

STANDARDS OF RIGOR IN ANTHROPOLOGY

by Susan Weller

The standards of rigor in anthropology are essentially the same as standards of scientific rigor across all of the social sciences. (This, of course, does not mean that all anthropologists adhere to these standards.) Although qualitative methods of measuring outcome variables can be used in experimental and evaluation research, these types of research designs are rarely used in anthropology. Anthropological research most often uses observational research designs (as opposed to true experiments with random assignment of subjects) with convenience sampling, and so I focus on those types of studies here. The most important standard is the statement of purpose (research question to be answered or hypothesis) and use of a research design that allows for the question to be answered (to contribute meaningful information, possibly allowing for falsification of a hypothesis). A second issue concerns interview methods and the reliability of the data that are collected. A final issue concerns the coding and analysis of data. A key issue is recognizing the strength of qualitative methods, in contrast to more structured interviewing methods. Qualitative methods are most appropriate for areas where little is known.

With regard to the research design, researchers must specify the number of groups in the study (whether a one group or multiple group study), the rationale for selecting a group, and consideration for extraneous variables. Often, projects have only a single group of informants and it is impossible to tell if themes are unique to that group or are general themes. For example, qualitative research with diabetic patients where diabetic patients are interviewed about their disease without a comparison group, either another group of patients with a chronic illness or some systematic subgroup of diabetic patients, can only find very general themes (see, for example: Quatromoni, Carballeira, Milbauer, Brunt, et al 1994; Anderson, Goddard, Garcia, et al 1998). In contrast, a study that, by design, interviews diabetic patients with good glycemic control and another group with poor glycemic control can identify themes that may be relevant to the management of diabetes (Savona, Miller, Quandt 2004). Themes mentioned with similar frequency in both groups are of less interest than are themes mentioned more by one group than the other group. Also, extraneous factors can be controlled for by using matching procedures in the grouping design. Matching was used by Rubel et al (1984) to control for age and gender in his study of the folk illness *susto*. In the case of diabetes, in order to focus on lifestyle variables, those with good control and those with poor glycemic control should be matched for their duration of disease and their medications (De Alba et al., under review). Thus, when informants in the two groups are approximately similar with regard to the matching characteristics, the effects of those extraneous factors are removed, and factors of interest can be highlighted.

A related issue with regard to research design is clear description and rationale for informant selection. There are a few outstanding source books for this, especially with

regard to qualitative research. The first is The Ethnographic Interview by Spradley (1979). Spradley is appropriate for research projects during their initial or preliminary stages. The focus is on how to get started, including who to interview and how to figure-out what questions to ask. The best single source on selecting informants is the Sage publication in the Qualitative Series by Jeffrey Johnson (1990), Selecting Ethnographic Informants. Johnson describes how groups of informants should be selected according to theoretical interests, e.g., the purpose of the study. Although there are many books on representative sampling, this is the only one that discusses how to select a stratified convenience sample.

Besides an appropriate research design, studies need to use interviewing methods that obtain reliable and valid data. A big issue seems to be whether to use group or individual interviews. Focus groups are very popular right now, but few researchers seem to understand that their effective sample size is somewhere between the number of groups and the number of participants, and is not the total number of participants. Evidence for reaching the saturation point and adequacy for the sample should be provided. Saturation is the point at which there are no more or few new items. Work by Fern (1982, also discussed in Morgan 1996) indicates that individual interviews are more efficient in eliciting topics; focus groups get about 60% of the amount of information as individual interviews with the same number of informants. For highly shared domains, free-recall listing interviews (Weller & Romney 1988) can reach saturation with as few as 10 to 20 people. Similarly, focus groups on highly shared domains are reported to reach saturation with as few as 4-6 groups of about 8 people each. One thing is clear: the total number of hours spent in focused interviewing is directly related to the amount of information obtained. Two studies that used focus groups to interview Latino diabetic patients about their disease using four groups of approximately eight people each, only had six to eight total hours of interviewing (Quatromoni, et al 1994; Anderson, et al 1998). In contrast, Hunt, Pugh, and Valenzuela (1998) conducted similar interviews individually with diabetic patients with slightly over 100 total hours of interviewing (51 patients, with about two hours per patient). The amount of information obtained by Hunt et al is evident in the series of papers published from those interviews.

The coding and analysis of data needs to include an assessment of the reliability of coding procedures. Some techniques use minimal interpretation of responses, and use verbatim words or phrases. Others interpret responses, and form categories of responses. Methods that require interpretation need to include some method of inter-rater assessment of coding responses in to those categories. A standard way to handle such data is to specify the coding rules and criteria and then have at least two people independently code the text responses. Although this step is sometimes ignored, it is essential if the codes will be used in further analysis. Hruschka, Schwartz, St. John, et al (2004) describe an iterative procedure for bringing coders to acceptable levels of inter-rater agreement.

CROSS-DISCIPLINE EMPHASIS

Qualitative interviewing methods can be used at any stage in a research project. Open-ended interviews can be used in the initial stages of a project, and the results can be used to design structured interview materials. Open-ended interviews can be conducted after a survey has been completed, to better understand responses structured questions (Kempton et al; Baer 1996). Or, a study may rely exclusively on qualitative interviewing. The difference across the social sciences is mainly in the emphasis and where it is integrated into the research process. Using broad generalizations, psychologists have “pilot” studies or an “instrument development” phase where open-ended questions are used to elicit themes for scale development. In psychology and sociology, may use qualitative interviews during a “pilot” phase to explore how respondents interpret survey questions (what are they really being asked?) and what responses might be expected (see, for example, the cognitive survey work). Anthropology, in contrast, may use qualitative methods for a study that is descriptive.