partnership with participants, keeping dual accounts of their own observations alongside participant comments. More commonly, ethnographers request reactions to working analyses or processed material from selected informants (e.g., Wolcott, 1973). In this way confirmation may be sought for various levels of the collection and analysis process: description of events and interactions, interpretation of participant meanings, and explanations for overall structures and processes.

Peer examination. Corroboration of findings by researchers operating in similar settings proceeds in three ways. First, ethnographers may integrate descriptions and conclusions from other fieldworkers in their presentations (e.g., Borman, 1978; Clement & Harding, 1978; Sieber, 1979). If discrepancies occur, explanations are proffered (Kaplan & Manners, 1972). Second, findings from studies conducted concurrently at multiple sites, such as those discussed above, may be analyzed and integrated. Independent generation or confirmation of results support the reliability of observation and enhance cross-site validity of conclusions (Campbell, 1979). Finally, the publication of results constitutes an offering of material for peer review. Wolcott's admonition (1975) to fieldworkers to include sufficient primary data in published accounts recognizes the significance of review by colleagues in the evaluation of ethnographic reports. Magoon (1977) cites Scriven's position (1972) that the reliability of various categories of so-called subjective material rests, to an extent, on the observer's established reputation for truthfulness and accuracy. The issue is not, then, to expurgate the subjective experience of the researcher, but to draw on it for insight as well as to provide information regarding its predictions, biases, and possible influences. In this way, ethnographers study themselves within the setting and their influence on it, as well as the setting itself (Wax, 1971).

Mechanically recorded data. Ethnographers use a variety of mechanical devices to record and preserve data. Mehan (1979) argues for the use of observational techniques

that record as much as possible and preserve to the greatest extent the raw data, so that the veracity of conclusions may be confirmed by other researchers. Video and audio tape recorders, cameras, and moving-picture cameras are becoming standard equipment in the collection of ethnographic data (e.g., Collier, 1973; Eddy, 1969; Mehan, 1979). Such devices do possess serious limitations. Although cameras and recorders register much that a researcher could forget or ignore, and consequently may increase the reliability of a study, they preserve all data in uncodified and unclassified form and record only that data chosen by the researcher to be preserved. They are an abstraction and yet they may preserve too much data. Thus coding and analysis are imperative to render them usable.

Validity

Validity necessitates demonstration that the propositions generated, refined, or tested match the causal conditions which obtain in human life. There are two questions involved in matching scientific explanations of the world with actual conditions in it.

First, do scientific researchers actually observe or measure what they think they are observing or measuring? This is the problem of internal validity; solving it credibly is considered to be a fundamental requirement for any research design (e.g., Campbell & Stanley, 1963; Cook & Campbell, 1979).

Second, to what extent are the abstract constructs and postulates generated, refined, or tested by scientific researchers applicable across groups? This addresses the issue of external validity; it poses special problems for ethnographers because of the nature of their research designs and methods. Contrasting approaches to these problems are discussed below.

Although the problems of reliability threaten the credibility of much ethnographic work, validity may be its major strength. This becomes evident when ethnography is compared to survey studies, experimentation, and other quantitative research designs for assessment of internal validity (Crain, 1977; Erickson, 1977; Reichardt & Cook, 1979). The claim of ethnography to high internal validity derives from the data collection and analysis techniques used by ethnographers (see Denzin, 1978, for comparison of research designs). First, the ethnographer's common practice of living among participants and collecting data for long periods provides opportunities for continual data analysis and comparison to refine constructs and to ensure the match between scientific categories and participant reality. Second, informant interviewing, a major ethnographic data source, necessarily is phrased more closely to the empirical categories of participants and is formed less abstractly than instruments used in other research designs. Third, participant observation, the ethnographer's second key source of data, is conducted in natural settings that reflect the reality of the life experiences of participants more accurately than do contrived settings. Finally, ethnographic analysis incorporates a process of researcher self-monitoring, termed disciplined subjectivity (Erickson, 1973), that exposes all phases of the research activity to continual questioning and reevaluation.

Although internal and external validity are interrelated issues, they customarily are separated (e.g., Campbell & Stanley, 1963; Cook & Campbell, 1979) to clarify procedures, and this convention is discussed below. Among the measures of scientific credibility—internal and external reliability and internal and external validity—the

problems of external validity most frequently are ignored by ethnographers. Reasons for this derive from three common characteristics of the ethnographic process.

First, ethnography focuses on recording in detail aspects of a single phenomenon, whether that phenomenon is a small group of humans or the operation of some social process. Traditionally, ethnographers have concentrated on single research settings. However, studies of a phenomenon, particularly an organizational innovation, over a number of sites have become more common (e.g., Cassell, 1978; Herriott, 1977; Herriott & Gross, 1979; Wax, in press; Rist, Note 2). The task is to reconstruct, in what Lofland (1971) calls loving detail, the characteristics of that phenomenon. Consequently, the ethnographic researcher begins by examining even commonplace groups or processes in a fresh and different way, as if they were exceptional and unique (Erickson, 1973).

In doing this, a second characteristic of ethnographic inquiry emerges. One school of ethnography advocates that researchers enter their fields with an assumption of ignorance or naiveté about the phenomena under investigation; other researchers simply attempt to suspend preconceived notions and even existing knowledge of the field under study. Although they may be familiar with related empirical research and use general theoretical frameworks to initiate studies, fieldworkers assume that detailed description can be constructed more accurately by not taking for granted facets of the social scene (Erickson, 1973).

Third, the problems, goals, and applications of ethnographic research affect how issues of external validity are defined and resolved. As indicated previously the credibility of research, which is contextual, theoretically eclectic, and comparative, is threatened by and grounded in factors different from those pertaining to experimentation and other forms of quantitative research.

Issues pertaining to the validity of ethnographic research, both internal and external, are addressed by fieldworkers operating from the perspective of these characteristics. The following discussion presents the threats to credibility of ethnographic design and their remedies.

Internal Validity

The definition of internal validity presented earlier subsumes the problem of whether conceptual categories understood to have mutual meanings between the participants and the observer actually are shared. For internal validity, the threats that Campbell and Stanley (1963) and Cook and Campbell (1979) describe as posing difficulties for experimental research are equally applicable to ethnographic research, although they present somewhat different problems and may be resolved differently. These threats include history and maturation, observer effects, selection and regression, mortality, and spurious conclusions.

History and maturation. The extent to which phenomena observed at entry or at other initial occasions are the same as those observed subsequently becomes salient when process and change are the focus of the research project. Unlike the experimenter who uses various strategies to hold constant the effects of time, the ethnographer conducts research in natural settings where the clock advances. Changes that occur in the overall social scene are what experimenters designate as history; changes that involve progressive development in individuals are considered to be maturation.

Ethnographers assume that history affects the nature of the data collected and that

phenomena rarely remain constant. The ethnographic task is to establish which baseline data remain stable over time and which data change (LeCompte & Goetz, in press). Such change may be recurrent, progressive, cyclic, or aberrant; sources of change and their operation also need to be specified (Appelbaum, 1970; Lofland, 1971). This is facilitated by systematic replication and comparison of baseline data, analogous to the pretest data collected by experimenters.

In order to assess the rate and direction of change, ethnographers establish long-term residence in their fields—extending from 6 months to 3 years. This permits time-sampling procedures, the identification of factors intervening in the social scene across some period of time, and the retrospective tracing of phenomena isolated in the terminal phases of a study. In situations where data are required from the preentry period of a field study, ethnographers use informant reconstructions and information located in a variety of documents. They may revisit sites at subsequent intervals in order to verify the time-dependent nature of various phenomena.

The classic instance in educational ethnography of site revisiting is Hollingshead's return (1975) to his Elmtown site and the accompanying analysis of changes that occurred over a 30-year period (cf. Mead, 1956; Wylie, 1974). Wolcott, in his examination of education in a Kwakiutl Indian village in Canada (1967), supplemented his 12-month participant observation with extended visits the following two summers and by retrospective interviews with village informants and educators who had taught in the village school prior to his tenure (cf. Hostetler & Huntington, 1971; King, 1967; Modiano, 1973). Ogbu's study (1974) of the inner-city neighborhood traces the 10-year history and development of the education rehabilitation movement in the community's schools through interviews and the collection and analysis of pertinent documents. These researchers used replication and time-sampling strategies to distinguish phenomena subject to change from phenomena that remained relatively stable.

Many of the techniques used by ethnographers to control for the effects of history are applicable to controlling for maturation. Experimenters manage these variables through such constraints as designing projects of limited duration and assigning subjects randomly to control and experimental groups. When effects of treatments are being measured, maturation may be regarded as a source of contamination. For an experimental study, a biological or quasi-biological model with universal stages of development is posited. Maturation is conceptualized as a universal, normative process, proceeding through well-defined stages. Ethnographers, however, view maturational stages as varying according to cultural norms. Fieldworkers attempt to control for the effects of maturation by identifying explicitly what behaviors and norms are expected in different sociocultural contexts. They are less concerned with what people actually are capable of doing at some developmental stage than with how groups specify appropriate behavior for various developmental stages.

Maturation and development frequently become the focus of ethnographic studies (e.g., Howard, 1970; Moore, 1973). Leemon (1972) and Burnett (1969) used Van Gennep's model (1960) of passage rites to analyze maturation of students in the United States. Other researchers (e.g., Becker et al., 1961) have reconstructed the maturation process through the perceptions of the participants involved. Constant comparison (Glaser & Strauss, 1967), discrepant case analysis (Erickson, 1973; Robinson, 1951; Wolcott, 1975; Znaniecki, 1934), and a variety of logico-deductive strategies (e.g., Scriven's modus operandi, 1974) can be used to distinguish maturation

effects from other intervening phenomena in order to identify possible causes, their interactions, and their probable impacts (e.g., Eddy, 1969; Ward, 1971).

Observer effects. The threat to validity posed by observer effects in ethnography is parallel to the threats to experimental and survey studies posed by testing and instrumentation effects. The reactivity of instrumentation (discussed elsewhere, e.g., Campbell & Stanley, 1963; Cook & Campbell, 1979; Phillips, 1971) is as problematic for ethnographers as it is for other social researchers. Participant observation and informant interviewing pose particular problems of their own. The difficulty is amplified by the common practice in ethnography of supplementing these strategies with a variety of standardized instruments.

When data are being gathered through participant observation and informal informant interviewing, reactivity must be assessed. Possible and probable effects of the observer's presence on the nature of the data gathered must be considered (Schwartz & Schwartz, 1955). Such effects operate in a number of ways.

As noted earlier, what observers see and report is a function of the position they occupy within participant groups, the status accorded them, and the role behavior expected of them. Direct observer effects may occur when informants become dependent on the ethnographer for status enhancement or the satisfaction of psychological needs. In such cases, a symbiotic relationship may develop between researcher and informant that precludes obtaining data from other than a single source or that distorts data obtained from other informants who are affected by what they perceive as a special relationship. Ethnographers address this threat by establishing several field relationships (Kahn & Mann, 1952; Miller, 1952; Vidich, 1955), by gradually disengaging themselves from informant relationships (Powdermaker, 1966), and by including in their presentation of results a retrospective analysis of their field positions and relationships (see *Researcher status position* above).

Attempting to avoid problems of entanglement by assuming a position of neutrality can lead the researcher into other distortions. Detachment can destroy rapport and cause informants to infer indifference or even hostility on the part of the researcher. Consequent paranoiac reactions may seriously affect the quality of data (Miller, 1952; Vidich, 1955; Wax, 1971). In settings such as schools, participants may expect, even demand, advocacy from the ethnographer as a condition of rapport (e.g., Cusick, 1973; Goetz, 1976).

Participants may behave abnormally (Argyris, 1952). This may be a consciously planned show in which subjects seek to reveal themselves in the best possible light. Or it may be an unconscious distortion performed to provide what participants believe the researcher wants to see. Interactive situations, in which participants respond spontaneously to the researcher's presence and attention, may result in phenomena comparable to the halo effect documented in experiments and in quasi-experimental field studies (Cook & Campbell, 1979).

Parallel to this problem in observation is the credibility of informant reports in interviewing. Informants may lie, omit relevant data, or misrepresent their claims (Dean & Whyte, 1958). Independent corroboration from multiple informants (e.g., Fuchs, 1969) or other fortuitous observers of the social scene (e.g., Smith & Keith, 1971), sufficient residence in the field to reduce artificial responses (e.g., Wolcott, 1973), and coding participant responses according to situations expected to elicit contrived responses (e.g., Becker et al., 1961; McCall, 1969) are techniques used by ethnographers to control for such distortions in the data.

Unusual observer effects (discussed above as informal social experiments) also may threaten the validity of ethnographic studies. Contrivance effects may distort data gathered: this obtains in situations where the ethnographer plans and executes some exceptional act in order to elicit responses from subjects. Such strategies may violate the research ethics of participant consent (cf., e.g., Denzin, 1978; Jorgensen, 1971; Rynkiewich & Spradley, 1976), although inadvertent faux pas and gaffes are less controversial than deliberate manipulations and do provide valuable information on norms and sanctions. Here the researcher must establish that it is the act itself that elicits the responses rather than the act as performed by the researcher (Webb, Campbell, Schwartz, & Sechrest, 1966).

Two problems are associated with intensive, long-term studies. Research exhaustion, or the saturation of a setting for research purposes (Wolcott, 1975), occurs when the investigation ceases to reveal further new constructs. The ethnographer has become so familiar with the setting that new or discrepant data are no longer observable. Related to this may be the classic problem of going native: ethnographers participate to such a degree in groups that they can no longer maintain sufficient distance from the group role to observe and analyze objectively. Some observers (e.g., Everhart, 1977) interpret these difficulties as an indication that field residence should be terminated; other ethnographers (e.g., Whyte, 1955) advocate periodic temporary withdrawals from the field in order to defamiliarize themselves with the social scene, to reconfirm their primary status as dispassionate researchers, and to provide a respite for participants.

Finally, in cases where presentation of the perspective of participants is important, ethnographers must demonstrate that the categories are meaningful to the participants, reflect the way participants experience reality, and actually are supported by the data. Even where participant-derived constructs are less important, researcher-designated constructs still should be grounded in and congruent with actual data.

In essence, researchers must guard against their own ethnocentrisms and perceptual biases. Disciplined subjectivity (Erickson, 1973) uses the tension arising from the investigator's emotive and affective responses to participant behavior and practice (Wax, 1971) as an indicator of salient phenomena. Through what Wax defines as resocialization, the ethnographer searches for the group's perspectives toward and meanings for significant phenomena (Schatzman & Strauss, 1973), emerging with a dual identity as an outsider-insider which permits authentic presentation of the participant world.

Especially where formal instrumentation is used, ethnographers try to establish the extent to which the measure has the same meaning for both researcher and subject (e.g., Goodman, 1957; Spindler, 1973, 1974). Assumptions underlying instrument items, how they are assessed, and the choice of who scores them, as well as overt meanings of the items and the overall test, should be shared between tester and testee (cf., e.g., Mehan, 1976; Phillips, 1971). Demonstrating equivalence of meaning between researcher and subject is difficult (e.g., Gay & Cole, 1967; Modiano, 1973) and this problem is highlighted in interdisciplinary research where the task is complicated by the necessity for equivalence across different disciplines (Petrie, 1976).

Although sociocultural theories and analytic models provide ethnographers with perspectives for monitoring themselves as members of both participant groups and the scientific community (Schatzman & Strauss, 1973), biases resulting from academic

training also may distort data. For example, disciplinary biases may appear, however implicitly, in the categories an investigator chooses as salient for analysis and coding of ethnographic data, regardless of whether participant-derived categories or researcher-designated constructs are used. Researchers with different theoretical backgrounds may choose to focus on quite different aspects of the data. The strategies discussed above for enhancing the reliability of analytic constructs and premises and for ensuring the internal reliability of ethnographic studies also contribute to controlling and managing observer analytic biases. Of these, participant reaction and confirmation—conducted through all levels of the ethnographic process—may be most effective in revealing researcher-induced distortions (Wax, 1971).

Selection and regression. In experimental research, control of selection and regression effects attempts to ensure that measured differences between treatment and control groups are caused by the treatment rather than by differences inherent in the groups. Although ethnographers rarely grapple with the problem of isolating treatment effects, they do cope with distortions in their data and conclusions created by the selection of participants to observe and informants to interview. Wax (1971) emphasizes that the disciplined investigator seeks and maintains contact with a diversity of participants—despite personal preferences and prejudices—as a strategy for correcting bias and distortion. Selectivity becomes a serious problem in situations where the number of participants necessitates gathering data from some sample of the population or where the social scene is sufficiently complex that continual observation of all events, activities, and settings is precluded. Failure to complete an adequate inventory (cf., e.g., LeCompte & Goetz, in press; Schatzman & Strauss, 1973) of subgroups, factions, events, and social scenes in the field site may result in findings representative only of certain participants or of particular circumstances.

Ethnographers commonly initiate investigations by establishing the range of possible informants and participants in a group so as to obtain data from all participant types (e.g., Dean, Eichhorn, & Dean, 1967). In his study of attitudes toward formal education held by American Indian parents and students, Riner (1979) first identified the categories of families sending their children to school and then sampled from that typology. Conclusions reported by Clement and Harding (1978) in their analysis of student relationships in a desegregated elementary school and by Becker et al. (1961) in their study of student culture in a medical school are based on observations sampled from the range of events, activities, and settings identified in the field sites.

Although marginal individuals and other extreme types among a population may serve as liaisons for entry and initial investigation (Kahn & Mann, 1952; Vidich, 1955), ethnographers try to maintain contacts and relationships with as diverse a group of participants as possible. In her analysis of the impact of a state-mandated curriculum on the staff of a school district, Brown (Note 4) verified perceptions of the innovation reported initially by a few teacher informants through subsequent questionnaires administered to all involved teachers.

The tendency for the exotic to be more obvious than the commonplace affects the events and activities the ethnographer notes as well as the selection of informants. Khleif (1974) and Erickson (1973) suggest that this may be remedied by using strategies such as the aforementioned discrepant case analysis, by constantly questioning commonly assumed meanings, and by making comparisons with cross-cultural data and cases.

Because most ethnographers study characteristics and behavior of human groups rather than the effects of specific treatments, ethnographic subjects are chosen for relevance to specific interests. Glaser and Strauss's (1967) use of theoretical sampling—collecting data chosen for relevance to emerging theoretical constructs—is one purposive strategy for implementing this process systematically (for alternative forms of purposive sampling, see Patton, 1980). Following successful access to and entry into particular groups, methodical sampling assures that data adequately represent the population being investigated. Such sampling may take the form of cross-informant interviewing for confirmation and validation of interviews, structured questionnaires, or findings derived from participant observation across the spectrum of subgroups and factions. These strategies are as useful for ensuring external validity as they are for internal validity: if cross-group comparisons are to be credible, they must be grounded in accurate data from individual groups.

Mortality. The ways in which groups change over time as a result of losses and gains in membership pose special difficulties for ethnographers. Although experimenters may replace subjects who are lost from their studies, ethnographers assume that the naturalistic approach precludes the interchangeability of human informants and participants. Loss and replacement as they naturally occur become topics of study in themselves. Growth and attrition are assumed to be normal processes in most group settings, so the ethnographic task becomes the identification of their effects. This requires careful attention to baseline data (discussed above) so that the researcher may compare events and activities occurring across time.

In his study of enculturation, Jocano (1969) examined mobility of young people into and out of a Philippine barrio, as well as the treatment of birth and death, in order to establish cycles of growth and attrition as defined and interpreted by the community. By studying the spring enrollment of a new child into a third-grade classroom, Goetz (1976) was able to validate socialization practices and goals observed among students earlier in the school year. Smith and Keith (1971) approached staff attrition and turnover in a similar manner to illuminate the social dynamics of innovation in an elementary school program. In each of these instances, collection of baseline data enabled researchers to analyze the effects of subsequent loss and replacement.

Spurious conclusions. However thoroughly an ethnographer may have accounted for effects of history and maturation, observer impact, selection and regression, and mortality, relationships posited among observed phenomena nevertheless may be spurious. This problem is comparable to Cook and Campbell's formulation (1979) of statistical conclusion validity. They define this construct as (a) the extent to which a treatment actually caused a predicted effect and (b) the extent to which presumed phenomena actually covary or are causally related. Statistical conclusion validity alerts researchers to search for spurious relationships and to resist assuming relationships where there may be none or assuming nonrelationships where they may be obscured by an artifact of instrumentation or treatment. These issues are paramount to experimental researchers whose designs customarily preclude laborious post hoc examinations of sources of error except where intuition or insight suggest such errors might exist.

In contrast, ethnographic design mandates what Scriven (1974) has designated a modus operandi perspective in which the geneses of observed data are traced

retrospectively. All plausible causes are delineated by examination of collected data and through discussion with informants. Postulating associations among phenomena depends on elimination of alternative explanations (Campbell, 1979). Denzin (1978) conceptualizes the adequate support of relational generalizations as requiring establishment of time order, covariance, and elimination of rival hypotheses. He assesses participant observation as excellent, good, and fair, respectively, on these three factors.

Elimination of rival explanations mandates control of factors threatening internal validity. It also requires effective and efficient retrieval systems for ethnographic data and the scrupulous use of corroboratory and alternative sources of data. These serve to support the fieldworker's search for negative instances of tentatively postulated relationships and disconfirming evidence for emergent constructs (Mehan, 1979; Robinson, 1951; Znaniecki, 1934). Although no research design can identify the precise cause of an observed datum, ethnographic data may be quite effective in delineating the most probable causes and in specifying an array of those most plausible.

Participant explanations of events are central among those that ethnographers examine. Factors that many researchers designate as causal may not be so designated by participants. Although Rist (1970) rejected teachers' explanations for student failure in his 3-year study of a group of elementary schoolchildren, he demonstrated that those students who failed were those that the teachers expected to fail. In contrast, Smith and Keith (1971) expanded upon participant accounts and interpretations to explain the failure of innovation in an elementary school.

Longevity in the research site, presupposed in ethnographic research design, facilitates the search for causes and consequences. Ethnographers are likely to have witnessed personally the temporal antecedents of events; where this is impossible, data from informants, documents, and other sources may be substituted. Similarly, long-term field residence permits identification of the covariance of phenomena in natural settings. Nevertheless, Cook and Campbell's counsel (1979) to experimenters is applicable as well to ethnographers:

> Estimating the internal validity of a relationship is a deductive process in which the investigator has to systematically think through how each [factor] ...may have influenced the data .... In all of this process, the researcher has to be his or her own best critic, trenchantly examining all of the threats he or she can imagine. (p. 55)

For the ethnographer, the process is an inductive one as well; sources of bias or contamination must be discovered as the study proceeds.

External Validity

In most ethnographic studies, as well as in many quantitative studies, the strictures required for statistical generalization may be difficult to apply. Problems of access may preclude the use of random samples, or random assignments may have to be made from available groups rather than from entire populations. Statistical sampling may even be irrelevant where initial description of a little known or singular phenomenon is desired, where social constructs (to be tested later in more stringently controlled designs) are to be generated, where the goal of the research is explication of meanings or microsocial processes, or where the subject of an investigation is an

entire population. To researchers studying special institutions, regions, or populations, selection criteria are different from those required to generate a representative or stratified, random sample. The goal under these circumstances is the development of findings that may be compared and contrasted with many other groups.

Threats to the external validity of ethnographic findings are those effects that obstruct or reduce a study's comparability and translatability. The fieldworker's problem is to demonstrate what Wolcott (1973) conceptualizes as the typicality of a phenomenon, or the extent to which it compares and contrasts along relevant dimensions with other phenomena. Consequently, external validity depends on the identification and description of those characteristics of phenomena salient for comparison with other, similar types. Once the typicality of a phenomenon is established, bases for comparison may be assumed.

This problem is addressed to an extent by multisite ethnographic designs. The classic model for this approach in educational anthropology, Whiting's investigation (1963) of child-rearing practices in six different cultures, incorporated ethnography into a multimethod investigation (Whiting & Whiting, 1975). Although each of the six teams of field researchers produced an independent ethnography, preentry planning and collection of standardized data for other phases of the study resulted in six investigations of comparable phenomena. More recent multisite studies are variations of the Whitings' design (e.g., Cassell, 1978; Herriott, 1977; Herriott & Gross, 1979; Stake, 1978; Tikunoff, Berliner, & Rist, 1975; Wax, in press; Rist, Note 2). Time spent on site, central integration of data collection and analysis processes, number of field researchers per site, and the nature of final products vary across these studies so they cannot be regarded as a homogeneous solution to threats to external validity. All fail to meet the selection requirements for statistical generalization. Nevertheless, the increase in sample size over single-site studies does strengthen the external validity of their findings (Campbell, 1979). Because sample size is an insufficient condition for confident generalization (Reichardt & Cook, 1979), investigators in multisite studies must address threats to external validity as carefully as do single-site ethnographers. Four factors may affect the credibility of a study for cross-group comparisons: selection effects, setting effects, history effects, and construct effects.

Selection effects. Some constructs cannot be compared across groups because they are specific to a single group or because the researcher mistakenly has chosen groups for which the construct does not obtain. This may occur more frequently when researcher-designated categories are used. Here, the researcher's initial task is to determine the degree of match between the categories and the reality of the group, culture, or setting under investigation. When this is neglected, the categories operate on an assumptive level, and invalid comparisons may be drawn. In circumstances in which a researcher is investigating the cross-group occurrence of participant-derived categories, this may be less likely to obtain: awareness of the participant derivation of the constructs may function as a control for threats to validity. Finally, the discovery that data are absent for the support of a construct may be useful information in itself. In some cases, the ethnographer then may reanalyze the data for contrasts across groups.

The ethnographer's virtual obsession with identifying distinct characteristics of investigated populations derives from a recognition of the significance of this information for comparative purposes. Although characterization may be rendered

partially in subjective qualities, quantitatively measured attributes of populations are essential. Socioeconomic status, levels of education attained, and racial composition are population characteristics that are readily reported in quantitative terms. In his ethnographic analysis of the role of the principal, Wolcott (1973) typifies the individual studied by describing the individual in comparison with the modal category of a nationwide survey of elementary school principals. Goetz (1981) notes that the cultural broker role assumed by the teachers in her investigation of sex-role enculturation may be dependent on their particular relationship with the community serviced by their school. Cusick (1973) limits his explanations for patterns of student-teacher exchange to schools servicing student populations that are comparable to the groups he examined.

Setting effects. Simply by studying a group, culture, or setting, the investigator affects it in some ways. Constructs generated in one context may not be comparable in others because they are a function of context-under-investigation rather than of context only. The reactive observer effects, discussed above as threats to internal validity, are equally serious when cross-group comparisons are conducted. When the construct is a function of observer-setting interaction, it may be treated as equivalent only for groups being observed in a comparable manner, and the interactive dynamics should be identified clearly. Limitations attributed to school ethnographies conducted by participant observers who functioned as teachers (e.g., King, 1967; McPherson, 1972; Rosenfeld, 1971; Wolcott, 1967) stem from the possibility that findings were distorted by observation-setting interaction effects. Smith and Geoffrey (1968) sought to avoid this problem by collecting observations from two perspectives: teacher and nonparticipant researcher. Wolcott (1974) and King (1974) addressed the issue with retrospective analyses of the dynamics of interactions in their respective settings.

Oversaturation of settings is a second facet of this problem. It relates to group history, a third threat to external validity. Groups and cultures that attract continual or intermittant investigation by social scientists may be assumed to be different from groups and cultures with few or no such experiences. Educational researchers are familiar with this problem as it arises in school districts adjacent to research centers; research activities become so integrated with ongoing teaching and administration that the population is altered permanently. Caudill (1963) cites Appalachia as a subculture that has experienced cyclic attention from scholars. He claims that mountaineers have developed a cautious, cynical response to researchers and practitioners based on repeated experiences of disappointed expectations.

History effects. Cross-group comparison of constructs may be invalid due to the unique historical experiences of groups and cultures. Researchers are cautioned, for example, in making comparisons between black slavery in the United States and in Latin America. Nevertheless, careful identification of differing historical variables and subsequent comparison of discrepant cases have proved fruitful (e.g., Elkins, 1959).

In his investigation of schooling in a small German village, Spindler (1974) outlined the community's ongoing conversion from rural to urban orientation and the school's introduction of a nationally disseminated curriculum innovation. Cross-site comparison of Spindler's findings with Warren's (1967) earlier study of a school in a similar village would locate urbanizing developments in both places, but would have to take into account variations in school curriculum stemming from the 10-year time differential. Failure to consider differences between groups resulting from

historical factors may result in the misapplication of constructs and the assumption that phenomena are equivalent across groups.

The opposite assumption, that all group phenomena are unique, is equally misleading. Studstill (1979) has noted the ethnocentric restriction of school studies to complex technological societies. He attributes this to the unquestioned assumption that schools in nonliterate societies have little or nothing in common with the bureaucratic organizations predominant in industrial cultures, despite evidence to the contrary (cf., e.g., Hansen, 1979). Studstill suggests that the failure to identify clearly both common and contrastive features of schools in nonliterate and literate societies has led to the attribution of undeserved uniqueness to schools in complex technological societies.

Construct effects. Construct validity is defined by Cook and Campbell (1979) as the extent to which abstract terms, generalizations, or meanings are shared across times, settings, and populations. This can be interpreted in several ways. Definitions and meanings of terms and constructs can vary (see *Analytic constructs and premises* above).

A second interpretation of construct validity concerns how the effects of observed phenomena are construed. Explanations regarded as valid among some groups are discounted by others (see *Spurious conclusions* above). Construct validity also may refer to the degree to which instructions for and formats of instruments are mutually intelligible to the instrument designer, to the instrument administrator, and to the subjects to whom the instrument is applied (see *Observer effects* above).

The comparability of ethnographic studies may be reduced or obstructed by idiosyncratic use of initial analytic constructs or by generation of constructs so peculiar to a particular group that they are useless for cross-group examinations. Cook and Campbell (1979) accord sufficient gravity to threats to the construct validity of instruments used in experimental research that they discuss the issue independently of internal and external validity. A number of the effects discussed above affect construct validity. Because a major outcome of ethnographic research is the generation and refinement—through cross-group applications—of constructs, ethnographers must consider issues of construct validity as critical to the credibility of their results.

The elicitation techniques used by ethnographers are designed specifically to intensify equivalence in meaning and interpretation between researchers and their subjects. A variety of strategies, including listing, Q-sorting, and constant cross-checking in discussions with informants, support this objective (LeCompte & Goetz, in press). Triangulating many data sources (Denzin, 1978) formalizes the meanings which participants attribute to phenomena. Team research and peer review serve as audits to ensure that interpretation of mundane phenomena are examined rather than assumed.

Where disparities are identified, ethnographers report them as attributes of the groups being examined (e.g., Ogbu, 1974; Smith & Keith, 1971; Wolcott, 1973). This sensitizes other researchers, directing them to examine comparable effects in other populations. In cases where particular group dimensions require the customizing of instruments or initial analytic constructs (e.g., Modiano, 1973; Spindler, 1973, 1974), a common requirement in cross-cultural studies, modifications are included in the presentation of results.

Conclusion

A serious problem in assessing the credibility of ethnographic research, which may be peculiar to this investigative tradition, is that addressing all of the categories of contamination and bias may appear to mandate contradictory measures. Although the term ethnography has been used throughout this discussion as referring to a research process, ethnography also refers to the product of a research effort. It is defined by anthropologists as an analytic description of an intact cultural scene (Spradley & McCurdy, 1972). It delineates the shared beliefs, practices, artifacts, folk knowledge, and behaviors of a group of people. Its objective is the holistic reconstruction of the culture or phenomena investigated.

Given this goal, the ethnographer's primary commitment is to a faithful and accurate rendition of the participant's lifeways. To the extent that these may be eccentric, singular, or idiosyncratic when compared to other groups, they still require reporting. As a consequence, some ethnographers may resist formulation of the constructs and postulates applicable to other groups that are prerequisite for establishing external validity and external reliability. Tailoring these abstractions for cross-group comparisons may appear to distort their derivation. Thus, while confronting the possibility of obtaining noncomparable data and results is a risk undertaken both by experimental and ethnographic researchers, it is perhaps more serious for the latter.

Such dilemmas, which are frequently discussed as dichotomous choices between subjective or objective data and data analysis processes, between replicability or authenticity, between representativeness of samples or purposive sampling, or between generalizability or uniqueness of results (Filstead, 1979; Rist, 1977; Wilson, 1977), are shared across social science research designs. While dichotomous conceptualization of these issues may be useful pedagogically, it may distort and mislead when used to assess and design research activity (Reichardt & Cook, 1979). Educational anthropologists and other educational researchers who use qualitative strategies may be susceptible to viewing these alternatives as mutually exclusive choices because, in their shared commitment to the improvement of curriculum, instruction, and other factors in education, they seek research designs that purport to demonstrate clear-cut causality and that allege to be distortion free. In addition, public and academic concern for direct applicability of educational research overshadows investigation in the field and provokes what are frequently simplistic interpretations of designs and results. The inevitable outcome is that research results rarely are functional in or applicable to real classrooms.

Our position is that the transformation of such issues into dichotomous choices is unnecessary, inaccurate, and ultimately counterproductive. Many research studies include the collection of both objective and subjective data. Similarly, the same investigations may employ data-analysis strategies that range from subjective to objective (Scriven, 1972). Replicability, often viewed as merely a function of standardization of instruments and procedures, is a complex issue that must be addressed by various strategies. If sampling is viewed as a collection of overlapping alternatives to a variety of design problems, then its assessment will depend on how well problems are solved rather than on conformity to a randomness seldom achieved even in quantitative studies. The extent to which results are generalizable or unique depends

on such factors as the level of abstraction addressed and will vary by particular construct or relationship posited.

Attaining absolute validity and reliability is an impossible goal for any research model. Nevertheless, investigators may approach these objectives by conscientious balancing of the various factors enhancing credibility within the context of their particular research problems and goals. For decades, reputable ethnographers have used a variety of strategies to reduce threats to reliability and validity. This has been a major source for one of the defining characteristics of present-day ethnography—its multimodality.

Reference Notes

1. Goetz, J. P., & LeCompte, M. D. *The role of theory in the design of educational ethnography.* Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, April 1981.
2. Rist, R. C. *The youthwork national policy study.* Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, April 1981.
3. Wolcott, H. F. *How to look like an anthropologist without really being one.* Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.
4. Brown, M. J. M. *Triangulation: An approach to the analysis of how one school system implemented a state mandated curriculum innovation.* Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

References

Appelbaum, R. P. *Theories of social change.* Chicago: Markham, 1970.

Argyris, C. Diagnosing defenses against the outsider. *Journal of Social Issues*, 1952, *8*(3), 24–34.

Barker, R. G., & Wright, H. F. *Midwest and its children.* Evanston, Ill.: Row, Peterson, 1954.

Becker, H., Geer, B., & Hughes, E. *Making the grade: The academic side of college life.* New York: Wiley, 1968.

Becker, H., Geer, B., Hughes, E., & Strauss, A. *Boys in white: Student culture in medical school.* Chicago: University of Chicago Press, 1961.

Berreman, G. D. Behind many masks: Ethnography and impression management in a Himalayan village. *Monographs of the Society for Applied Anthropology*, 1962, No. 4.

Biddle, B. J. Methods and concepts in classroom research. *Review of Educational Research*, 1967, *37*, 337–357.

Borman, K. Social control and schooling: Power and process in two kindergarten settings. *Anthropology and Education Quarterly*, 1978, *9*, 38–53.

Bossert, S. T. *Tasks and social relationships in classrooms: A study of instructional organization and its consequences.* Cambridge: Cambridge University Press, 1979.

Burnett, J. H. Ceremony, rites, and economy in the student system of an American high school. *Human Organization*, 1969, *28*, 1–10.

Burns, A. F. On the ethnographic process in anthropology and education. *Anthropology and Education Quarterly*, 1976, *7*(3), 25–33.

Campbell, D. T. "Degrees of freedom" and the case study. In T. D. Cook & C. S. Reichardt (Eds.), *Qualitative and quantitative methods in evaluation research.* Beverly Hills, Calif.: Sage, 1979.

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

Carroll, T. G. Work and play: A probe of the formation, use and intersection of adult and child activity domains (Doctoral dissertation, SUNY, Buffalo, 1976). *Dissertation Abstracts International*, 1977, *38*, 5211A–5212A. (University Microfilms No. 77-3520).

Cassell, J. *A fieldwork manual for studying desegregated schools.* Washington, D.C.: The National Institute of Education, 1978.

Caudill, H. M. *Night comes to the Cumberlands: A biography of a depressed area.* Boston: Little, Brown, 1963.

Cicourel, A. V. *Method and measurement in sociology.* New York: Free Press, 1964.

Cicourel, A. V., & Kitsuse, J. I. *The educational decision-makers.* Indianapolis: Bobbs-Merrill, 1963.

Clement, D. C., & Harding, J. R. Social distinctions and emergent student groups in a desegregated school. *Anthropology and Education Quarterly*, 1978, *9*, 272–282.

Collier, J., Jr. *Alaskan Eskimo education: A film analysis of cultural confrontation in the schools.* New York: Holt, Rinehart, & Winston, 1973.

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

Crain, R. L. Racial tensions in high schools: Pushing the survey method closer to reality. *Anthropology and Education Quarterly*, 1977, *8*, 142–151.

Cusick, P. A. *Inside high school: The student's world.* New York: Holt, Rinehart, & Winston, 1973.

Dean, J. P., Eichhorn, R., & Dean L. R. Fruitful informants for intensive interviewing. In J. T. Doby (Ed.), *An introduction to social research* (2nd ed.). New York: Meredith, 1967.

Dean, J. P., & Whyte, W. F. How do you know if the informant is telling the truth?. *Human Organization*, 1958, *17*(2), 34–38.

Denzin, N. K. *The research act: A theoretical introduction to sociological methods* (2nd ed.). New York: McGraw-Hill, 1978.

Eddy, E. M. *Becoming a teacher: The passage to professional status.* New York: Teachers College Press, 1969.

Elkins, S. M. *Slavery: A problem in American institutional and intellectual life.* New York: Grosset & Dunlap, 1959.

Erikson, E. H. Observations on the Yurok: Childhood and world image. *Publications in American Archaeology and Ethnology*, 1943, *35*(10), I-VII, 257–302.

Erickson, F. What makes school ethnography "ethnographic?". *Anthropology and Education Quarterly*, 1973, *4*(2), 10–19.

Erickson, F. Some approaches to inquiry in school-community ethnography. *Anthropology and Education Quarterly*, 1977, *8*, 58–69.

Evans-Pritchard, E. E. Social anthropology: Past and present. In *Essays in social anthropology.* New York: Free Press, 1962.

Everhart, R. B. Between stranger and friend: Some consequences of "long term" fieldwork in schools. *American Educational Research Journal*, 1977, *14*, 1–15.

Filstead, W. J. Qualitative methods: A needed perspective in evaluation research. In T. D. Cook & C. S. Reichardt (Eds.), *Qualitative and quantitative methods in evaluation research.* Beverly Hills, Calif.: Sage, 1979.

Flanders, N. A. *Analyzing teacher behavior.* Reading, Mass.: Addison-Wesley, 1970.

Fuchs, E. *Pickets at the gate.* New York: Free Press, 1966.

Fuchs, E. *Teachers talk: Views from inside city schools.* New York: Anchor, 1969.

Furlong, V. Interaction sets in the classroom: Towards a study of pupil knowledge. In M. Stubbs & S. Delamont (Eds.), *Explorations in classroom observations.* London: Wiley, 1976.

Gay, J., & Cole, M. *The new mathematics and an old culture: A study of learning among the Kpelle of Liberia.* New York: Holt, Rinehart, & Winston, 1967.

Gearing, F. O. Where we are and where we might go: Steps toward a general theory of cultural transmission. *Anthropology and Education Quarterly*, 1973, *4*(1), 1–10.

Gearing, F. O. Overview: A cultural theory of education. *Anthropology and Education Quarterly*, 1975, *6*(2), 1–9.

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

Goetz, J. P. Configurations in control and autonomy: A microethnography of a rural third-grade classroom (Doctoral dissertation, Indiana University, 1975). *Dissertation Abstracts International*, 1976, *36*, 6175A. (University Microfilms No. 76-6275).

Goetz, J. P. Sex-role systems in Rose Elementary School: Change and tradition in the rural-transitional South. In R. T. Sieber & A. J. Gordon (Eds.), *Children and their organizations: Investigations in American culture*. Boston: G. K. Hall, 1981.

Goetz, J. P., & LeCompte, M. D. Ethnographic research and the problem of data reduction. *Anthropology and Education Quarterly*, 1981, *12*, 51–70.

Gold, R. Roles in sociological field observations. *Social Forces*, 1958, *36*, 217–223.

Goodman, M. E. Values, attitudes, and social concepts of Japanese and American children. *American Anthropologist*, 1957, *59*, 979–999.

Hammond, D., & Jablow, A. *Women in cultures of the world*. Menlo Park, Calif.: Cummings, 1976.

Hansen, J. F. *Sociocultural perspectives on human learning: An introduction to educational anthropology*. Englewood Cliffs, N.J.: Prentice-Hall, 1979.

Harris, M. *Culture, man, and nature*. New York: Crowell, 1971.

Henry, J. A cross-cultural outline of education. *Current Anthropology*, 1960, *1*(4), 267–305.

Herriott, R. E. Ethnographic case studies in federally funded multidisciplinary policy research: Some design and implementation issues. *Anthropology and Education Quarterly*, 1977, *8*, 106–115.

Herriott, R. E., & Gross, N. (Eds.). *The dynamics of planned educational change: Case studies and analyses*. Berkeley, Calif.: McCutchan, 1979.

Hilgar, M. I. *Field guide to the ethnological study of child life* (2nd ed.). New Haven, Conn.: Human Relations Area Files Press, 1966.

Hollingshead, A. B. *Elmtown's youth and Elmtown revisited*. New York: Wiley, 1975.

Hostetler, J. A., & Huntington, G. E. *Children in Amish society: Socialization and community education*. New York: Holt, Rinehart, & Winston, 1971.

Howard, A. *Learning to be Rotuman: Enculturation in the South Pacific*. New York: Teachers College Press, 1970.

Ianni, F. A. J. Anthropology and educational research: A report on federal agency programs, policies and issues. *Anthropology and Education Quarterly*, 1976, *7*(3), 3–11.

Janes, R. W. A note on phases of the community role of the participant-observer. *American Sociological Review*, 1961, *26*, 446–450.

Jocano, F. L. *Growing up in a Philippine barrio*. New York: Holt, Rinehart, & Winston, 1969.

Jorgensen, J. G. On ethics and anthropology. *Current Anthropology*, 1971, *12*, 321–334.

Kahn, R., & Mann, F. Developing research partnerships. *Journal of Social Issues*, 1952, *8*(3), 4–10.

Kaplan, D., & Manners, R. A. *Culture theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Khleif, B. B. Issues in anthropological fieldwork in the schools. In G. D. Spindler (Ed.), *Education and cultural process: Toward an anthropology of education*. New York: Holt, Rinehart, & Winston, 1974.

King, R. A. *The school at Mopass: A problem of identity*. New York: Holt, Rinehart, & Winston, 1967.

King, R. A. The teacher as participant-observer: A case study. In G. D. Spindler (Ed.), *Education and cultural process: Toward an anthropology of education*. New York: Holt, Rinehart, & Winston, 1974.

LeCompte, M. D., & Goetz, J. P. Ethnographic data collection in evaluation research. *Educational Evaluation and Policy Analysis*, in press.

Leemon, T. A. *The rites of passage in a student culture: A study of the dynamics of transition.* New York: Teachers College Press, 1972.

Lewis, O. *Life in a Mexican village: Tepoztlan restudied.* Urbana, Ill.: University of Illinois Press, 1951.

Linton, R. *The cultural background of personality.* New York: Appleton-Century-Crofts, 1945.

Lofland, J. *Analyzing social settings: A guide to qualitative observation and analysis.* Belmont, Calif.: Wadsworth, 1971.

Magoon, A. J. Constructivist approaches in educational research. *Review of Educational Research*, 1977, *47*, 651–693.

McCall, G. J. Data quality control in participant observation. In G. J. McCall & J. L. Simmons (Eds.), *Issues in participant observation: A text and reader.* Reading, Mass.: Addison-Wesley, 1969.

McPherson, G. *Small town teacher.* Cambridge, Mass.: Harvard University Press, 1972.

Mead, M. *New lives for old: Cultural transformations—Manus, 1929–1953*, New York: Morrow, 1956.

Mehan, H. Assessing children's school performance. In M. Hammersley & P. Woods (Eds.), *The process of schooling: A sociological reader.* London: Routledge & Kegan Paul in association with the Open University Press, 1976.

Mehan, H. *Learning lessons: Social organization in the classroom.* Cambridge, Mass.: Harvard University Press, 1979.

Mehan, H., & Griffin, P. Socialization: The view from classroom interactions. *Sociological Inquiry*, 1980, *50*(3–4), 357–398.

Miller, S. M. The participant observer and "over-rapport." *American Sociological Review*, 1952, *17*, 97–99.

Modiano, N. *Indian education in the Chiapus highlands.* New York: Holt, Rinehart, & Winston, 1973.

Moore, A. *Life cycles in Atchálan: The diverse careers of certain Guatemalans.* New York: Teachers College Press, 1973.

Naroll, R. *Data quality control: A new research technique.* New York: Free Press, 1962.

Ogbu, J. U. *The next generation: An ethnography of education in an urban neighborhood.* New York: Academic Press, 1974.

Patton, M. Q. *Qualitative evaluation methods.* Beverly Hills, Calif.: Sage, 1980.

Paulme, D. (Ed.). *Women of tropical Africa.* Berkeley: University of California Press, 1963.

Pelto, P. J., & Pelto, G. H. *Anthropological research: The structure of inquiry* (2nd ed.). Cambridge: Cambridge University Press, 1978.

Peshkin, A. *Growing up American: Schooling and the survival of community.* Chicago: University of Chicago Press, 1978.

Petrie, H. G. Do you see what I see? The epistemology of interdisciplinary research. *Educational Researcher*, 1976, *5*(2), 9–15.

Phillips, D. L. *Knowledge from what? Theories and methods in social research.* Chicago: Rand McNally, 1971.

Powdermaker, H. *Stranger and friend.* New York: Norton, 1966.

Redfield, R. *Tepoztlan—A Mexican village.* Chicago: University of Chicago Press, 1930.

Reichardt, C. S., & Cook, T. D. Beyond qualitative versus quantitative methods. In T. D. Cook & C. S. Reichardt (Eds.), *Qualitative and quantitative methods in evaluation research.* Beverly Hills, Calif.: Sage, 1979.

Reiter, R. R. *Toward an anthropology of women.* New York: Monthly Review Press, 1975.

Riner, R. D. American Indian education: A rite that fails. *Anthropology and Education Quarterly*, 1979, *10*, 236–253.

Rist, R. C. Student social class and teacher expectations: The self-fulfilling prophecy in ghetto education. *Harvard Educational Review*, 1970, *40*, 411–451.

Rist, R. C. On the relations among educational research paradigms: From disdain to detente. *Anthropology and Educational Quarterly*, 1977, *8*, 42–49.

Robinson, W. S. The logical structure of analytic induction. *American Sociological Review*, 1951, *16*, 812–818.

Rosenfeld, G. *"Shut those thick lips!": A study of slum school failure.* New York: Holt, Rinehart & Winston, 1971.

Rynkiewich, M. A., & Spradley, J. P. (Eds.). *Ethics and anthropology: Dilemmas in fieldwork.* New York: Wiley, 1976.

Schatzman, L., & Strauss, A. *Field research: Strategies for a natural sociology.* Englewood Cliffs, N.J.: Prentice-Hall, 1973.

Schwartz, M. S., & Schwartz, C. G. Problems in participant observation. *American Journal of Sociology*, 1955, *60*, 343–354.

Scriven, M. Objectivity and subjectivity in educational research. In L. G. Thomas (Ed.), *Philosophical redirection of educational research* (71st Yearbook, Part I). Chicago: National Society for the Study of Education, 1972.

Scriven, M. Evaluation perspectives and procedures. In J. W. Popham (Ed.), *Evaluation in education: Current applications.* Berkeley, Calif.: McCutchan, 1974.

Sieber, R. T. Classmates as workmates: Informal peer activity in an elementary school. *Anthropology and Education Quarterly*, 1979, *10*, 207–235.

Sieber, R. T. Many roles, many faces: Researching school-community relations in a heterogeneous American urban community. In D. A. Messerschmidt (Ed.), *Anthropologists at home: Toward an anthropology of issues in America.* New York: Cambridge University Press, in press.

Smith, L. M. Reflections on trying to theorize from ethnographic data. *Anthropology and Education Quarterly*, 1974, *5*(1), 18–24.

Smith, L. M. An evolving logic of participant observation, educational ethnography, and other case studies. *Review of Research in Education*, 1979, *6*, 316–377.

Smith, L. M., & Brock, J. A. M. *"Go, bug, go!": Methodological issues in classroom observational research.* St. Louis: Central Midwestern Regional Educational Laboratory, 1970.

Smith, L. M., & Geoffrey, W. *The complexities of an urban classroom: An analysis toward a general theory of teaching.* New York: Holt, Rinehart, & Winston, 1968.

Smith, L. M., & Keith, P. M. *Anatomy of educational innovation: An organizational analysis of an elementary school.* New York: Wiley, 1971.

Spindler, G. D. *Burgbach: Urbanization and identity in a German village.* New York: Holt, Rinehart, & Winston, 1973.

Spindler, G. D. Schooling in Schönhausen: A study in cultural transmission and instrumental adaptation in an urbanizing German village. In G. D. Spindler (Ed.), *Education and cultural process: Toward an anthropology of education.* New York: Holt, Rinehart, & Winston, 1974.

Spradley, J. P. *The ethnographic interview.* New York: Holt, Rinehart, & Winston, 1979.

Spradley, J. P. *Participant observation.* New York: Holt, Rinehart, & Winston, 1980.

Spradley, J. P., & McCurdy, D. W. *The cultural experience: Ethnography in complex society.* Chicago: Science Research Associates, 1972.

Stake, R. E. *Case studies in science education.* (2 vols.). Washington, D.C.: U.S. Government Printing Office, 1978. (ERIC Document Reproduction Service Nos. ED 166 058 & ED 166 059.)

Studstill, J. P. Education in a Luba secret society. *Anthropology and Education Quarterly*, 1979, *10*, 67–79.

Tikunoff, W. J., Berliner, D., & Rist, R. C. *Special study A: An ethnographic study of 40 classrooms of the BTES known sample.* San Francisco: Far West Laboratory for Educational Development, 1975. (ERIC Document Reproduction Service No. ED 150 110).

Van Gennep, A. *The rites of passage.* Chicago: University of Chicago Press, 1960.

Vidich, A. J. Participant observation and the collection and interpretation of data. *American Journal of Sociology*, 1955, *60*, 354–360.

Ward, M. C. *Them children: A study of language learning.* New York: Holt, Rinehart, & Winston, 1971.

Warren, R. L. *Education in Rebhausen: A German village.* New York: Holt, Rinehart, & Winston, 1967.

Wax, M. L. (Ed.). *When schools are desegregated: Problems and possibilities for students, educators, and the community.* New Brunswick, N.J.: Transaction Books, in press.

Wax, R. H. *Doing fieldwork: Warnings and advice.* Chicago: University of Chicago Press, 1971.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. *Unobtrusive measures: Nonreactive research in the social sciences.* Chicago: Rand McNally, 1966.

Whiting, B. B. (Ed.). *Six cultures: Studies of child rearing.* New York: Wiley, 1963.

Whiting, B. B., & Whiting, J. W. M. *Children of six cultures: A psychocultural analysis.* Cambridge, Mass.: Harvard University Press, 1975.

Whiting, J. W. M., Child, I. L., & Lambert, W. W. *Field guide for a study of socialization.* Huntington, N.Y.: Krieger, 1966.

Whyte, W. F. *Street corner society: The social structure of an Italian slum* (2nd ed.). Chicago: University of Chicago Press, 1955.

Williams, T. R. *Field methods in the study of culture.* New York: Holt, Rinehart, & Winston, 1967.

Wilson, S. The use of ethnographic techniques in educational research. *Review of Educational Research,* 1977, *47,* 245–265.

Wolcott, H. F. *A Kwakiutl village and school.* New York: Holt, Rinehart, & Winston, 1967.

Wolcott, H. F. Handle with care: Necessary precautions in the anthropology of schools. In M. L. Wax, S. Diamond, & F. O. Gearing (Eds.), *Anthropological perspectives on education.* New York: Basic Books, 1971.

Wolcott, H. F. *The man in the principal's office: An ethnography.* New York: Holt, Rinehart, & Winston, 1973.

Wolcott, H. F. The teacher as an enemy. In G. D. Spindler (Ed.), *Education and cultural process: Toward an anthropology of education.* New York: Holt, Rinehart, & Winston, 1974.

Wolcott, H. F. Criteria for an ethnographic approach to research in schools. *Human Organization,* 1975, *34,* 111–127.

Wolcott, H. F. *Teachers versus technocrats: An educational innovation in anthropological perspective.* Eugene, Ore.: Center for Educational Policy and Management, University of Oregon, 1977.

Wylie, L. *Village in the Vaucluse* (3rd ed.). Cambridge, Mass.: Harvard University Press, 1974.

Znaniecki, F. *The method of sociology.* New York: Farrar & Rinehart, 1934.

AUTHORS

MARGARET D. LECOMPTE, Associate Director, Program Evaluation, Department of Research and Evaluation, Houston Independent School District, 3830 Richmond Avenue, Houston, TX 77027. *Specialization:* Research methods; classroom ethnography; educational sociology.

JUDITH PREISSLE GOETZ, Associate Professor, Social Science Education, Dudley Hall 200, University of Georgia, Athens, GA 30602. *Specialization:* Educational anthropology.