AK: Still, I am astonished to find that in some circles you have come to be seen as a reactionary.

ERL: I don't find it so strange. there can be guile in these matters. Here in King's I am certainly seen as very conservative. It has already been forgotten that women became members of the College for the first time (after 525 years) while I was Provost. It is an interesting story.

To make the proposal plausible, we had to have allies. There was much backstage negotiation with our neighbours, Clare and Churchill. Both these colleges eventually first admitted women at the same time as King's. But they had problems. In Clare the Master was enthusiastic; at Churchill the Master was opposed. The Clare story was that the Old Guard dug up a retired fellow who had not been seen for 20 years. He turned up in a Bath chair and cast his vote against the admission of women. So they were back at square one.

We took things slowly and talked it out for a couple of years or perhaps more. The opposition were clearly in a minority, but we had to win by a two-thirds majority of the whole Governing Body (that is, all the Fellows, about 100 individuals). At the last minute, when I thought all was lost, the leader of the opposition got up and made a highly emotional speech to the effect that this was the most important vote that the College had faced during the last 500 years. He could see that his supporters were in a minority, so he asked them to change sides and let this historic decision go through on a unanimous vote. One splendid old boy refused to take that one and said he must abstain. So King's voted for the admission of women *nem. con.!*

That could only have happened in King's. But there was some anthropology in it, too. If I had not already begun to acquire the reputation of being much more conservative than my erratic predecessor, I doubt if we would have made it at all. I am no fisherman, but it was all rather like landing a salmon.

References Cited

EVANS-PRITCHARD, E. E. 1937. *Witchcraft, oracles, and magic among the Azande*. Oxford: Clarendon Press.

———. 1940*a*. *The Nuer*. Oxford: Clarendon Press.

———. Editor. 1940*b*. *African political systems*. London: Oxford University Press.

FIRTH, RAYMOND. 1946. *Malay fishermen: Their peasant economy*. London: Routledge and Kegan Paul.

HOWORTH, HENRY H. 1876–1927. *History of the Mongols from the 9th to the 19th century*. 5 vols. London: Longmans, Green.

———. 1887. *The mammoth and the flood: An attempt to confront the theory of uniformity with the facts of recent geology*. London: Sampson Low.

———. 1917. *The golden days of the early English church from the arrival of Theodore to the death of Bede*. 3 vols. London: John Murray.

LEACH, E. R. 1940. *Social and economic organisation of the Rowanduz Kurds*. London: London School of Economics.

———. 1945. Jinghpaw kinship terminology. *Journal of the Royal Anthropological Institute* 75:59–72.

———. 1951. The structural implications of matrilateral cross-cousin marriage. *Journal of the Royal Anthropological Institute* 81:23–55.

———. 1954. *Political systems of highland Burma*. London: Bell.

———. 1976. *Culture and communication: The logic by which symbols are connected*. Cambridge: Cambridge University Press.

LÉVI-STRAUSS, CLAUDE. 1949. *Les structures élémentaires de la parenté*. Paris: Presses Universitaires de France.

———. 1953. "Social structure," in *Anthropology today*. Edited by A. L. Kroeber. Chicago: University of Chicago Press.

RICHARDS, AUDREY. 1952. Review of: *Les structures élémentaires de la parenté*, by Claude Lévi-Strauss (Paris: Presses Universitaires de France, 1949). *Man* 52:12–13.

SAHLINS, MARSHALL. 1985. *Islands of history*. Chicago: University of Chicago Press.

VALERI, V. 1985. *Kingship and sacrifice: Ritual and society in ancient Hawaii*. Chicago: University of Chicago Press.

YALMAN, NUR. 1967. *Under the bo tree: Studies in caste, kinship, and marriage in the interior of Ceylon*. Berkeley: University of California Press.

The Construction of Primary Data in Cultural Anthropology[1]

by H. RUSSELL BERNARD, PERTTI J. PELTO, OSWALD WERNER, JAMES BOSTER, A. KIMBALL ROMNEY, ALLEN JOHNSON, CAROL R. EMBER, and ALICE KASAKOFF[2]

c/o Department of Anthropology, University of Florida, Gainesville, Fla. 32611, U.S.A. 2 IV 86

In April 1985, a conference was held at the National Science Foundation in Washington, D.C., on the construction of primary data in cultural anthropology. We were concerned with data gathering rather than with data analysis. We felt that, while much attention had been focused on data analysis, particularly on complex statistical analysis, insufficient attention had been paid to issues of validity, reliability, and accuracy of our basic raw materials. Recognizing that once bias has entered into data, statistical analysis will not get it out, we were concerned with improving the construction of primary data in cultural anthropology. To address this issue we felt it useful to focus discussion on the four main types of data construction in our discipline: (1) relatively open-ended, unstructured interviews with key informants, (2) structured interviews of respondents who, in the case of surveys, may number in the hundreds or thousands, (3) direct observation of behavior and environmental features, and (4) extraction of information from existing records such as native texts, court proceedings, marriage records, and so on. In the case of the Human Relations Area Files, this last method involves the coding of the ethnographic work of colleagues.

This concern with the quality of primary data is certainly not new in anthropology. Malinowski addressed it forcefully in the introduction to *Argonauts of the Western Pacific,* and it has remained in the forefront of British social anthropology (see Epstein 1967). Boas made it one of the cornerstones of his research program. The ethnoscience perspective of the '60s was based largely on a concern for validity (what do the *natives really think?*) and for reliability (do question frames produce the same results over a series of informants or from a single informant interviewed by a series of anthropologists?). More recently, anthropologists have become concerned also with accuracy and precision of data. For example, a question such as "How often does it rain around here?" may tap an informant's cognition about the periodicity of rainfall, and if that is what the researcher *intends* to get at, then it is a valid

[1]

[2] With John Roberts, William Dressler, Roy D'Andrade, Patricia Draper, John W. Adams, Lee Sailer, Daniel Gross, Ronald Cohen, Christopher McCarty, Joel Gittelsohn, and Brian Johnstone. This paper is the result of a conference convened by Bernard and Pelto, who contributed the introduction and conclusions and edited the draft sections of the conference report. The section on unstructured interviewing was contributed by Boster and Werner and the section on structured interviewing by Romney; Johnson wrote the section on direct observation and Ember and Kasakoff the section on the use of archival sources. All components of the manuscript were sent to all participants in the conference for comment and emendation. The final product is truly a joint effort.

instrument for measuring that particular cognition. But if the researcher really wants to know how much it rains in the area, then the instrument may be "poorly calibrated" and may produce inaccurate results when measured against an objective index of rainfall. Furthermore, if an informant says "It rains a lot," then that is a very imprecise result as measurements go. But if the informant says "It rains 60 inches a year" (a precise answer) when it really rains 120 inches, then the less precise measure may cause less mischief than the precise one. (Experiments on the extent of informant accuracy are reported in Killworth and Bernard 1976, 1979; Bernard and Killworth 1977; Bernard, Killworth, and Sailer 1980, 1982; the issues are reviewed in Bernard et al. 1984.)

Anthropologists are increasingly aware of these problems and of the fact that every field situation is different and can never be duplicated. Our relationships with our informants are unique, and we study communities that are changing rapidly. Thus, each of us has the responsibility to make sure that data are collected in such a way as to document fully the influence of the particular situation on the results of the research and in such a way that others can use them. The focus at the conference, then, was on the state of the art of data collection:

What are the problems involved in getting valid, reliable, accurate, and precise data from the field or extracting it from archival sources?

What research do we have concerning data quality control under various conditions?

What are the current solutions to these problems?

What training do cultural anthropologists receive to ensure data quality?

What needs to be done in the area of research and training to improve data quality?

We assume, first of all, that the logic of data construction is the same in all sciences and that facts about the world exist and can be discovered. We assume that one does not simply discover facts through a pristine encounter with the world. Rather, the acquisition of primary data—the act of observation and measurement—is driven by theory, even if the "theory" is only a loosely assembled series of biases and hunches.

By "construction of primary data" we mean an interactive process between a researcher, a theory, and the research materials under study, whether they be people in the field or documents to be examined. Moreover, we acknowledge that data construction goes on after one leaves the field. The constructs and variables around which papers are written take their final shape through qualitative, formal and statistical analysis. "Theory," then, operates at all levels of observation and measurement: in deciding what to observe, what to measure, how to measure it, and what the final forms and names of variables should be after data have been gathered. We leave aside for the moment the issue of theory in data generation. Regardless of the relative merits of alternative theories, there seems to be general agreement that primary data should be as reliable and valid as possible. The basic questions, then, are: How well is cultural anthropology doing in data quality control? How can we improve our performance?

There are many lessons to be learned from work in other disciplines. In sociology, for example, the literature on response effects in surveys is extensive (see Sudman and Bradburn 1974, 1982 and Cannell, Marquis, and Laurent 1977 for reviews). This literature does not deal with problems of translation, however, and few anthropologists have found them interesting (see Bohannan 1954, Casagrande 1954, Colby 1966, Nida 1964, Phillips 1959, Schoua 1985, Voegelin and Voegelin 1967, Werner and Campbell 1970). Psychologists, too, have conducted many experiments on interaction effects between observers and their subjects (Rosenthal 1966, Rosenthal and Jacobson 1968). Epidemiologists and other medical researchers have developed a considerable literature on "decay rates" in respondents' recall of past illness events and related topics (see Barron 1977; Brown and Harris 1978; Herrman 1985*a, b;* Tilley et al. 1985). Nutrition researchers have documented the problems of reliability and precision that beset the various methods of measuring food intakes in natural settings (see Beaton 1982, Beaton et al. 1983, Block 1982, Byers et al. 1983, Chu et al. 1984, Humble, Samet, and Skipper 1984, Karinpaa and Seppanen 1983, Lee, Kolonel, and Hankin 1983, Liu et al. 1978, Lyon et al. 1983, McGee et al. 1982, Moller Jenson et al. 1984, Mullen et al. 1984, Rohan and Potter 1984, Rush and Kristal 1982, Samet, Humble, and Skipper 1984, Sempos et al. 1985, Stuff et al. 1983, Todd, Hudes, and Calloway 1983, Tremblay et al. 1983, C. M. Young 1981).

By and large, though, these methodological works have not been taken very seriously by anthropologists, perhaps because of communication gaps between disciplines and perhaps because many anthropologists consider them irrelevant to the rough-and-tumble circumstances of cross-cultural fieldwork. Perhaps anthropologists feel that quality control is not so much an issue because of the implicit controls arising from long-term participant observation or believe that the standards of precision are simply different in our brand of research (see Holy 1984). Whatever the reason, anthropologists have not systematically addressed the key questions about the quality of their construction of primary data. We held this conference to address these questions and to inventory the strengths and weaknesses of various data collection methods. We hope that among the results of this report will be:

1. Greater awareness of the data quality control literature already available, including not only the pioneering work of Naroll (1962) but also that coming from other behavioral sciences (see Bradburn and Sudman 1979, Cannell, Marquis, and Laurent 1977, Cook and Campbell 1979, Ciminero, Calhoun, and Adams 1977, Kent and Foster 1977, Sudman and Bradburn 1974).

2. A heightened understanding of the need for purely methodological research that focuses on issues of validity, reliability, accuracy, and precision of various methods of data construction, particularly in cross-cultural work (see Werner, Schoepfle, et al. 1985–86).

3. Agreement among our colleagues that studies are needed to improve the construction of primary data in order to enhance the credibility and usefulness of our discipline.

4. A preliminary inventory of the principal issues and research priorities in data quality control.

5. An awareness on the part of research funding organizations of the need for support, both individual and institutional, for appropriate research on data quality control in cultural anthropology.

Cultural anthropology has the potential to become the social science which synthesizes qualitative and quantitative data to address important social issues. To fulfill this potential, we need to do a great deal of work to develop data of the quality necessary to test theories. Anthropologists are not short of theories; the field is filled with good ideas. It is also filled with people who have the stamina and resourcefulness to venture out and observe the social world in its natural settings. The creativity, ideas, and energy of cultural anthropologists should be harnessed to the goal of building a science of humanity. Such a science depends vitally on credible data. Achieving such a science will require a great deal of research and financial support, but the goal is well worth striving for.

Unstructured Interviewing

Unstructured interviewing furnishes data which are usually not strictly comparable across informants. It is the principal data collection procedure in all ethnographic work and, in its various aspects, merges with all other types of anthropological data-collecting techniques. We distinguish four types of unstructured interviewing.

In a typical unstructured interview, topics are explored in depth in a fluid, open-ended conversation with a key informant or informants, with a high degree of interaction between ethnographer and informant and with the ethnographer formulating grand-tour questions (Spradley 1979, 1980) and probing questions on the basis of theoretical interests or of informant responses to past inquiries (Burgess 1982).

In other types of unstructured interviews, the ethnographer elicits data from informants in culturally predefined chunks, including (*a*) personal documents, such as life histories and histories of particular events which are comparable to diaries and personal journals (see du Boulay and Williams 1984, Faraday and Plummer 1979, Kluckhohn 1945, Langness 1965, Lewis 1961, Shaw 1980); (*b*) narratives, such as myths, tales, and other folklore (Dundes 1965, Lévi-Strauss 1963, Tedlock 1971, Vansina 1965); (*c*) cases, such as accounts of disputes, legal transactions, and public events and rituals, which are comparable to court records and journalistic accounts (Mitchell 1984, van Velsen 1967).

A third kind of unstructured interviewing merges with the direct observation of behavior, as when the ethnographer queries an informant about the meaning of some event in progress. This is sometimes called "opportunistic" or "on-the-spot" interviewing or "grass-hut interviewing" (see Frake 1964, Werner 1966).

Finally, there are semistructured interviews in which the ethnographer's queries follow a predetermined pattern but their substantive focus shifts according to the responses of the informant. This form of data elicitation has been explored by sociologists and social psychologists (see Cannell and Kahn 1968, Houston and Sudman 1975, Hyman and Cobb 1975, Merton 1956). Many of the eliciting procedures of classic ethnoscience are of this form, including the elicitation of taxonomies (Frake 1964, Metzger and Williams 1966), componential analyses (Goodenough 1956), folk definitions (Keen 1985, Manes 1976), and decision models (C. Gladwin 1976, 1980; H. Gladwin and Murtagh 1980; Quinn 1978). (See also Casagrande and Hale 1967, Goodenough 1956, Lounsbury 1956, Wallace and Atkins 1960, Schensul 1969, Werner, Schoepfle, et al. 1985–86.)

The strength of unstructured interviewing is that informants have great freedom to express themselves using their own cultural constructs independently of the presuppositions of the ethnographer. In other words, its strength is the trustworthiness or validity of the data collected. However, this trustworthiness is not guaranteed; it depends on a number of factors, including the motives and capabilities of the key informant and the degree of openness in the interaction between the researcher and the persons he or she interviews. Some ethnographers have worked intensively over relatively long periods of time with certain informants. Majnep and Bulmer's (1977) work is an example from ethnobiology of close collaboration that resulted in a joint publication by the ethnographer and the informant. Berlin (1984) has described the extensive contributions of his Mayan and Jivaro informants. Long-term, intensive collaboration with individual key informants is particularly likely in linguistics and ethnoscience research (see Hale 1969, Werner, Begishe, and Austin-Garrison n.d., Bernard and Salinas 1975, Salinas and Bernard 1978, n.d.).

There are at least three main roles that key informants assume:

The key informant can serve as tutor or guide, answering grand-tour questions, identifying topics of potential concern, helping the ethnographer discover what questions are important, and aiding in the design of standardized interviews for systematic survey of the rest of the community. Unfortunately, very little such work has been reported (but see Werner, Schoepfle, et al. 1985–86), although we suspect that many anthropologists have actually built survey instruments in close collaboration with native informants but have not measured the effects of these efforts on producing effective questionnaires. This is tedious and difficult methodological work, requiring comparisons and controls, but it is vital research if we are to understand the relative importance of the insider versus the outsider approach in questionnaire design.

Alternatively, the key informant can act directly as a data gatherer either by interviewing other members of the society or by recording his or her own insights. There is a minor tradition in anthropology of working with native people who are actually writing the ethnography of their own people in collaboration with an anthropologist (for a report of Boas's work with George Hunt, see Rohner 1966; for modern examples, see Hale 1969; Werner, Begishe, and Austin-Garrison n.d.; Salinas and Bernard 1978, n.d.).

Finally, the key informant or native member of a research team can serve as an interpreter and analyst of data collected and as a judge of conclusions drawn by the ethnographer (see Werner, Manning, and Begishe 1983).

The principal weaknesses of unstructured interviewing are that the data are usually not comparable across informants and are incomplete because systematic coverage of all topics is impossible. (D'Andrade [1970] has suggested, however, that if three informants independently agree on a point of information, chances are that all informants within a social system will agree.) It is this weakness, of course, that makes survey research attractive. Increasing the reliability of ethnography and the validity of survey research through use of the two techniques to inform one another is a high priority in the social sciences. As noted above, a great deal of research is needed to determine how much ethnography increases the validity of survey research and how much surveys increase the reliability of ethnography. (Claims that they do so are attractive but insufficient.)

Ethnographic description is a multistage process. The observed verbal and physical behavior of a community is reduced to a collection of records which are then analyzed, and the conclusions of that analysis are presented in an ethnographic report. There are three types of error that can be introduced at each stage in this process of information reduction and transformation: random error on the part of both informants and investigators (as noted above, this can sometimes be handled with statistical techniques), bias, or systematic distortion, on the part of both informants and investigators (for example, ethnocentric bias of either the informant or the ethnographer or both), and misrepresentation through the reduction of complexity on the part of investigators. In order to reduce the magnitude of error in ethnographic reporting and to estimate the amount of residual error, we recommend that anthropologists:

1. Improve procedures of selecting key informants, choosing those who are knowledgeable, accessible, and communicative. As a first step we recommend getting informants' assessments of who might be a group's significant actors or subgroups; a further step would be a systematic census of all persons present on the ethnographic site, when possible. Recent work by Romney, Weller, and Batchelder (n.d.) describes a procedure for selecting knowledgeable informants.

2. Use mechanical means, such as tape recorders and video, whenever possible, to make verbatim records of the informant's words (on film recording, see Blacking 1984, Collier 1967, Mead 1970; on audio recording, see Werner 1961, Werner, Schoepfle, et al. 1985–86; on the use of videotape, see Dehavenon 1978, Sharff 1979, Reiss 1983).

3. Accurately record the circumstances of the collection of data and report both the research procedures used and the social context. Whenever possible, and without compromising informant confidentiality, anthropologists should adopt the historian's practice of citing sources.

4. Improve data management techniques, preferably with field computers. Efficient procedures for systematic manage-

ment and reduction of complexity are particularly important when dealing with verbatim recording of informants' reports in their native language. Much can be done by adapting already existing data base management systems to the needs of fieldwork. A data base management system should provide adequate storage and efficient retrieval of information, procedures to ensure completeness of data collection and data analysis, procedures to search for contradictions and variation in the data, and procedures to facilitate the systematic reduction of complexity (see Bernard and Evans 1983, Podolefsky and McCarthy 1983, Sailer 1984, Werner 1983).

5. Cross-check conclusions drawn from unstructured interviews with direct observation, structured interviews, and archival materials in order to establish validity. Any success at resolving (rather than adjudicating) discrepant accounts improves the ethnographic description.

6. Make public primary-data records, preferably in the form of the data base, without compromising the confidentiality of informants. The archiving of these primary records would allow second opinions on the interpretation of data, as well as replication of research. It would also improve studies of social change and provide for systematic exploration of the strengths and limitations of alternative data collection procedures.

7. Improve language skills, cross-cultural understanding, and *self*-understanding (see Naroll 1962 on the value of language skills; see Agar 1980:42 for discussion of ethnographers who have gone through analysis in order to increase their self-understanding; see also Goward 1984, Powdermaker 1966, Tonkin 1984).

Our goal is to improve data quality while reducing the cost of its collection. Research is needed to:

1. Discover the best data base management systems for handling anthropological field notes.

2. Find ways of increasing the speed and reducing the difficulty of transcribing verbal data through the use of Dvorak-like keyboards.

3. Improve methods of selecting key informants and of informant sampling. We suggest that fieldworkers be polled for their heuristics for estimating the sample size necessary to establish various kinds of data on social stratification, division of labor in society, and other complex topics (see Poggie 1972 for a consideration of the issues).

4. Test the validity and reliability of alternative data collection procedures (such as participant observation, direct observation, questionnaire surveying) in order to discover which methods are most robust in what contexts.

5. Reveal the consequences for data quality, by topic, of the use of interpreters and of improving the efficiency of language learning by ethnographers. Some topics will almost certainly require less language fluency than others.

6. Determine the implications for the conduct of interviews of research on small-group interaction, including the effects of informant fatigue and stress on recall, body-language cueing, etc. (see Bower, Monteriot, and Gilligan 1978, Bower and Hilgard 1981, Fischoff 1983, Fisk and Schneider 1984, Gregg 1972, Nelson, Bajo, and Casanueva 1985).

7. Develop extralinguistic aids to unstructured interviewing such as the use as prompts of tape recordings, videos, photos, and movies of other informants.

8. Identify ways of manipulating the context of the interview by acting out cultural scripts or otherwise replicating culturally appropriate contexts of information transfer.

Structured Interviewing

Structured interviewing is any eliciting technique that systematically asks the same question of every individual in a sample of informants. The number of techniques is great, so we will outline only some of the better-known. Anthropologists have tended to pay little attention to the problems of interviewing in general and to the use of questionnaires and other systematic data collection techniques in particular. Fortunately, other social scientists have already produced an enormous literature on these topics, including Coombs's (1964) seminal work, *A Theory of Data;* test and measurement theory as well as psychometrics, beginning with Lord and Novik (1968) and including the work of Nunnally (1978); the work of Lazarsfeld and Henry (1968) on latent structure analysis; the literature on response effects in survey research, especially work by Sudman and Bradburn (1974, 1982); and, finally, the variety of models available in multidimensional scaling as reviewed by Carroll and Arabie (1980). Interview formats include structured domains, "pick"-type formats (free listing, nominal-data generators, ratings, triads and balanced incomplete block designs), "order"-type formats (rank orders, paired comparisons, modified rank orders), and other useful formats such as pile sorts, construction of tree graphs, other graphs, and successive sorts, and fill-in-the-blank.

Structured domains. Structured-domain interviewing involves questioning a group of informants, who have been more or less systematically chosen, about the same set of topics. It may be considered a bridge between unstructured and structured interviewing. The questions in this form of interview are generally open-ended, but the range of topics covered is preset through prior work by oneself or others. The topics and their subdivisions are determined by conventionalized categories (those current in the cross-cultural study of the particular domain under investigation) or by the culturally specific categories into which this domain is currently divided in the society in which the work is being carried out. Work with key informants helps to elicit the culturally specific categories, but structured-domain investigation is carried out on larger samples to establish a range of variation and of agreement about a cultural domain.

This kind of interviewing is common in the gathering of basic census materials, dietary information, reports of illness episodes, data on agricultural and other economic activities and household composition, and all the other types of data for which systematic protocols are administered in samples of households. In recent years it has become more common to establish basic descriptive information about individual communities through such structured-domain interviewing. Despite variations of content from one culture to another, basic household-level descriptions about economic activities, family composition, social activities, and so on, are obtained using interview schedules that require only moderate adjustments from one field site to another. Structured interviewing on specialized topics and the eliciting of specialized cultural domains as cognitive information are sometimes added onto the "routine" household interview (see Johnson 1978, Pelto and Pelto 1978; for a review of household interviewing, see Schmink 1984).

Interviewing in terms of structured domains may serve two quite different purposes: to identify *variations* in economic, social, and other behaviors in a population and to delineate "norms" or other areas of consensus (or near-consensus) in aspects of culture. It has most of the advantages of unstructured interviewing but adds a systematic dimension that is often lacking in participant observation or in completely open-ended work. It does not, however, permit informants and interviewers to roam freely across topics. In general, free, open-ended interviewing should be done before establishing a protocol for structured-domain investigation.

"Pick"-type formats. Free listing has several useful applications in anthropology. Perhaps the most important is to delineate the boundaries of a semantic or cultural domain and to ensure that one is dealing with culturally relevant items. For example, Weller (1984*a, b*) used free listing for this purpose in her comparative study of disease concepts. Free listing has other uses as well. In a study of English kinship terminology,

Romney and D'Andrade (1964) asked their subjects to "list all the names for kinds of relatives and family members you can think of in English." They were able to learn a great deal about the cognitive structure of kin terms by examining the order and frequency of recall of the terms (see also Sanday 1968). Formats of this kind ensure that the concepts are culturally relevant and can provide insights into cognition through frequency and order properties. It is sometimes difficult, however, to find an appropriate superordinate term to start the process. (An appropriate superordinate term may be identified through unstructured interviewing.) Moreover, the elicited terms may be too sparse for analysis, and an appropriate statistical model is often difficult to specify.

Nominal-data generators are of a variety of kinds, perhaps the most common of which is the checklist used for establishing material inventories or for developing composite variables such as an acculturation index. Often, checklists are part of the structured household interviews mentioned above, though their simplicity makes them useful in other contexts as well.

Not as common in anthropological work are true-false and multiple-choice questions. The format that is perhaps the most useful for true-false questions is frame elicitation in ethnoscience research. A typical example is the D'Andrade et al. (1972) study of disease terms. J. Young's (1980) study of diseases and health-care decision making in a Mexican community, in which respondents are asked whether or not each of a series of diseases is characterized by a given symptom or quality, is also a good illustration of the technique. Others who have applied the technique are Furbee and Benfer (1983), Weller (1984*a, b*), and Garro (1983). Where true-false questions are appropriate the frame format is superior to unrelated questions, giving the investigator tighter definition of the relevant domain and better criteria for the relevance of questions.

When there are a limited number of possible answers to a question, the multiple-choice format is appropriate. For example, Sankoff (1971) used this format when he asked people which of six clans a given piece of land belonged to. Multiple-choice questions are often used in aptitude tests, since they produce more information than true-false questions and therefore fewer are needed to reach reasonable reliability. However, it requires a great deal of cultural knowledge to write alternatives that are equally attractive.

Nominal-data generators are typically easy to construct and to administer to a large number of informants, and the data are easily coded and analyzed. Furthermore, frame formats help ensure that questions are comparable across a set of concepts. Test and measurement research has developed good models for analysis of true-false and multiple-choice data. Since true-false and multiple-choice questions are independent of each other, the data lend themselves to analysis of both informants and stimuli. Frame questions can, however, be tedious for both subjects and investigators. They require that informants fill in blanks for combinations of stimuli in sentences. The practical limit appears to be 25 concepts and 25 characteristics, involving 625 different questions to which an informant must respond. The technique requires good definition of both concepts and characteristics and cannot be applied in very many domains. One needs a thorough ethnographic grounding in any culture before selecting a set of terms on which to conduct such tests or, for that matter, to write an unbiased multiple-choice questionnaire.

Ratings are one of the more common ways of collecting data in the social sciences. Usually a simple scale of 1 to 7 or 1 to 9 is used. Some researchers believe that they can improve the measurement level of ratings by helping the informant with anchor points or by adding an apparent metric. For example, in studying the cultural convergence of Korean immigrants in Hawaii, Kincaid, Yum, and Woelfel (1983) used a task they describe as follows: "The 11 concepts (such as success, happiness, me, divorce, children, saving face, etc.) required 55 paired-comparison judgments per respondent, according to the following instructions: 'if Love and Hate are 100 units apart, how far apart are _____ and _____ ?' Respondents were instructed to keep this standard measure in mind as a guide for making direct magnitude (ratio) estimates of the distances among the 11 value concepts." This method may or may not be an improvement over a nine-point scale; in any event, there is no guarantee that the data can be considered to be "meaningfully" measured at the ratio level. The work of F. S. Roberts (1979) on measurement theory demonstrates that certain uses of ratings are "meaningless." Anthropologists have used modifications of Osgood's semantic differential procedures (Osgood, Suci, and Tannenbaum 1957) for obtaining ratings of foods, diseases, kinship, community characteristics, and many other domains (see, for example, Romney and D'Andrade 1964). Rating scales are very easy to use, but they have serious problems from the viewpoint of measurement theory, including anchor-point comparability both within and between subjects.

Triads and balanced incomplete block designs (BIBDs) have been used extensively in anthropology, psychology, and other social science fields. The method involves giving informants groups of three stimuli (e.g., three kin terms, the names of three persons, three actual plants) and asking them, for example, to choose two of the three that are most alike. One of the early uses was by Romney and D'Andrade (1964) on English kin terms. Burton and Nerlove (1976) developed BIBDs for the triads test that greatly increased the number of concepts that can be practically explored with that test. The complete method becomes impractical with around 12 or 13 concepts (13 concepts require 286 questions), while the balanced block design allows up to 25 concepts (a lambda-one design of 25 concepts requires 100 questions). Burton and Nerlove (1976) illustrated their methods on the domains of vegetables and kinship. Many other domains have been studied using the triads method, including animal terms (Henley 1969) and occupations (Burton 1968, 1972). Kirk and Burton (1977) studied 240 Maasai informants to investigate variations in the meaning of personality-trait descriptions across changes in the social-identity terms (e.g., "older boy," "warrior") that the trait descriptors modify. Lieberman and Dressier (1977) used the triad-sort method to examine differentials in grouping of common diseases among St. Lucians. Triad tests are good for small numbers of concepts, and many informants find them fun to do when the number of concepts is small. They can be administered easily to nonliterates, and they do not predetermine answers. Very good models exist for the analysis of triad data, including some from Coombs. Their major disadvantage is that they cannot be used for a large number of concepts.

"Order"-type formats. Some domains lend themselves to being ranked by informants on one or more dimensions. For example, the ranking of occupations in terms of prestige or income is a natural and easy task for many informants. In these favorable situations, the rank-order method produces rich information and would be the method of choice. Hammel has used it in both Yugoslavia (1970) and Peru (1962). A recent comparative study by Burton, Ferreira-Pinto, and Magana (1985) employed it in Guatemala and Mexico. Coxon and Jones (1978, 1979*a, b*) have made extensive studies using a number of methods, including ranking, on the perception of occupational cognition. Silverman (1966) used the method to obtain social rankings in an Indian community and DeWalt (1979) to examine socioeconomic ranking in a Mexican rural village. Rank ordering produces a large amount of information per unit of informant time and lends itself to individual-difference and intracultural-variability studies. However, it is difficult to use with nonliterate subjects (despite recent successes using pictures and other concrete stimuli), and there are limits on the number of objects that can be ranked. Successive rankings of the same objects on different variables are not independent.

Paired comparisons may be thought of as ranking each pair

of objects on some variable, and triples, quadruples, etc., can be treated in the same way (see BIBDs above). Weller (1984*b*) has pointed out that these designs are easier to use than rank orders in oral interviews with nonliterates. Nerlove and Walters (1977) have presented a way to combine information from different informants when informants have ranked different subsets of objects. Paired comparisons can be used with nonliterate informants; block designs allow a larger number of objects to be ranked, and within-informant consistency can be measured. However, not all kinds of data lend themselves to ranking, successive rankings are probably not independent, and there are some practical limits to the number of objects that can be ranked.

Other useful formats. Pile sorts have been used extensively in anthropology (see Werner and Fenton 1970). Perchonock and Werner (1968) used it for taxonomic elicitation of foods. Roberts and collaborators have studied eight-ball pool (Roberts and Chick 1979), pilot error (Roberts, Golder, and Chick 1980), women's trapshooting (Roberts and Nuttrass 1980), and tennis (Roberts et al. 1981). These investigators have perfected an ethnographic method that uncovers the structure of relevant behavior events in a domain through use of the pile sort. Each domain is coded into a "high-concordance code" that informants readily understand. The method has been used to collect data on concepts of success and failure in the United States (Romney et al. 1979) and in Guatemala (Freeman et al. 1981). In a study of 68 role terms, Burton and Romney (1975) showed that by rescaling subsets of roles (e.g., occupations or kin roles) they could better reveal the fine grain in the data. A study on occupation names by Burton (1972) also used the pile sort.

Burton (1975) has worked out a valuable standardization procedure for the pile sort that makes adjustments for different numbers of piles among informants and different numbers of concepts in each pile. In a pair of articles, Boorman and Arabie (1972; Arabie and Boorman 1973) derived a metric for comparing different pile sorts, reporting that the difference between "lumpers" and "splitters" overwhelms all others. The implication is that, unless one's interest is limited to "lumpers" and "splitters," the pile-sort technique is not well suited to the comparison of individuals.

Pile sorts can be used to compare a large number of concepts, and the analysis can be automated with the aid of computer cards. On the other hand, they cannot normally be used with nonliterates (but see White 1978), they cannot be used to compare individuals, and, despite work of Boorman and Arabie (1972), there is no general model for pile-sort data. (New developments in computer graphics will make on-screen sorts of pictures possible and will make pile sorts more appropriate with nonliterates.)

A number of variations on the pile sort have been used by a variety of researchers. Successive partitioning of objects automatically results in a taxonomy of the objects. Boster (n.d.), for example, used this method where the stimuli were prepared bird specimens. Fillenbaum and Rapaport (1971) studied the structure of several semantic domains including color terms, kinship terms, pronouns, emotion names, prepositions, conjunctions, "have" verbs, verbs of judging, and good-bad terms. They had subjects draw trees by giving them a list of terms in alphabetical order, letting them read the list several times, and then asking them to pick the two words that were most similar, write them on a blank sheet of paper, and connect them with a line labeled "1." Next they were to pick the next most similar pair of words (which could include one already used), add it to the blank sheet, and label the line "2," and so on, continuing until all the words were connected. Werner, Manning, and Begishe (1983) report a tree-drawing exercise with uneducated informants. They showed that taxonomies could be built upward, starting from the most specific of plants in Navajo; they also showed that all conceivable taxonomic variations in informant disagreements actually occur in a relatively limited domain of plants and animals in Navajo. People seem to catch on very quickly to tree-graph representation of semantic domains. The method produces a complete tree for each individual and seems to work well when data are aggregated, but it prejudges the structure of the domain under investigation (that is, it assumes that a strict tree exists). Also, there is no way to analyze (that is, compare statistically) individual differences (see Fillenbaum and Rapaport 1971:25–28). Finally, graph drawing cannot be used easily with uneducated informants and is impractical with more than 30 concepts.

Fill-in-the-blank is an extremely powerful method not often used in anthropology nowadays. All of the early work in ethnoscience was of this type, including Metzger's various uses of question-and-answer frames. Tyler (1969) presents a variety of uses of fill-in-the-blank studies. More recently, Boster (1985) has used it to study Aguaruna manioc classification. He concluded that the more an informant agreed with others the more knowledge that informant had about manioc. Since he was able to assess differences in knowledge among informants, he was able to establish that women knew more than men and that women in the same kin and residential group were more similar in knowledge than unrelated women. Systematic use of this method has led Boster to findings bearing on some basic questions about intracultural variability and culture theory in general. The cultural consensus model (Romney et al. 1985) applies to this kind of data and in favorable cases allows some very precise inferences to be drawn. Fill-in-the-blank formats have the advantages of providing a model for the analysis of data, producing information on both informants and cultural concepts, and yielding several times as much information as do true-false questions.

One basic problem with structured-interview formats involves the selection of a measurement model for the analysis of a particular domain. For example, some domains, such as illness terms, seem to be well specified by verbalizable cultural attributes, permitting the use of frames and true-false judgments. Other domains, such as color terms, seem to involve less verbalizable properties, and models involving simple naming of color chips and similarity judgments on color chips seem more appropriate. At present there is no clear theory about how to choose the right "tool" for particular "materials." The justification for selection of one or another kind of model seems best handled by a kind of trial-and-error process, in which the aptness of the model is judged by the reliability and validity of the measurements derived.

Other issues, such as the appropriateness of structured interviewing for collecting information about past actions or experiences of the informant, are even harder to formulate. While ideally one would prefer observational data about behavioral events, often memory-based measures are the only feasible means of data collection. However, an extensive body of research has demonstrated that certain kinds of memory-based measures are badly biased (see Bernard et al. 1984). A general recommendation would be that other memory-based measures for behavioral events be used and that at least some behavioral observation be undertaken to determine the validity of informant reports. Theoretically, informants and ethnographers should show great agreement in their immediate recording of observations and great diversity in their later recall of the same observations. This needs to be demonstrated experimentally, however.

Finally, it should be stressed that, even though working out an appropriate measurement model is difficult, if no coherent model can be stated the results are unlikely to be acceptably reliable and valid. Some of the problems which have emerged in trying to measure mental illness using the Rorschach, TAT, and MMPI should serve as a caution: that a technique will generate a set of numbers does not mean that it is a good one—

the numbers generated must have a sound relation to something in the world.

Direct Observation

The ideal instrument for social science is an unobtrusive, omniscient observer who describes without omission or distortion all the environmental conditions of a particular field site, all the behavior of the people there, and all their utterances. Of course, this is not possible (and profound questions of ethics would arise if it were), but much more direct observation is possible than has been previously thought. The most frequently cited drawback to direct observation is that it is labor-intensive and hence an expensive research method (Whiting and Whiting 1970:284). Another criticism is that the physical presence of the observer is intrusive (Webb et al. 1966). Thus, less direct methods of describing natural behavior are more common in actual use. "Indirect" methods include experiments of various kinds, techniques that use the research subject as observer of his or her own behavior via questionnaires, interviews, life histories, tape recordings by nonliterates, and diaries, and descriptions based on field notes written up by the observer some time after the observed event. We would expect these last to be more ethnocentrically biased than direct observation (Shweder and D'Andrade 1980).

Experimental observations are "indirect" because it is necessary to infer that behavior in an experimental setting accurately reflects some natural, real-world behavior, and any description based on "long-term" memory (that is, where the description is made after some seconds or, at most, a minute or two have elapsed) is of questionable validity. People are frequently unreliable in describing their own behavior (Bernard et al. 1984), most likely because long-term memory distorts observations in the direction of preexisting cognitive patterns (D'Andrade 1974). We need a great deal more research to find out what it is that people are actually describing when we ask them to recall events. For example, Weisner, Gallimore, and Tharp (1982) reported that when Hawaiian children under observation were asked whether or not they were caring for other children, boys underreported their caretaking (in comparison with the observer's judgment), whereas girls overreported. The distortion is in the predicted direction: Hawaiian boys do not consider caretaking to be appropriate to their role, whereas girls do. There was, however, some concordance between observer's and subject's descriptions, varying from about 50% to about 80%. Furthermore, we could ask in this case who is the more accurate: the observer who sees no activity and codes the child as "not cared for" or the subject who feels responsible and is "keeping an eye out" for the child and hence describes himself or herself as a caretaker? The answer depends on how "caretaking" is operationalized by the researcher.

Yet, some informants are more accurate at describing their own behaviors than the behaviors of others (Romney and Weller 1984), and if these can be identified, their self-reports can constitute a reliable subsample of a population; it is also the case that careful questioning can elicit more accurate descriptions (McSweeney 1979). On the other hand, the evidence is clear that informants are not the "videotape-like creatures with near-perfect retrieval systems" that those who use indirect methods might assume (D'Andrade 1974:124). It is ironic that researchers who spend such heroic efforts to present questionnaires and establish samples to reduce error to a tiny fraction of the responses overlook informant long-term memory distortion, which can sometimes cause informant reports to bear no statistical relation at all to their directly observed behavior (Bernard et al. 1984).

Direct observation by anthropologists is common, but usually in the form of loosely structured, opportunistic note-taking by individual fieldworkers (Whiting and Whiting 1970:282–86). Resistance to more systematic observation may reflect practical concerns such as cost but may also reflect a fear of "dehumanizing" the research. In the category of "dehumanized accounts," however, should be included accounts unfaithful to the lives of the people we study. Accurate descriptions of behavior should be considered an essential part of anthropological fieldwork, and therefore we need a toolkit of observational methods.

Participant observation in natural settings probably produces the most valid descriptions of human behavior, usually at some cost of reliability (Pelto and Pelto 1978:33–34). Here we may identify six inevitably somewhat arbitrary categories of observational methods: direct measurement of the environment, nonreactive measures, time allocation studies, time frame studies, structured observations, and participant observation.

Direct measurement of the environment. It is extremely common for anthropologists to measure natural features of the environment such as rainfall, soils, game and useful plants, and material cultural features such as houses (including interior layout and furnishings), shrines, equipment, and food supplies. Although actual measurements may sometimes be difficult to carry out, standardized procedures for measuring environmental features exist and in many cases require relatively little expertise. Yet our training of graduate students in such procedures is casual, at best. We know of no systematic compendia that describe the basic measurement techniques for such relatively straightforward tasks as sampling soils in an agricultural field, describing the tools used by craftsmen, taking rainfall measures, or mapping a community. What levels of accuracy are needed? Which details are essential and which peripheral? Standard works in other fields, such as nutrition (see Jelliffe 1966) or human biology (see Cameron 1984), suggest rich possibilities in this area.

Nonreactive measures. Nonreactive measures include "unobtrusive measures" (Webb et al. 1966) for observing the residual effects or consequences of the behavior of individuals, noninvasive observation (for example, aerial photography), and other, indirect measures that may be very intrusive but can still claim to give accurate indicators of behavior. These measures are limited only by the imagination of the investigator. Simply being present when food is cooked gives an observer the opportunity to collect accurate data on how much food is used by a particular household to prepare meals, provided that one has access to the utensils, such as gourds and spoons, that are used for measuring quantities. Even counting a sample of handfuls (as with grain) can provide much better data than we now have for most of the world on food preparation and consumption.

A well-known example of indirect, nonreactive food measurement is the program of "refuse analysis" conducted by Rathje and his associates (see Rathje and Ritenbaugh 1984). Consisting of the analysis of the refuse of living people, the technique has been called "an archeology of us" (Gould and Schiffer 1981). People who are observed eating or who are asked to write down everything they eat generally change their eating patterns. Yet careful analysis has shown that people do not change their refuse discard patterns when they know their garbage is being studied (the single exception being whiskey bottles) (Rathje and Ritenbaugh 1984).

An intrusive measure that is probably nonreactive is the measurement of obesity, for instance, by weighing subjects or measuring skinfolds. Although such measures make obvious to subjects what aspect of their behavior is being observed, it is unlikely that even a series of such observations over time will alter their overall patterns of food consumption (except when weighing is done in conjunction with a weight control program!).

Time allocation studies. Johnson (1975) and Gross (1984) provide an overall framework for the application of time allocation studies (previously the province of organizational man-

agement) to the field settings in which anthropologists generally work. The goal of time allocation studies should be a representative sample of all the behavior of all the members of the community being studied. Obviously, if the community is of any size, no research project will have the manpower to make continuous direct observations. Also, the research subjects will choose not to be observed at certain times, as is their right; hence indirect methods will be necessary.

The usual technique is to make spot checks (instantaneous scan sampling) of randomly chosen individuals at randomly chosen times, each observation being like a snapshot or "instantaneous slice" of the stream of behavior. Large numbers of such random observations make up a picture of average time allocation, in the same way that a newspaper reproduction of a photograph, made up of thousands of dots, can represent a face. Such descriptions, based on the researcher's direct observation of an activity at the moment it occurs, make it possible to describe the division of labor by sex, age, and social or economic status; patterns of social interaction as seen in visiting, gift or meal exchange, and cooperation in tasks; and how particular individuals divide their time among alternative activities open to them (see Rogoff 1978, Topper 1972).

Among the most important limitations of such time allocation methods are that they do not describe complex sequences of activities because they are based on random spot checks of behavior rather than continuous periods of observation; that they usually give only sketchy descriptions of single individuals (since the number of observations per individual tends to be fairly small); and that they do not distinguish activities that are the focus of the research from activities that are peripheral (the methods being specifically designed so as to not bias observations toward particular kinds of activities or settings).

Time frame studies. The usual way of circumventing the limitations of time allocation methods is direct observation of the activities of central interest to the research. As in studies of animal behavior, the research focuses on a specific individual for a specific period of time and describes that individual's activities in terms of the central concerns of the research. For example, the 5- to 30-minute observations widely used in cross-cultural research on child care and socialization have produced quantitative descriptions that have allowed detailed cross-cultural comparisons (e.g., Whiting and Edwards 1973). The unit of time need not, of course, be as restricted as 5 or 30 minutes: it may be any bounded time frame. Common units include a day or 24-hour period and a week. Conceivably, a year could be such a unit, though it is impractical to describe an individual's activities continuously over the course of a year either by direct observation or by informant recall, and thus the sampling procedures of time allocation studies become increasingly relevant the longer the time frame being studied. Another kind of time frame is the "event," such as a marriage, a public meeting, or a sports contest, which is generally bounded though variable in length. Technical problems are apt to arise concerning how a limited number of observers can fully describe the activities of all participants in a complex event, and therefore sampling is likely to be needed.

A multitude of specific measurement issues arise here: How detailed must our description of the subject's activities be? Ethologists sometimes break behavior down to bits of muscle movement (Blurton-Jones 1972, Harris 1964). Different research goals will imply collecting data of varying detail, depending on the level of analysis for which they are intended. How do we know when one activity has ended and another has begun? The problem of dividing the continuous stream of behavior into discrete "behaviors" has plagued observational research from the outset (Barker and White 1951, Chapple and Arensberg 1940, Harris 1964). Interestingly, informants seem to be much less equivocal than outside observers about the boundaries of their behavior (Werner 1966). How do we describe behavior that is simultaneously two or more "behaviors" at once? For example, a nursing mother who is preparing food or a man who is teaching his son mechanics while repairing the car is doing two things at once, but including both activities in the description leads to either an "inflated *N*" or cumbersome multiple or composite codes (Gross 1984). Perhaps something akin to musical notation for multiple, simultaneous activities needs to be developed for field observation.

The main value of focused observations is that they allow the description of behavior in context: for example, weeding a field, making a pot, or nursing a baby is not just an "activity" but a complex of constituent acts, with definite beginnings, middles, and endings that must be described as such. Furthermore, making many such "time-motion" descriptions of such activities over the course of fieldwork allows for quantitative descriptions of patterns, such as "average area weeded per hour of labor" or "average delay of response to baby's crying."

Structured observations. Structured observations are common in psychology and sociology but relatively underutilized in anthropology; it is possible for the researcher to construct or take advantage of existing standardized settings in which to observe behavior (see Aronson 1980, Ciminero, Calhoun, and Adams 1977, Kent and Foster 1977, Lindzey and Aronson 1968, Milgram 1977, Milgram and Shotland 1973). Examples are observing play areas supplied with toys (e.g., Henry and Henry 1944), having a meal with the whole family together (Johnson and Bond 1974), and placing videocameras at strategic locations in residences or workplaces (O. Johnson 1980, Dehavenon 1978, Reiss 1983). Although research subjects in such settings are relatively free and do not act out a script written by the observer, the standardized features of the setting help control variation to some degree, and a researcher gains experience through repeated use of such settings in standardizing the way in which behavior is recorded (for example, by keeping track of which individuals use which toys or of who serves food to which family members).

Participant observation. The least structured and systematic of direct observation methods—jotting down anything that strikes the observer as important in a pocket-size field notebook or dictating data into a pocket-size tape recorder—has greater possibilities for quantification than is usually recognized. There is nothing so "anthropological" as taking observational notes in the field, and there is no greater opportunity than in this technique for anthropologists to improve the quality of their data. Researchers can develop work habits that well enable them to extract data in the field for quantitative manipulation once fieldwork is over. For example, fieldworkers can make efforts to move widely throughout their research community, even into areas where they feel shy and unwelcome, in order to make their field notes cover a representative sample of community settings. Time allocation studies try to achieve representativeness with rigorous sampling; researchers doing time allocation studies find themselves in a representative sample of "scenes" throughout the course of their research, and their field notes reflect this comprehensiveness. Furthermore, researchers can be more self-conscious in describing in the fullest possible detail those aspects of behavior that are central to the research, so that patterns can be demonstrated quantitatively during later analysis. Because the observational task is so labor-intensive, this aspect of research must be theory-driven. One cannot simply observe everything and hope that patterns will emerge. Since formal, mathematical theories are generally not available in anthropology to tell us what to observe, the next best thing is to isolate a domain of behavior in which we are interested and then observe as much behavior as possible that *may* have something to do with that domain. For example, if political influence is a central focus of research, then everything relevant to political influence, such as hosting of meals, giving of gifts, garnering and achieving signs of support in public meetings, etc., should be recorded as to time, place, personnel, goods/services exchanged, etc. Later, it will be pos-

sible to tabulate these observations and assign numerical values such as frequency of interaction, tendency to give more than receive (e.g., Henry 1951), and so forth.

In the abstract, the ideal in direct-observation research is an omniscient observer who describes all behavior and from whom informants do not hide. Of course, achieving this ideal is impractical (some would say it would be unethical, too). Nevertheless, the costs and benefits of compromises should be carefully weighed.

The reliability of observational data must be assessed. Many behavioral observations are done by lone fieldworkers, and no reliability checks are available. Field assistants often do not understand the need for them and resist tests for intercoder reliability (Munroe and Munroe 1969). Reliability is perhaps the least-stressed scientific principle in our training of graduate students.

Closely related is the standardization of measurement cross-culturally, especially important when different researchers in different research projects are collecting similar data. Standardization would be enhanced by sharing raw data more widely, making available an explicit codebook describing procedures for coding raw data into published form, and publishing descriptions of how primary observations were made. (Journal editors frequently excise such descriptions in order to conserve space; authors may help solve this problem by making methodological appendices available by mail to interested readers or, better still, by depositing their raw data, along with descriptions of how they were collected, into archives such as the Human Relations Area Files.)

Because the short-term memories of observers (and informants) are more reliable (and possibly more accurate) than long-term memories (D'Andrade 1974), field methods should stress immediate recording of observed behavior.

Sampling procedures should be explicit and, ideally, based on random selection of the units of analysis so that data can be generalized to the community or population sampled.

Anthropologists have been reluctant to standardize field observations, seemingly without giving the matter much thought. It is one thing to recognize that each naturalistic research setting has unique characteristics but another simply to discard the research of others and reinvent the wheel in each new research project. As in most systems of measurement, direct observation is a skill that can be taught (Boice 1983, Levine et al. 1980).

Use of Archival/Documentary Data

Archival/documentary data include all materials collected or written up by persons other than the investigator. These may be ethnographic materials or documentary sources such as census records, household registers, and diaries. Ethnographic resources include photographic and film collections (Bateson and Mead 1942, Sorenson 1975) and the Human Relations Area Files, the largest and most widely used archival resource in our discipline.

The Human Relations Area Files. Cross-cultural or holocultural researchers are the main users of the Human Relations Area Files (HRAF); the files provide a growing worldwide sample (now including more than 300 societies) on which hypotheses can be tested. The major assumption in cross-cultural research is that if a theory has merit, the presumed causes should be associated with their presumed effects (Naroll, Michick, and Naroll 1976, Otterbein 1969, Rohner 1977, Whiting 1961). Most cross-cultural researchers also assume that the ethnographic record is probably largely accurate and therefore, if there is lawfulness with regard to cultural variation, we should generally be able to find it. This belief is bolstered by the enormous number of significant correlations we now have in the cross-cultural literature (for a review, see Levinson and Malone 1980). A third assumption is that most of the errors in ethnographic data are random errors; since random errors tend to *lower* correlations, most of the statistically significant correlations found so far are probably even higher.

All cross-culturalists agree that there are errors in this kind of research, but there seem to be two different approaches to error. One set of researchers considers errors unlikely to be threats to validity (see Murdock 1977 for this point of view); another set worries more about data and spends considerable time and effort testing for possible errors, particularly systematic errors that could lead to false rejection of the null hypothesis (this point of view is represented by Naroll 1962, 1970, 1977). Both styles of research are based, of course, on assumptions rather than facts about error. We have relatively little research documenting the frequency or kinds of error in data.

Possible errors can be classified in terms of their point of origin: informant, ethnographer, investigator, and coder. There can be random or nonrandom (systematic) errors at each point of origin.

Informant and ethnographer errors might seem beyond the control of the cross-cultural researcher, but Naroll (1962) proposed an indirect method, which he called "data quality control," for dealing with them. The basic procedure is to identify factors, such as language proficiency or length of stay in the field, that might produce biases in certain variables. A short stay in the field, for example, would presumably result in underreporting of things like witchcraft. If length of field stay, a data quality control factor, is significantly correlated with both the independent and the dependent variable in a study, the researcher is obliged to control statistically for that factor.

There are some major problems with this strategy. First, that a data quality control factor is correlated with both an independent and a dependent variable does not necessarily mean that the data are biased. For example, Whyte (1978), testing for sex bias, found significant correlations between sex of ethnographer and certain measures of women's status, but several alternative explanations might account for them, including the possibility that female anthropologists selected research sites in which women were thought to be better off. Second, as Naroll (1977) himself noted, it is very expensive for researchers to code for a large number of factors that could but probably do not produce false correlations. Most important, of the large number of studies done by Naroll, his students, and others (see Levinson 1978 for studies employing data quality controls), hardly any have found that a data quality control factor accounts for the correlations that have been discovered (but see Rohner, DeWalt, and Ness 1973 and Divale 1976). Instead of testing for a large number of data quality control variables, we recommend that researchers control for only those variables which they have strong theoretical reasons to believe may artificially inflate correlations between independent and dependent variables.

Investigator error arises in the process of trying to operationalize theoretical constructs. The error here is validity error—the operational measures may not be good measures of what we think we are measuring. For example, most ethnographers have not done (or published) censuses that would be suitable for arriving at a precise estimate of fertility for the groups that they have studied. The cross-cultural researcher often encounters statements like "They have very large families" or "The average family size is 6" and has to make a decision: to stick with operational definitions such as those used by demographers and lose many cases because they cannot be measured, or to opt for cruder ordinal measures and thereby retain a higher proportion of cases. Measures with the highest face validity and the highest precision are desirable, but the most obvious choices may not be possible if the data are not available. C. Ember (1983) used generally cruder data from HRAF and more precise data from Nag (1962) to measure fertility and found virtually no difference in results. We recommend that

cross-cultural researchers construct a set of measures for each variable of interest that range from the more to the less direct and then report the results for each. If the results do not change substantially using measures of different precision, the researcher can use cruder measures with confidence.

The experienced cross-cultural researcher is probably the most accurate coder. For example, the experienced researcher is more likely to be aware that different ethnographers mean different things when they use certain words and is unlikely to take words at their face value. But reliance on the investigator runs counter to the principle of employing a coder who is innocent about the theories being tested. After all, the experienced investigator may be subject to unconscious "cheating," a serious systematic error (Rosenthal 1966, Rosenthal and Jacobson 1968). Many cross-cultural researchers, trying to protect themselves against the charge of unconscious cheating, prefer to use theory-blind coders. There is no easy solution to this problem, but the best strategy may be to use an experienced researcher *and* a theory-blind coder in order to see if there is substantial reliability.

Coders may introduce systematic errors because of their sex, their political ideology, their personality, or their faulty assumptions about different types of societies. For example, the assumption that hunter-gatherer societies do not exhibit much aggression as compared with more "modern," commercialized societies might lead to erroneous results. One suggested solution to these systematic distortions is a kind of "self-analysis," in which researchers and their assistants explore (that is, hypothesize about) biases they might introduce into coding and then test for these biases by means of the usual quality control procedures. Another source of coder error arises when the coder is not attuned to the problem of "time" and "place" focus. Cultural anthropologists often minimize the possibility that culture change can occur in a short span of time. This lack of attention to time and place has spawned controversy about the accuracy of Mead's and Freeman's descriptions of Samoa, for example, but those two investigators may not have been studying the same thing (see M. Ember 1985 for this position and Freeman n.d. for a rebuttal). In the usual cross-cultural study the investigator seeks to evaluate whether the presumed independent and dependent variables are associated. But if coders sometimes rate the independent variable for later time periods and the dependent variable for earlier time periods, then the strength of the association will be reduced, so much perhaps that the association will not prove to be satistically significant. Divale (1975) has shown with a few examples how lack of time and place focus, which is random error, tends to lower correlations.

Despite considerable attention directed to possible biases in cross-cultural research, few researchers have found evidence that systematic distortions account for the correlations that have been found in variables. If anything, it appears that random error is more of an impediment to our understanding than systematic error: deflation of correlations seems more likely than inflation (see Divale 1975, Kang and Kang 1981, Witkowski 1978). This is not to suggest that we should not study error; on the contrary, we should try to measure as directly as possible whatever error we suspect may be present. The study of error in all areas of anthropological data collection and coding is in its infancy. When we have empirical evidence about the effects of various kinds of bias, and when we understand more about how coders actually make their coding decisions, we will be able to make more intelligent choices about cross-cultural procedures.

Other archival/documentary materials. Most anthropologists have made use of archival or documentary materials in field research. The judicious use of nonethnographic archival data can allow researchers to place their communities in a broader context. It also allows study of events too rare to be observed during the short time fieldwork lasts. If archives go back in time, they allow the researcher to chart temporal cycles which could affect observations and provide clues about the effects on culture of powerful forces—the government, the military, migration, or even the weather. Collected by someone else, they can be a check on both the fieldworker's and the informants' biases. Informants may want to hide events whose recollection could divide the community, such as feuds, incest, etc. These archival materials include (1) person-based records collected by agencies and government, including (*a*) census and vital statistics (birth, marriage, and death records, for example), which purport to cover entire populations; (*b*) records dealing with specific subpopulations (medical records, school and military records, wills, deeds, court case records, and so on); and (*c*) bureaucratic action records (for example, letters, reports); (2) household-based records, such as household registers, land holding and tax records, household consumption and expenditure records; (3) specimens of popular culture and art forms, such as films, records, art collections, newspapers, magazines, books, artifact collections, etc. (Alan Lomax's music and dance archive is a highly developed archive of this sort); (4) personal documents of individuals, such as letters, diaries, and personal collections of photographs (these may be of persons in the culture under study or of missionaries, travelers, or administrators); (5) aggregated secondary data, including (*a*) regional or national reports by economists, political scientists, historians, and others referring to but not necessarily confined to the research population and (*b*) location-specific reports such as weather reports, government surveys, reports of religious missions, trading company records, etc.; (6) linguistic descriptions, including dictionaries, grammars, educational materials, etc.; and (7) photographic archives, including films and videotapes.

The many sources of error here can be divided into the same major categories identified in the use of the HRAF materials; errors can be introduced by informants, reporters or recorders of field data, researchers, and other coders. Many of the suggestions we made earlier about ways to deal with these errors apply equally to research using other archival or documentary materials. We confine our comments here to the special problems in the use of these other materials.

Informants do distort some information given to ethnographers, and ethnographers do introduce some of their own biases (see Bernard et al. 1984). Usually we assume that bias is more likely and more frequent in documents written by casual observers or by political officials and agencies than in those written by anthropologists. While it is comforting to assume this, however, good sense dictates that we find out not only whether our assumption is correct but to what extent. Some biases are fairly predictable: white supremacists are likely to describe cultural features of nonwhite populations in negative terms, administrators will suppress signs of disorder, generals will minimize their own casualties and inflate those of their enemies, informants may distort their reports if they will gain from incorrect ones. Some informant bias is not so obvious, however. Much research is needed on the sources of informant bias and on how to control for it. Researchers using archival material need actively to consider potential biases and then, whenever possible, test for them using procedures similar to those suggested by Naroll (1962). Source criticism is basic to the use of all archival materials, and historians have produced a large literature concerning this mode of error analysis (see Barzun and Graff 1970; Bloch 1963; Fischer 1970; Handlin 1979; Handlin et al. 1954:22–23). Anthropologists should make use of the methodological guidelines this literature suggests. With all types of documents, the anthropologist should carefully consider how, why, under what conditions, and by whom the materials were collected. Such source criticism, contained within monographs and articles as part of the scholarly effort, can help to identify possible systematic biases.

In addition to errors of conceptualization and operationalization, errors can be introduced by the investigator when he or

she does not properly attend to potential problems in the use of these other archives. Among these problems are insufficient familiarity with the sources used or the language in which they are written; reliance on easily available sources at the expense of less accessible ones that might correct erroneous impressions; failure to recognize gaps in the record because of wars, floods, fires, or departure without replacement of recording personnel; failure to notice changes in categories or units of analysis in an archive (for example, changes in the rules for census taking over time or in definitions of disease and causes of death); and failure to adjust or calibrate data to reflect demographic and other changes, such as changes in the value of monetary units. Investigators must attend to shifts in the units of analysis of different archives and ascertain whether or not data in the archive accurately represent the population being studied. For example, some records may pertain only to the upper classes, others only to taxpayers, still others only to those who attend school. If individuals or households are the units of analysis, investigators often want to link one archive to another because different types of information are available in different archives. Here the proper identification of individuals in different data bases becomes an arduous but necessary task in order to minimize error.

Cross-cultural researchers have been leaders in systematic efforts to minimize or control for possible sources of error in studies using the HRAF or other collections of ethnographies. Less attention has been paid to data quality control in the use of other archives (although the use of censuses by demographers constitutes a notable exception). In addition to making their source evaluations explicit, anthropologists using archival data need to develop specific procedures for transforming those data into operationally defined variables. If possible, they should employ theory-blind coders and assess intercoder reliability as a check on unconscious cheating. Rather than eliminating all problematic data, archival researchers should try to test for the effects of the suspected imprecisions, just as is recommended for cross-cultural researchers.

Because both comparativist and noncomparativist anthropologists can profit considerably from access to archives that are not usually available, HRAF should be enhanced through the addition of materials such as censuses, time allocation data, demographic surveys of local populations, aerial maps, local climatic data, etc. Many anthropologists have excellent field notes covering a lifetime of effort; they should be encouraged to give or will their notes to archival institutions. New archives, such as data banks of censuses, art materials, children's readers, grammars, and dictionaries, should also be developed to enhance comparative, hypothesis-testing research.

Anthropologists make a unique contribution to the social sciences through participant observation. By its very nature, the technique requires a large investment of time in a small community, one which may, in fact, be only a tiny fraction of the population carrying the culture under study. Assessing the degree to which a particular community, at a particular point in time, is representative of the culture at large is a basic problem of measurement for anthropologists. Very often, the choice of fieldwork site is out of the investigator's control—the result of personal contacts, available transportation, local politics, and so on. This makes it even more important that anthropologists set the communities they study in a broader context. The problem is compounded when, to shield the identity of informants, the name of the community is changed, making it more difficult to link it to other information from documentary sources.

Anthropologists need to develop a check list of basic information that should be collected and published about each community chosen for study so that its representativeness may be assessed. Besides the currently obligatory maps and dialect groupings, we need to have population sizes of all communities of the culture, location of roads, location of government services, lists of local cyclical events such as markets, etc. Controlled comparisons need to be made to show how communities that play different roles in larger regions differ culturally. These comparisons cannot now be made. The choice of a particular community may facilitate or hinder certain kinds of research: while a small community is more easily covered, a large one provides a bigger informant pool and may be less factionalized. These differences need to be studied so that researchers can choose communities suited to their topics and determine the effects of these choices on the data reported. Reanalysis of field notes of the "big" projects of the '60s and '70s, which had teams in several communities, may provide some answers to these questions.

As the communities anthropologists study come to be the subject of written records—and there are few that are not—anthropologists need to develop guides to different types of records and the problems associated with each. The anthropologist is often in a unique position to evaluate these records for the host country and for other academic disciplines that depend on them. The advantages they offer anthropologists are much greater time and space coverage than is usual for us, the opportunity to assess the representativeness of our data more fully, and the possibility of studying processes that reach outside the local communities and shallow time frames we typically examine.

Summary and Conclusions

Data quality issues in anthropology have become urgent because of our rapidly growing involvement in interdisciplinary research (and in applications) with investigators from the health fields, economic development, agricultural development, and so on. This frequently involves complex, multivariate statistical designs in which sociocultural variables are examined in interaction with biomedical, physiological, economic, and technological variables. In a growing number of cases anthropologists are asked to provide cultural information to improve or refine variables designed by other disciplines. Thus, the language of data quality control is part of the common denominator of communication across disciplinary boundaries.

Our most general conclusions, therefore, are that more methodological training is needed, not only for graduate students but for experienced field anthropologists as well, in all four major research techniques and that anthropologists need to take their place, along with other social scientists, in basic research on key methodological issues. We are convinced that anthropology can make unique and useful contributions to the toolkit of social science. In particular, anthropologists need:

1. Exposure to advances in techniques that have been made in other disciplines that improve validity and reliability. The latest findings and methods concerning data quality control in informant interviewing, structured interviewing, direct observation, and the use of archival materials need to be assembled for easy reference. Succinct, practical training manuals for cultural anthropologists that focus on specific aspects of data quality need to be compiled.

2. More detailed knowledge of the error bounds of our instruments (for example, the questions we ask informants). We need to know when data must be more precise and when they do not have to be so precise. We know little about the relative efficiency or appropriateness under varying conditions of different methods of gathering data. We recognize that improved data quality carries a cost and that the improvements we seek must be cost-effective. In some cases, insistence on more refinement of data can constitute "overkill" that consumes resources without adding to new knowledge. Nonetheless, we insist that anthropologists can make fundamental contribu-

tions to understanding when different methods make a difference.

3. Understanding of the critical importance of sampling in data collection and training in sampling procedures. Some innovative sampling designs focus on behaviors, events, or actions rather than on persons as the units of analysis. This kind of sampling requires special attention to definition of the appropriate universe. New sampling techniques are easy to master, even if the statistical rationale for some of them is devilishly difficult.

4. Training in the use of available tools for management of large amounts of qualitative and quantitative data. Microcomputer techniques for storage of field notes, entry of coded data for later processing, and qualitative and statistical analysis of data have been developed, but knowledge of the latest techniques is uneven, at best, in the profession. Most anthropological research should involve a fine-tuned blending of quantitative and qualitative data. Training in both qualitative and quantitative methods for data gathering and data analysis must be more widespread and at a much higher level.

In the long run, we hope, cultural anthropology will fulfill its promise as a science of humanity. If it is to do so, we must seek ways to produce, to integrate, and to disseminate knowledge about methodology. This will require the development of summer institutes and, eventually, several centers of excellence in this field—centers where scholars can spend a year or more doing research, training others, and being trained in the techniques that are constantly being developed on the fastest-moving edge of our discipline.

References Cited

AGAR, M. H. 1980. *The professional stranger: An informal introduction to ethnography*. New York: Academic Press.

ARABIE, P., and S. A. BOORMAN. 1973. Multidimensional scaling of measures of distance among partitions. *Journal of Mathematical Psychology* 10:148–203.

ARONSON, E. 1980. 3d edition. *The social animal*. San Francisco: Freeman.

BARKER, R. G., and H. WRIGHT. 1951. *The boy's day*. New York: Harper and Row.

BARRON, B. A. 1977. The effects of misclassification on the estimation of relative risk. *Biometrics* 33:414–18.

BARZUN, J., and H. GRAFF. 1970. Revised edition. *The modern researcher*. New York: Harcourt, Brace and World.

BATESON, G., and M. MEAD. 1942. *Balinese character: A photographic analysis*. New York: New York Academy of Sciences.

BEATON, G. H. 1982. "What do we think we are estimating?" in *Proceedings of the Symposium on Dietary Data Collection and Significance*. Edited by V. A. Beal and M. J. Laus, pp. 36–48. Massachusetts Agricultural Experiment Station Research Bulletin 675.

BEATON, G. H., J. MILNER, V. MCGUIRE, T. E. FEATHER, and J. A. LITTLE. 1983. Sources of variance in 24-hour dietary recall data: Implications for nutritional design and interpretation. *American Journal of Clinical Nutrition* 37:986–95.

BERLIN, B. 1984. Contributions of Native American collectors to the ethnobotany of the neotropics. *Advances in Economic Botany* 1:24–33.

BERNARD, H. R., and M. J. EVANS. 1983. New microcomputer techniques for anthropologists. *Human Organization* 42:182–85.

BERNARD, H. R., and P. D. KILLWORTH. 1977. Informant accuracy in social network data. 2. *Human Communication Research* 4:3–18.

BERNARD, H. R., P. D. KILLWORTH, and L. SAILER. 1980. Informant accuracy in social network data. 4. A comparison of clique-level structure in behavioral and cognitive data. *Social Networks* 2:191–218.

———. 1982. Informant accuracy in social network data. 5. An experimental attempt to predict actual communication from recall data. *Social Science Research* 11:30–66.

BERNARD, H. R., P. KILLWORTH, D. KRONENFELD, and L. SAILER. 1984. The problem of informant accuracy: The validity of retrospective data. *Annual Review of Anthropology* 13:495–517.

BERNARD, H. R., and J. SALINAS. 1976. *Otomi folk tales and jokes*. (Native American Texts Series, International Journal of American Linguistics.) Chicago: University of Chicago Press.

BLACKING, J. 1984. "Audio-visual equipment in general ethnographic studies," in *Ethnographic research: A guide to general conduct*. Edited by R. F. Ellen, pp. 199–206. London: Academic Press.

BLOCH, M. 1963. *The historian's craft*. New York: Knopf.

BLOCK, G. 1982. A review of validations of dietary assessment methods. *American Journal of Epidemiology* 115:492–505.

BLURTON-JONES, N. G. 1972. *Ethnological studies of child behavior*. Cambridge: Cambridge University Press.

BOHANNAN, P. J. 1954. Translation: A problem in anthropology. *The Listener*, May 13, pp. 815–16.

BOICE, R. 1983. Observational skills. *Psychological Bulletin* 193:3–219.

BOORMAN, S. A., and P. ARABIE. 1972. "Structural measures and the method of sorting," in *Multidimensional scaling: Theory in the behavioral sciences*, vol. 1. Edited by R. Shepard, A. K. Romney, and S. B. Nerlove, pp. 226–50. New York: Academic Press.

BOSTER, J. S. 1985. "Requiem for the omniscient informant: There's life in the old girl yet," in *Directions in cognitive anthropology*. Edited by J. Dougherty. Urbana: University of Illinois Press.

———. n.d. Cross-cultural sharing in biological classification. MS.

BOWER, G., and E. R. HILGARD. 1981. 5th edition. *Theories of learning*. Englewood Cliffs: Prentice-Hall.

BOWER, G., K. P. MONTERIOT, and S. G. GILLIGAN. 1978. Emotional mood as a context for learning and recall. *Journal of Verbal Learning and Verbal Behavior* 17:573–85.

BRADBURN, N. M., and S. SUDMAN. 1979. *Improving interview method and questionnaire design: Response effects to threatening questions in survey research*. San Francisco: Jossey-Bass.

BROWN, G. W., and T. HARRIS. 1978. *Social origins of depression*. New York: Free Press.

BURGESS, R. G. 1982. "The unstructured interview as conversation," in *Field research: A sourcebook and field manual*. Edited by R. G. Burgess. London: Allen and Unwin.

BURTON, M. L. 1968. *Multidimensional scaling of role terms*. Ann Arbor: University Microfilms.

———. 1972. "Semantic dimensions of occupation names," in *Multidimensional scaling: Applications in the behavioral sciences*, vol. 2. Edited by A. K. Romney, R. N. Shepard, and S. B. Nerlove, pp. 55–71. New York: Academic Press.

———. 1975. Dissimilarity measures for unconstrained sorting data. *Multivariate Behavioral Research* 10:4019–24.

BURTON, M. L., J. FERREIRA-PINTO, and R. MAGANA. 1985. Occupational cognition in three nations: Relationships between classification and evaluation. MS.

BURTON, M. L., and S. B. NERLOVE. 1976. Balanced design for triads tests. *Social Science Research* 5:247–67.

BURTON, M. L., and A. K. ROMNEY. 1975. A multidimensional representation of role terms. *American Ethnologist* 2:397–407.

BYERS, T. E., R. I. ROSENTHAL, J. R. MARSHALL, T. F. RZEPKA, K. M. CUMMINGS, and S. GRAHAM. 1983. Dietary history from the distant past: A methodological study. *Nutrition and Cancer* 5:69–77.

CANNELL, C. F., and R. L. KAHN. 1968. "Interviewing," in *The handbook of social psychology*, 2d edition, vol. 2. Edited by G. Lindzey and A. Aronson, pp. 526–95. Reading, Mass.: Addison-Wesley.

CANNELL, C. F., K. H. MARQUIS, and A. LAURENT. 1977. *A summary of studies of interviewing methodology*. (Vital and Health Statistics, series 2, 69.) Rockville, Md.: U.S. National Center for Health Statistics.

CASAGRANDE, J. B. 1954. The ends of translation. *International Journal of American Linguistics* 20:335–40.

CASAGRANDE, J. B., and K. L. HALE. 1967. "Semantic relations in Papago folk definitions," in *Studies in Southwestern linguistics*. Edited by D. Hymes and D. E. Bittle, pp. 81–105. The Hague: Mouton.

CARROLL, J. D., and P. ARABIE. 1980. Multidimensional scaling. *Annual Review of Psychology* 31:607–49.

CHAPPLE, E. D., and C. M. ARENSBERG. 1940. Measuring human relations: An introduction to the study of the interaction of individuals. *Genetic Psychology Monographs* 22:3–147.

CHU, S. Y., L. N. KOLONEL, J. H. HANKIN, and J. LEE. 1984. A comparison of frequency and quantitative dietary methods for epidemiologic studies of diet and disease. *American Journal of Epidemiology* 119:323–34.

CIMINERO, A. R., K. S. CALHOUN, and H. E. ADAMS. Editors. 1977. *Handbook of behavioral assessment*. New York: Wiley.

COLBY, B. N. 1966. Ethnographic semantics: A preliminary survey. CURRENT ANTHROPOLOGY 7:3–17.

COLLIER, J., JR. 1967. *Visual anthropology: Photography as a research method*. New York: Holt, Rinehart and Winston.
COOK, T. D., and D. T. CAMPBELL. 1979. *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.
COOMBS, C. H. 1964. *A theory of data*. New York: Wiley.
COXON, A. P. M., and C. J. JONES. 1978. *The images of occupational prestige*. London: Macmillan.
———. 1979*a*. *Class and hierarchy*. London: Macmillan.
———. 1979*b*. *Measurement and meaning: Techniques and methods of studying occupational cognition*. London: Macmillan.
D'ANDRADE, R. G. 1974. "Memory and the assessment of behavior," in *Measurement in the social sciences*. Edited by H. Blalock, pp. 159–86. Chicago: Aldine.
D'ANDRADE, R. G., N. QUINN, S. B. NERLOVE, and A. K. ROMNEY. 1972. "Categories of disease in American-English and Mexican-Spanish," in *Multidimensional scaling: Applications in the behavioral sciences*, vol. 2. Edited by A. K. Romney, R. N. Shepard, and S. B. Nerlove, pp. 19–54. New York: Academic Press.
DEHAVENON, A. L. 1978. Superordinate behavior in urban homes: A video analysis of request-compliance and food control behavior in two black and two white families living in New York City. Ph.D. diss., Columbia University, New York, N.Y.
DEWALT, B. 1979. *Modernization in a Mexican ejido: A study in economic adaptation*. New York: Cambridge University Press.
DIVALE, W. T. 1975. Temporal focus and random error in cross-cultural hypothesis tests. *Behavior Science Research* 10:19–36.
———. 1976. Female status and cultural evolution: A study of ethnographer bias. *Behavior Science Research* 11:169–212.
DU BOULAY, J., and R. WILLIAMS. 1984. "Collecting life histories," in *Ethnographic research: A guide to general conduct*. Edited by R. F. Ellen, pp. 247–57. London: Academic Press.
DUNDES, A. 1965. *The study of folklore*. Englewood Cliffs: Prentice-Hall.
EMBER, C. R. 1983. The relative decline in women's contribution to agriculture with intensification. *American Anthropologist* 85:285–304.
EMBER, M. 1985. Evidence and science in anthropology: Reflections on the Freeman-Mead controversy. *American Anthropologist* 87:906–10.
EPSTEIN, A. L. Editor. 1967. *The craft of social anthropology*. London: Tavistock.
FARADAY, A., and K. PLUMMER. 1979. Doing life histories. *Sociological Review* 27:773–98.
FILLENBAUM, S., and A. RAPOPORT. 1971. *Structures in the subjective lexicon*. New York: Academic Press.
FISCHER, D. H. 1970. *Historians' fallacies: Toward a logic of historical thought*. New York: Harper Torchbooks.
FISCHOFF, B. 1983. Predicting frames. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 9:103–16.
FISK, A. D., and W. SCHNEIDER. 1984. Memory as a function of attention, level of processing, and automatization. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10:181–97.
FRAKE, C. O. 1964. Notes on queries in ethnography. *American Anthropologist* 66(3, pt. 2):132–45.
FREEMAN, D. n.d. Reply to Ember. *American Anthropologist*. In press.
FREEMAN, H., A. K. ROMNEY, J. FERREIRA-PINTO, R. E. KLEIN, and T. SMITH. 1981. Guatemalan and U.S. concepts of success and failure. *Human Organization* 40:140–45.
FURBEE, L., and R. A. BENFER. 1983. Cognitive and geographic maps: Study of individual variation among Tojolabal Mayans. *American Anthropologist* 85:305–34.
GARRO, L. 1983. Individual variation in a Mexican folk medical belief system: A curer-lay comparison. Ph.D. diss., University of California, Irvine, Calif.
GLADWIN, C. 1976. A view of Plan Puebla: An application of hierarchical decision models. *American Journal of Agricultural Economics* 59:881–87.
———. 1980. "A theory of real-life choice: Applications to agricultural decisions," in *Agricultural decision making*. Edited by P. Bartlett, pp. 45–85. New York: Academic Press.
GLADWIN, H., and M. MURTAGH. 1980. "The attentive-preattentive distinction in agricultural decision making," in *Agricultural decision making*. Edited by P. Bartlett, pp. 115–36. New York: Academic Press.
GOODENOUGH, W. 1956. Componential analysis and the study of meaning. *Language* 32:195–216.
GOULD, R. A., and M. B. SCHIFFER. Editor. 1981. *The archaeology of us*. Albuquerque: University of New Mexico Press.
GOWARD, N. 1984. "Personal interaction and adjustment," in *Ethnographic research: A guide to general conduct*. Edited by R. F. Ellen, pp. 178–87. London: Academic Press.
GREGG, L. Editor. 1972. *Cognition in learning and memory*. New York: Wiley.
GROSS, D. 1984. Time allocation: A tool for the study of cultural behavior. *Annual Review of Anthropology* 13:119–58.
HALE, K. L. 1969. "Some questions about anthropological linguistics: The role of native knowledge," in *Reinventing anthropology*. Edited by D. Hymes, pp. 382–97. New York: Random House.
HAMMEL, E. A. 1962. Social rank and evolutionary position in a coastal Peruvian village. *Southwestern Journal of Anthropology* 18:199–215.
———. 1970. The ethnographer's dilemma: Alternative models of occupational prestige in Belgrade. *Man* 5:652–70.
HANDLIN, O. 1979. *Truth in history*. Cambridge: Harvard University Press.
HANDLIN, O., et al. 1954. *Harvard guide to American history*. Cambridge: Harvard University Press.
HARRIS, M. 1964. *The nature of cultural things*. New York: Random House.
HENLEY, N. M. 1969. A psychological study of the semantics of animal terms. *Journal of Verbal Learning and Verbal Behavior* 8:176–84.
HENRY, J. 1951. The economics of Pilagá food distribution. *American Anthropologist* 53:187–219.
HENRY, J., and Z. H. HENRY. 1944. *Doll play of Pilagá children*. New York: Vintage.
HERRMAN, N. 1985*a*. Retrospective information from questionnaires. 1. Comparability of primary respondents and their next-of-kin. *American Journal of Epidemiology* 121:937–47.
———. 1985*b*. Retrospective information from questionnaires. 2. Intrarater reliability and comparison of questionnaire types. *American Journal of Epidemiology* 121:948–53.
HOLY, L. 1984. "Theory, methodology, and the research process," in *Ethnographic research: A guide to general conduct*. Edited by R. F. Ellen. London: Academic Press.
HOUSTON, M. J., and S. SUDMAN. 1975. A methodological assessment of the use of key informants. *Social Science Research* 4:151–64.
HUMBLE, C. G., J. M. SAMET, and B. E. SKIPPER. 1984. Comparison of self- and surrogate-reported dietary information. *American Journal of Epidemiology* 119:86–98.
HYMAN, H. H., and W. J. COBB. 1975. *Interviewing in social research*. Chicago: University of Chicago Press.
JELLIFFE, D. B. 1966. *The assessment of the nutritional status of the community*. Geneva: World Health Organization.
JOHNSON, A. 1975. Time allocation in a Machiguenga community. *Ethnology* 14:301–10.
———. 1978. *Quantification in cultural anthropology*. Stanford: Stanford University Press.
JOHNSON, A., and G. C. BOND. 1974. Kinship, friendship, and exchange in two communities: A comparative analysis of norms and behavior. *Journal of Anthropological Research* 30:55–68.
JOHNSON, O. 1980. The social context of intimacy and avoidance: A videotape study of Machiguenga meals. *Ethnology* 19:353–66.
KANG, G. E., and T. K. KANG. 1981. Evaluation and modification of the SES random error procedure. *Behavior Science Research* 16:1–26.
KARINPAA, A., and R. SEPPANEN. 1983. A comparison of a self-administered dietary assessment questionnaire and a short interview. *Nutrition Research* 3:285–91.
KEEN, I. 1985. Definitions of kin. *Journal of Anthropological Research* 41:62–90.
KENT, R. N., and S. N. FOSTER. 1977. "Direct observational procedures: Methodological issues in naturalistic settings," in *Handbook of behavioral assessment*. Edited by A. R. Ciminero, K. S. Calhoun, and H. E. Adams, pp. 329–64. New York: Wiley.
KILLWORTH, P. D., and H. R. BERNARD. 1976. Informant accuracy in social network data. *Human Organization* 35:269–86.
———. 1979. Informant accuracy in social network data. 3. A comparison of triadic structures in behavioral and cognitive data. *Social Networks* 2:19–46.
KINCAID, D. L., J. D. YUM, and J. WOELFEL. 1983. The cultural convergence of Korean immigrants in Hawaii: Empirical test of a mathematical theory. *Quality and Quantity* 18:519–78.
KIRK, L., and M. BURTON. 1977. Meaning and context: A study of contextual shifts in meaning of Maasai personality descriptors. *American Ethnologist* 4:734–61.
KLUCKHOHN, C. 1945. "The personal document in anthropological science," in *The use of personal documents in history, anthropology, and sociology*. Edited by L. Gottschalk et al., pp. 79–105. New York: Social Science Research Council.

LANGNESS, L. L. 1965. *The life history in anthropological science*. London: Holt, Rinehart and Winston.
LAZARSFELD, P. F., and N. W. HENRY. 1968. *Latent structure analysis*. Boston: Houghton Mifflin.
LEE, J., L. N. KOLONEL, and J. H. HANKIN. 1983. On establishing the interchangeability of different dietary-intake assessment methods used in studies of nutrition and cancer. *Nutrition and Cancer* 5:215–18.
LEVINE, H. G., R. GALLIMORE, T. S. WEISNER, and J. L. TURNER. 1980. Teaching participant-observation research methods: A skills-building approach. *Anthropology and Education Quarterly* 11:38–54.
LEVINSON, D. Editor. 1978. *A guide to social theory: Worldwide cross-cultural tests*. New Haven: HRAF Press.
LEVINSON, D., and M. J. MALONE. 1980. *Toward explaining human culture*. New Haven: HRAF Press.
LÉVI-STRAUSS, C. 1963. "The structural study of myth," in *Structural anthropology*. Edited by C. Lévi-Strauss, pp. 206–31. New York: Basic Books.
LEWIS, O. 1961. *The children of Sánchez: Autobiography of a Mexican family*. New York: Random House.
LIEBERMAN, D., and W. M. DRESSLER. 1977. Bilingualism and cognition of St. Lucian disease terms. *Medical Anthropology* 1(1): 81–110.
LINDZEY, G., and E. ARONSON. 1968. 2d edition. *The handbook of social psychology*. 2 vols. Reading, Mass.: Addison-Wesley.
LIU, K., J. STAMLER, A. DYER, J. MCKEEVER, and P. MCKEEVER. 1978. Statistical methods to assess and minimize the role of intra-individual variability in obscuring the relationship between dietary lipids and serum cholesterol. *Journal of Chronic Disorders* 31:399–418.
LORD, F. M., and M. R. NOVICK. 1968. *Statistical theories of mental test scores*. Reading, Mass.: Addison-Wesley.
LOUNSBURY, F. G. 1956. A semantic analysis of the Pawnee kinship usage. *Language* 32: 158–94.
LYON, J. L., J. W. GARDNER, D. W. WEST, and A. M. MAHONEY. 1983. Methodological issues in epidemiological studies of diet and cancer. *Cancer Research*, suppl. 43:2392S–96S.
MCGEE, D., G. RHOADS, J. HANKIN, K. YANO, and J. TILLOTSON. 1982. Within-person variability of nutrient intake in a group of Hawaiian men of Japanese ancestry. *American Journal of Clinical Nutrition* 36:657–63.
MCSWEENEY, B. G. 1979. Collection and analysis of data on rural women's time use. *Studies in Family Planning* 10:3719–83.
MAJNEP, S., and R. BULMER. 1977. *Birds of my Kalam country*. Auckland: University of Auckland Press.
MANES, J. I. 1976. *Semantic relations in American English folk definitions*. Ann Arbor: University Microfilms.
MEAD, M. 1970. "The art and technology of fieldwork," in *A handbook of method in cultural anthropology*. Edited by R. Naroll and R. Cohen, pp. 246–65. New York: Natural History Press.
MERTON, R. K. 1956. *The focused interview*. New York: Harper and Row.
METZGER, D., and G. E. WILLIAMS. 1966. Some procedures and results in the study of native categories: Tzeltal firewood. *American Anthropologist* 68:389–407.
MILGRAM, S. 1977. *The individual in a social world: Essays and experiments*. Reading, Mass.: Addison-Wesley.
MILGRAM, S., and L. SHOTLAND. 1973. *Television and antisocial behavior: Field experiments*. New York: Academic Press.
MITCHELL, J. C. 1984. "Case studies," in *Ethnographic research: A guide to general conduct*. Edited by R. F. Ellen, pp. 237–41. London: Academic Press.
MOLLER JENSON, O., J. WAHRENDORF, A. ROSENQUIST, and A. GESER. 1984. The reliability of questionnaire-derived historical dietary information and temporal stability of food habits in individuals. *American Journal of Clinical Nutrition* 120:281–90.
MULLEN, B. J., N. J. KRANTZLER, L. E. GRIVETTI, H. G. SCHUTZ, and H. L. MEISELMAN. 1984. Validity of a food frequency questionnaire for the determination of individual food intake. *American Journal of Clinical Nutrition* 39:136–43.
MUNROE, R. H., and R. L. MUNROE. 1969. Local personnel as scientific observers. Paper presented to the annual meeting of the Southwestern Anthropological Association, Las Vegas, Nev.
MURDOCK, G. P. 1977. Major emphases in my comparative research. *Behavior Science Research* 12:217–22.
NAG, M. 1962. *Factors affecting fertility in nonindustrial societies: A cross-cultural study*. Yale University Publications in Anthropology 66.
NAROLL, R. 1962. *Data quality control*. New York: Free Press.
———. 1970. "Data quality control in cross-cultural surveys," in *A handbook of method in cultural anthropology*. Edited by R. Naroll and R. Cohen, pp. 927–45. New York: Natural History Press.
———. 1977. Cost-effective research vs. safer research: A reply to Murdock. *Behavior Science Research* 12:223–26.
NAROLL, R., G. L. MICHIK, and F. NAROLL. 1976. *Worldwide theory testing*. New Haven: HRAF Press.
NELSON, D. L., M. T. BAJO, and D. CASANUEVA. 1985. Prior knowledge and memory: The influence of natural category size as a function of intention and distraction. *Journal of Experimental Pscyhology: Learning, Memory, and Cognition* 11:94–105.
NERLOVE, S. B., and A. S. WALTERS. 1977. Pooling intra-cultural variation: Toward empirically based statements of community consensus. *Ethnology* 16:427–41.
NIDA, E. 1964. *Towards a science of translation*. Leiden: Brill.
NUNNALLY, J. C. 1978. *Psychometric theory*. New York: McGraw-Hill.
OSGOOD, C. E., G. J. SUCI, and P. H. TANNENBAUM. 1967. *The measurement of meaning*. Urbana: University of Illinois Press.
OTTERBEIN, K. 1969. Basic steps in conducting a cross-cultural study. *Behavior Science Research* 4:221–36.
PELTO, P., and G. PELTO. 1978. *Anthropological research: The structure of inquiry*. London: Cambridge University Press.
PERCHONOCK, N., and O. WERNER. 1968. Navajo systems of classification: Some implications of food. *Ethnology* 8:299–342.
PHILLIPS, H. P. 1959. Problems of translating and meaning in fieldwork. *Human Organization* 18:395–415.
PODOLEFSKY, A., and C. MCCARTY. 1983. Topical sorting: A technique for computer-assisted qualitative data-analysis. *American Anthropologist* 85:886–90.
POGGIE, J. J., JR. 1972. Toward quality control in key informant data. *Human Organization* 31:23–30.
POWDERMAKER, H. 1966. *Stranger and friend: The way of an anthropologist*. New York: Norton.
QUINN, N. 1978. Do Mfantse fish sellers estimate probabilities in their heads? *American Ethnologist* 5:206–26.
REISS, N. 1983. Speech act taxonomy as a tool for ethnographic description: An analysis based on videotapes of continuous behavior in two New York households. Ph.D. diss., Columbia University, New York, N.Y.
ROBERTS, F. S. 1979. *Measurement theory with applications to decision making, utility, and the social sciences*. Reading, Mass.: Addison-Wesley.
ROBERTS, J. M., and G. E. CHICK. 1979. "Butler County eight ball: A behavioral space analysis," in *Sports, games, and play*. Edited by J. H. Goldstein, pp. 65–99. Hillsdale, N.J.: Lawrence Erlbaum Associates.
ROBERTS, J. M., G. E. CHICK, M. STEPHANSON, and L. L. HYDE. 1981. "Inferred categories for tennis play: A limited semantic analysis," in *Play as context*. Edited by A. B. Cheska, pp. 181–95. New York: Leisure Press.
ROBERTS, J. M., T. V. GOLDER, and G. E. CHICK. 1980. Judgment, oversight, and skill: A cultural analysis of P3 pilot error. *Human Organization* 39:5–21.
ROBERTS, J. M., and S. NUTTRASS. 1980. "Women and trapshooting," in *Play and culture*. Edited by H. B. Schwartzman, pp. 262–91. West Point: Lei Press.
ROGOFF, B. 1978. Spot observations: An introduction and examination. *Quarterly Newsletter of the Institute of Comparative Human Development* 2, 2C:21–26.
ROHAN, T. E., and J. D. POTTER. 1984. Retrospective assessment of dietary intake. *American Journal of Epidemiology* 120:876–87.
ROHNER, R. 1966. "Franz Boas: Ethnographer of the Northwest Coast," in *Pioneers of American anthropology*. Edited by J. Helm, pp. 149–222. Seattle: University of Washington Press.
———. 1977. Why cross-cultural research. *Annals of the New York Academy of Sciences* 285:3–12.
ROHNER, R., B. R. DEWALT, and R. C. NESS. 1973. Ethnographer bias in cross-cultural research. *Behavior Science Research* 8:275–317.
ROMNEY, A. K., and R. G. D'ANDRADE. 1964. Cognitive aspects of English kin terms. *American Anthropologist* 66(3, pt.2):230–42.
ROMNEY, A. K., T. SMITH, H. E. FREEMAN, J. KAGAN, and R. E. KLEIN. 1979. Concepts of success and failure. *Social Science Research* 8:302–26.
ROMNEY, A. K., and S. C. WELLER. 1984. Predicting informant accuracy from patterns of recall among individuals. *Social Networks* 6:519–77.
ROMNEY, A. K., S. C. WELLER, and W. H. BATCHELDER. n.d. Culture as consensus: A theory of culture and informant accuracy. *American Anthropologist*. In press.

ROSENTHAL, R. 1966. *Experimenter effects in behavioral research.* New York: Appleton-Century-Crofts.
ROSENTHAL, R., and L. JACOBSON. 1968. *Pygmalion in the classroom.* New York: Holt, Rinehart and Winston.
RUSH, D., and A. R. KRISTAL. 1982. Methodologic studies during pregnancy: The reliability of the 24-hour dietary recall. *American Journal of Clinical Nutrition* 35:1259–68.
SAILER, L. 1984. Computer-assisted anthropology. *Practicing Anthropology* 6(2):5–17.
SALINAS, J., and R. BERNARD. 1978. *The Otomí.* Albuquerque: University of New Mexico Press.
———. n.d. The flora, fauna, and religious life of the Ñähñu. 3 vols. MS.
SAMET, J. M., C. G. HUMBLE, and B. E. SKIPPER. 1984. Alternatives in the collection and analysis of food frequency interview data. *American Journal of Epidemiology* 120:572–81.
SANDAY, P. 1968. The psychological reality of American English kinship terms: An information processing approach. *American Anthropologist* 70:508–23.
SANKOFF, G. 1971. Quantitative analysis of sharing and variability in a cognitive model. *Ethnology* 10:389–408.
SCHENSUL, S. 1969. *Marginal rural peoples: Behavior and cognitive models among northern Minnesotans and western Ugandans.* Ann Arbor: University Microfilms.
SCHMINK, M. 1984. Household economic strategies: Review and research agenda. *Latin American Research Review* 19(3):87–101.
SCHOUA, A. 1985. *An English-Spanish test of decentering for the translation of questionnaires.* Ann Arbor: University Microfilms.
SEMPOS, C. T., N. E. JOHNSON, E. L. SMITT, and C. GILLIGAN. 1985. Effects of intraindividual and interindividual variation in repeated dietary records. *American Journal of Epidemiology* 121:120–30.
SHARFF, J. W. Patterns of authority in two Puerto Rican households. Ph.D. diss., Columbia University, New York, N.Y.
SHAW, B. 1980. Life history writing in anthropology: A methodological review. *Mankind* 12:226–33.
SHWEDER, R. A., and R. G. D'ANDRADE. 1980. "The systematic distortion hypothesis," in *New directions for the methodology of social and behavioral science.* Edited by R. A. Shweder. San Francisco: Jossey-Bass.
SILVERMAN, S. F. 1966. An ethnographic approach to social stratification: Prestige in a Central Indian community. *American Anthropologist* 68:899–921.
SORENSEN, E. R. 1975. *Visual evidence: An emerging force in visual anthropology.* (National Anthropological Film Center, Occasional Papers 1.) Washington, D.C.: Smithsonian Institution.
SPRADLEY, J. P. 1979. *The ethnographic interview.* New York: Holt, Rinehart and Winston.
———. 1980. *Participant observation.* New York: Holt, Rinehart and Winston.
STUFF, J. E., C. GARZA, E. D. SMITH, B. L. NICHOLS, and C. M. MONTANDON. 1983. A comparison of dietary methods in nutritional studies. *American Journal of Clinical Nutrition* 37:300–306.
SUDMAN, S., and N. M. BRADBURN. 1974. *Response effects in surveys.* Chicago: Aldine.
———. 1982. *Asking questions.* San Francisco: Jossey-Bass.
TEDLOCK, D. 1971. On the translation of style in oral narrative. *Journal of American Folklore* 84:114–33.
TILLEY, B. C., A. B. BARNES, E. BEYSTRALH, D. LABARTHE, K. L. NOLLER, T. COLTON, and E. ADAMS. 1985. A comparison of pregnancy history recall and medical records: Implications for retrospective studies. *American Journal of Epidemiology* 121:269–81.
TODD, K. S., M. HUDES, and D. H. CALLOWAY. 1983. Food intake measurement: Problems and approaches. *American Journal of Clinical Nutrition* 37:139–46.
TONKIN, E. 1984. "Language learning," in *Ethnographic research: A guide to general conduct.* Edited by R. F. Ellen, pp. 100–118. London: Academic Press.
TOPPER, M. D. 1972. *The daily life of a traditional Navajo household: An ethnographic study of human daily activities.* Ann Arbor: University Microfilms.
TREMBLAY, A., J. SEVIGNY, C. LEBLANC, and C. BOUCHARD. 1983. The reproducibility of a three-day dietary record. *Nutrition Research* 3:819–30.
TYLER, S. Editor. 1969. *Cognitive anthropology.* New York: Holt, Rinehart and Winston.
VANSINA, J. 1965. *Oral tradition: A study in historical method.* Translated by H. M. Wright. Chicago: Aldine.
VAN VELSEN, J. 1967. "The extended case method and situational analysis," in *The craft of social anthropology.* Edited by A. L. Epstein, pp. 129–52. London: Tavistock.
VOEGELIN, C. F., and F. M. VOEGELIN. 1967. "Anthropological linguistics and translation," in *To honour Roman Jakobson: Janua Linguarum,* pp. 2159–90. The Hague: Mouton.
WALLACE, A. F. C., and J. ATKINS. 1960. The meaning of kinship terms. *American Anthropologist* 62:58–79.
WEBB, E., D. T. CAMPBELL, R. D. SCHWARTZ, and L. SECHREST. 1966. *Unobtrusive measures: Nonreactive research in the social sciences.* Chicago: Rand McNally.
WEISNER, THOMAS S., R. GALLIMORE, and R. G. THARP. 1982. Concordance between ethnographer and folk perspectives: Observed performances and self-ascription of sibling care-taking roles. *Human Organization* 41:237–44.
WELLER, S. C. 1984*a*. Cross-cultural concepts of illness: Variation and validation. *American Anthropologist* 86:341–51.
———. 1984*b*. Consistency and consensus among informants: Disease concepts in a rural Mexican town. *American Anthropologist* 86:966–75.
WERNER. O. 1961. Ethnographic photography. M.A. thesis, Department of Anthropology, Syracuse University, Syracuse, N.Y.
———. 1966. Pragmatics and ethnoscience. *Anthropological Linguistics* 8:42–65.
———. 1983. Microcomputers in cultural anthropology: APL programs for qualitative analysis. *BYTE* 7(7):250–80.
WERNER, O., K. Y. BEGISHE, and M. A. AUSTIN-GARRISON. n.d. *The anatomical atlas of the Navajo.* Pine Hill, N.M.: Pine Hill Community School. In press.
WERNER, O., and D. T. CAMPBELL. 1970. "Translating, working through interpreters, and the problem of decentering," in *A handbook of method in cultural anthropology.* Edited by R. Naroll and R. Cohen, pp. 398–420. New York: Natural History Press.
WERNER, O., and J. FENTON. 1970. "Method and theory in ethnoscience or ethnoepistemology," in *A handbook of method in cultural anthropology.* Edited by R. Naroll and R. Cohen, pp. 537–78. New York: Natural History Press.
WERNER, O., A. MANNING, and K. Y. BEGISHE. 1983. "A taxonomic view of the traditional Navajo universe," in *Handbook of North American Indians,* vol. 10. Edited by A. Ortiz, pp. 579–91. Washington, D.C.: Smithsonian Institution.
WERNER, O., G. M. SCHOEPFLE, et al. 1985–86. *Systematic fieldwork.* Vol. 1. *Foundations of ethnography and interviewing;* vol. 2. *Ethnographic analysis and data management.* Beverly Hills: Sage.
WHITE, G. 1978. Ambiguity and ambivalence in A'ara personality descriptors. *American Ethnologist* 5:334–60.
WHITING, B., and C. P. EDWARDS. 1973. A cross-cultural analysis of sex differences in the behavior of children aged three through eleven. *Journal of Social Psychology* 91:171–88.
WHITING, B., and J. WHITING. 1970. "Methods for observing and recording behavior," in *A handbook of method in cultural anthropology.* Edited by R. Naroll and R. Cohen, pp. 282–315. New York: Natural History Press.
WHITING, J. W. M. 1961. "The cross-cultural method," in *Readings in cross-cultural methodology.* Edited by F. W. Moore, pp. 287–300. New Haven: HRAF Press.
WHYTE, M. K. 1978. Cross-cultural studies of women and the male bias problem. *Behavior Science Research* 13:65–80.
WINKS, R. 1968. *The historian as detective: Essays on evidence.* New York: Harper Torchbooks.
WITKOWSKI, S. 1978. Ethnographic field work: Optimal versus nonoptimal conditions. *Behavior Science Research* 13:245–54.
YOUNG, C. M. 1981. "Dietary methodology," in *Assessing changing food consumption patterns.* Edited by the Committee on Food Consumption Patterns, Food and Nutrition Board, National Research Council, pp. 89–118. Washington, D.C.: National Academy Press.
YOUNG, J. C. 1980. A model of illness treatment decisions in a Tarascan town. *American Ethnologist* 7:106–31.