JEFFREY C. JOHNSON

Four

Research Design and Research Strategies

We need a powerful mode of argumentation, a mode that ensures we can represent our representations in credible ways. In such worlds, a systematic argument enjoys a star-spangled legitimacy. We need a way to argue what we know based on the process by which we came to know it. That's what I seek, not as the only possible representation that our field can offer, but as an essential lever to try and move the world.

Michael A. Agar (1996:13)

Introduction

In a complex world of competing arguments, who is to be believed or trusted? Are data themselves, independently of how they were conceived and collected, proper evidence for making a case? Although some may be swayed by the elegance of a well-written essay, for many it's crucial to know something about the author, his or her motivations, experiences, skills, methods of investigation, and so on before passing judgment on the conclusions. In Agar's statement above, we get the impression that a credible argument should be systematic and based on a process that informs us about how researchers came to know what they know.

It is the articulation of this "process by which we came to know it" that reflects the elements of research design. For Stinchcombe (1987:23), the observations produced by how a study was designed are fundamental to the proper assessment of empirical evidence: "We always want to reject evidence if it can be explained by the design of the research or by a large number of small, unorganized causes." Some things, like perceptual errors, that hinder our observation may be beyond our control. Some things, like site selection, sampling, measurement, and recording are at least partly within our control. The value of empirical evidence can only be properly evaluated by understanding the details of how the research was conducted.

According to Pelto and Pelto (1978:291): "Research design involves combining the essentials of investigation into an effective problem-solving sequence. Thus the plan of research is a statement that concentrates on the components that *must be present* in order for the objectives of the study to be realized." This statement illustrates at least two important elements of research design.

First, research design involves an a priori plan or strategy for all phases of the research (such as data collection and analysis) including, for some researchers, the production of the final product (like an ethnography). By definition, a plan cannot deal with the unanticipated or unknown realities of research, such as tragedies or acts of nature that disrupt fieldwork. A good understanding of the research problem and the research site allows us to plan for some contingencies, but there is no research design crystal ball. In fact, chance factors often lead to great discoveries or unexpected findings. Still, while luck plays a role in research, planning for such luck is not within the realm of research design (Kirk and Miller 1986).

Second, an idealized plan gives guidelines for linking theory to the methods of data collection and analysis that yield either valid or "defensible" results. I use "defensible" in addition to "valid," which I normally use, to make readers aware that I am broadening the traditional application of research design to include the variety of research strategies found in anthropology today. Interpretive, hermeneutic, and postmodern approaches make little explicit reference to ethnographic design issues, but well-written examples from ethnography may provide "moral evidence" to deal with current social problems, moving people (including politicians) in ways that numerical facts can't (Seidman 1994:134). Nevertheless, a well-articulated project design helps "to promote the effective conduct of research," whether one starts from a positivist or humanist perspective (Ellen 1984:158).

On a practical level, good research design is essential in the competition for research grants and contracts. There is much variation in what funding agencies and foundations expect regarding research design. One agency may require a detailed description of the proposed project paying attention to the research design logic of science (for example, validity, reliability, hypotheses, etc., see also Plattner [In press]); others may require a description of the research problem and site but require less detail about the methods of data collection and analysis. All funding agencies expect a well-organized outline of the proposed project—one that meets the design expectations of peer reviewers and agency personnel.

A distinction needs to be made between what's sometimes called the laundry-list component of research and research *design*. The laundry-list component is important. It involves details about getting into and out of the field situation, travel arrangements, getting proper government permissions, making contacts at the field site, arranging for living accommodations, and so on. *Design*, on the other hand, involves the methodological and analytical details that contribute to the credibility, validity, believability, or plausibility of any study. In this chapter I concentrate on elements of design related to the production of valid results or a believable ethnographic account.

The Need for Design

Evidence for the power of research design is all around us. The invention of the simple control/treatment design of clinical trials allowed researchers in this century to evaluate competing therapies and to select the ones that worked best. One result is that infectious childhood diseases that killed thousands of young people a century ago are today only a memory in industrialized countries. The lessons learned from controlled experimentation are applied today to the policy arena where groups are in conflict over resources or because of social inequalities (Johnson and Pollnac 1989; Porter 1995). Members of such competing groups—such as large-scale commercial producers, commodity producers, environmental groups, and real estate developers—believe strongly in their positions. They have evidence, often anecdotal, that their positions are credible. Without some unbiased means for assessing the evidence, the truth is only be a matter of who has the most political clout.

The outcry for a ban on nets in tuna fishing is a famous recent example. Environmental organizations launched campaigns to ban nets in tuna fishing because dolphins are often caught incidentally in that fishery. Media campaigns in the U.S. showing pictures of dolphins being caught in nets (generally not in U.S. waters), contributed to Florida's totally banning fishing nets—even though no marine mammals were threatened by the use of nets in Florida waters. Thus, policy emerges from interactions between groups of differing political, ideological, social, and economic backgrounds.

There has been similar concern over the incidental catch of harbor porpoises by net fishers in New England (Schneider 1996). This case led to a systematic test of a technology that might ameliorate the problem. Wildlife conservationists petitioned the U.S. federal government in 1991 to declare harbor porpoises a threatened species. In response, the fishing industry proposed the voluntary use of "pingers"—an underwater acoustic device—to keep porpoises from their nets. The effectiveness of the device, however, was in question, and there was no firm evidence in the literature about it. Fishers petitioned the federal government to fund a study of pinger effectiveness. The study used the classic control/treatment design in which catch rates for a set of nets with pingers were compared to catch rates for set of nets without pingers.

In the first experiment, the control net caught 10 porpoises while the treatment net caught none. Some conservationist groups claimed the study was biased in that the treatment nets were placed in areas known not to have large numbers of porpoises. So another study was conducted placing experimental treatment and control nets in the same proximity. This time, the treatment net caught only 1 porpoise while the control net caught 32. Some environmental groups were still concerned that evidence with more statistical power was needed. Lobbying efforts by fishers yielded more funds for a larger, more comprehensive study involving more than 10,000 fishing nets. Both control and treatment nets were outfitted with pingers, but only the pingers on treatment nets would activate once placed in the water. Thus, fishers were blind as to which nets were control and which were treatment—a classic double-blind experimental design. Again the evidence was impressive: The treatment nets caught 2 porpoises (1 was thought to be deaf), while the control nets caught 25.

The issue is still under debate, but this series of studies illustrates how the elements of research design help muster evidence in light of competing beliefs and philosophies. In each successive study, investigators tried to control for as many extraneous variables as possible so that the hypothesized effect could be assessed (that is, the effectiveness of pingers compared to not using pingers). The logic of the research design contributed to the production of credible results.

Although the power of experimental design is evident, concern for its application in anthropology—particularly cultural anthropology—has been limited. Some early exceptions include Brim and Spain's (1974) book on hypothesis-testing designs, Pelto and Pelto's (1978) book on research methodology in cultural anthropology, and Naroll and Cohen's (1973) *A Handbook of Method in Cultural Anthropology*, which has several chapters that address issues in research design (LeVine 1973; Sechrest 1973; Spindler and Goldschmidt 1973). Bernard (1994) has elaborated in more detail on issues of design, but his treatment is necessarily limited, given his task of describing the range of methods available to anthropologists.

If research design gets relatively little attention from anthropologists, other social scientists have written volumes about it. What should we make of this apparent dearth of specific treatments of research design in cultural anthropology? I don't think we should make too much of it because the important elements of research design—reliability, informant accuracy, validity, objectivity, and operationalization of theoretical concepts—have been present in the writings of cultural anthropologists even before Boas.

Boas, Malinowski, and Research Design in the Scientific Tradition

Boas and most of his students advocated a natural science logic in the collection of ethnographic materials and a true concern for the collection of reliable data that could lead to the production of valid theory. Yet, despite his concern for scientific method, Boas was more explicit about his methods of data analysis than about his methods of fieldwork and data collection (Ellen 1984; Boas 1920). Malinowski was also concerned with the aims of science and with methodological rigor. His earliest contributions, however, were more a demonstration of the value of ethnographic writing—his "unusual literary sense" (Lowie 1937:231)—rather than of methodological details of proper ethnographic fieldwork (Ellen 1984).

A good example of this tension between the stated early concerns for the methods of science and the actual use of such methods in ethnography comes from correspondence between Boas and his student Margaret Mead during her first fieldwork in Samoa. As Orans (1996) describes it, Mead wrote to Boas with her concerns about possible violations of scientific principles in the data she had collected to that point. She wrote of her doubts about the comparability of cases and about her ability, or even the need, to do a quantitative comparison of the similarity of attitudes among the adolescent girls in her study. She had concerns—and I believe she thought her mentor, Boas, would feel similarly—as to whether a valid comparison of this type could be made given the selection process for her sample of girls.

The constraints of field research may lead one to stray from the idealized prescriptions of a research design, but Mead was attempting to exert her authority without necessarily following the research procedures advocated by Boas and others. Orans says: "What she wants is permission to present data simply as 'illustrative material' for the representativeness of which one will simply have to take her word" (p. 127). What is most surprising is Boas's response to Mead. He writes:

> I am very decidedly of the opinion that a statistical treatment of such intricate behavior as the one that you are studying, will not have very much meaning and that the characterization of a selected number of cases must necessarily be the material with which you operate. Statistical work will require the tearing out of its natural setting, some particular aspects of behavior which, without that setting may have no meaning whatever. A complete elimination of the subjective use of the investigator is of course quite impossible in a matter of this kind but undoubtedly you will try to overcome this so far as that is all possible. (from Orans 1996:128)

This response is important for at least two reasons. First, it demonstrates the differences between the stated scientific objectives of ethnographic work as advocated by Boas and the actual practice of ethnographic research. There appears to be a perception that a systematic treatment of the data will have to be abandoned to preserve context and meaning. Ironically, this concern for context and meaning over methodological rigor, particularly for those in search of theoretical foundations (that is, the Boasian idea of data leading to the construction of theory), would ultimately hinder the comparability of data from different ethnographic sources (see Moran [1995] for a recent discussion of this issue and see Ember and Ember, this volume).

Second, Boas's concern for contextual meaning over the statistical analysis of data was prophetic. Rightly or wrongly, the preeminence of contextualization has been a consistent issue in ethnographic research and has often clouded issues in research design. The idea that quantification detracts from context and meaning in the ethnographic endeavor—evident even in the time of Boas—and a failure to understand that systematic methods—whether quantitative or qualitative—help minimize the subjectivity of the investigator have impeded the development of well-delineated research strategies in anthropology.

It's tempting to explain this as the consequence of the intensely personal nature of fieldwork, and the complexity of a holistic approach. However, this debate has its parallel in sociology where schools such as ethnomethodology and symbolic interactionism developed in response to the largely quantitative macro-level focus of the discipline. These micro-level approaches are attempts to get at a better understanding of meaning in everyday life (Cook 1994).

Boas's final sentence in his response to Mead illustrates that even at this early stage the issue of the subjectivity of ethnographic research was of concern. There was a faith, however, that awareness of the potential biases associated with the subjectivity of the investigator could be dealt with in some reasonable way. A further irony is that the one thing that might have lessened potential subjectivity biases—the use of standardized methods—was rejected outright because meaning might be compromised. Mead's position on these various elements of research design provided fuel for the continuing discussions about the validity of her original findings (Brim and Spain 1974; Freeman 1983; Orans 1996).

Thus, while early British and U.S. anthropologists advocated the scientific method in ethnographic research, there is little evidence that they considered appropriate design issues when they actually did the research. As Urry (1984) sees it:

> In Britain the claims that anthropology not only studied a distinctive body of data but also that it possesseda sophisticated methodology to collect these data, was an important factor in the establishment of anthropology as a discipline. This was less necessary in America where, by the late nineteenth century, anthropology was already established in universities, museums and government agencies. But in spite of claims to scientific methodology, particularly in the British tradition, there are surprisingly few details about actual methods anthropologists used in the field, beyond a few first principles and illustrative anecdotes. There was a wide belief among British anthropologists that fieldwork could not be taught to new recruits, but could only be experienced by individuals in the field. In the American tradition texts provided what was regarded as an objective body of data, whereas the British tradition was more a matter of subjective experience. It is a strange paradox in the development of field methods that the scientific study of other cultures has been built upon such a foundation. (p. 61)

There is much anecdotal evidence for a belief, across the British and U.S. traditions, in a trial-by-fire method of training for ethnographers. This belief supports the current lack of formal training in methods and research design in anthropology. Agar (1980) and Bernard (1994) relate stories about Kroeber's recommendations regarding the teaching and conduct of ethnographic research. In the stories, one concerning Wagley's teaching of a field methods course and one concerning a graduate student at Berkeley asking for advice before going to the field, Kroeber's response was a terse, one liner that reflected the attitude of the times. Even in the late 1960s, when concern for methodological rigor was probably at its peak in anthropology, many treatments of research methods and design in the literature played down the need for more systematic methods and design detail, particularly with respect to hypothesis-testing approaches (LeVine 1973). A good example of this is a book by Thomas Rhys Williams (1967) published in the Spindlers's series on field methods. Williams writes:

> I believe that only someone wholly involved and fully immersed in fieldwork can really communicate the essence of cultural anthropology to students or general readers. And since I have indicated here that research in culture involves a great deal of unique personal experience for the anthropologist, I have taken the position that it is probably unlikely there can be a rigorous, systematic, and formal presentation of methods in the study of culture like those of the natural sciences and that there are overriding concerns among many sociologists, psychologists, and economists. I find this stance comfortable, for it is my conviction that so long as prime theoretical concerns in the study of culture are an attempt to record and understand the native's view of his culture and the objective and historical realities of culture, then methods for field study will have to reflect the end purpose of making a whole account of a part of the human experience. (pp. 64–65)

LeVine (1973) and others (Johnson 1990) make the point that the nature of fieldwork, in terms of its requisite huge investments in time and geographical focus, has often limited the attractiveness of more formal research designs because of its commitment to studying specific problems in a specific way. The realities of fieldwork often dictate the need to change the problem focus or, finding that the proposed hypotheses are inappropriate to the cultural setting under study, the need to somehow salvage the research.

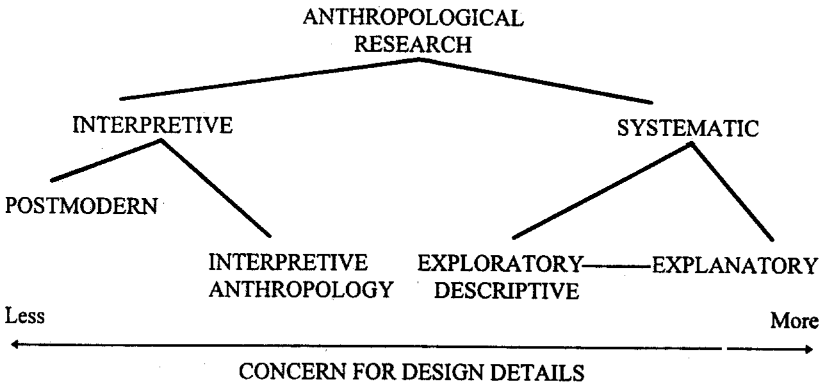
Laboratory and survey researchers have some flexibility to change the problem focus and study populations in light of emerging problems, but field workers are limited in their ability to do so. Thus, the idea of researchers "putting all their eggs in one basket" may have limited the a priori formulation of problems in fieldwork (LeVine 1973:184). Further, the huge investment in time and resources limited another important goal of science, that of replication, since an ethnographer couldn't realistically be expected to replicate someone else's work. The "my natives" or "my village" mentality of some and the fact that careers were made by discovering new theories or describing exotic less well-known cultures has certainly inhibited replication efforts (Johnson 1990).

Contemporary Design Issues in Cultural Anthropology

There is an ongoing debate in cultural anthropology concerning science and its role in contemporary research. A discussion of the basic arguments as related to epistemology, objectivity, reality, authority, and the like are beyond the scope of this chapter (see Schweizer in this volume). Suffice to say that traditionally, research design and its logic have been associated with science and an underlying belief in objectivity and explanation. The historical tension between interpretive and scientific approaches in anthropology has given way to an outright rejection by some anthropologists of science and its logic of design. To say that the research design logic of science has been replaced by something that is recognizable as the research design logic of, say, postmodernism would, I think, be misleading. It is not that interpretive approaches lack some form of research plan; but the term "design" itself smacks of the very formalism that is being rejected. A more appropriate term that would encompass the diversity currently found in cultural anthropology might be "research strategy."

Figure 1 is a taxonomic characterization of the different types of research strategies found in contemporary cultural anthropology. The figure distinguishes between strategies within the realm of interpretive studies and those using systematic strategies that have more of the elements of science. This is a highly simplified representation. Many examples of research in anthropology fall within the two extremes of the continuum. Under the systematic distinction are the two primary categories of exploratory and explanatory approaches, each entailing a specific design strategy. The light line connecting the two categories indicates their complementarity and interrelatedness in that a design may include both within an overall research design framework. These approaches are by no means mutually exclusive in approaching a research problem (see section on Research Design in Systematic Research, below).

In its most extreme form, systematic strategies tend to involve the search for explanations of phenomena and the pursuit of theoretical foundations. In searching for such foundations, there is a need for objectivity, replication, and control over possible sources of error leading to a valid assessment of a given theory. Epistemologically, systematic work is objectivist. Its practitioners are ultimately interested in research findings that approximate an external truth. As a result, the assessment of any theory involves research designs more heavily concerned with the means—the research process, rather than simply the way the study was written or argued—since the validity of study results depends on the scientific soundness of the research design. For any given research problem, it is the purpose of research design to ward off as many threats to validity as possible. This leads to designs that involve concern for a higher degree of methodological and analytical detail, whether quantitative or qualitative. In this line of thinking, the researcher is a field-worker-as-writer.



IN SEARCH OF: UNDERSTANDING MORAL TALES LOCAL RATIONALES	IN SEARCH OF: RULING OUT RIVAL HYPOTHESES OBJECTIVITY POSSIBLE SOURCES OF ERROR REPLICATION THEORETICAL FOUNDATIONS
PURPOSE: A BELIEVABLE ACCOUNT OR STORY	PURPOSE: VALID ASSESSMENT
EPISTEMOLOGY: SUBJECTIVIST VALUE MEDIATED FINDINGS OR CREATED FINDINGS	EPISTEMOLOGY: OBJECTIVIST FINDINGS APPROX. THE TRUTH
CONCERN FOR: THREATS TO BELIEVABILITY	CONCERN FOR: THREATS TO VALIDITY
MEANS FOR DIMINISHING THREATS: LITERARY	MEANS FOR DIMINISHING THREATS: METHODOLOGICAL

Exploratory: Exploratory approaches are used to develop hypotheses and more generally to make probes for circumscription, description, and interpretation of less well-understood topics. This is similar to the grounded theory ideas of Glaser and Strauss (1967), where exploratory descriptive research leads to the development of more meaningful theory and measures. Exploratory research can be the primary focus of a given design or just one of many components.

Explanatory: Explanatory approaches generally involve testing elements of theory that may already have been proposed in the literature or that have been informed by exploratory research. Research designs in this mode are determined a priori and their primary purpose is to eliminate threats to validity, where validity is concerned with whether things are what they appear to be or are the best approximation to the truth (Cook and Campbell 1979). In this enterprise, explanation can involve a general search for causality or prediction.

Figure 1. Types of anthropological research strategies and their features.

Interpretive strategies, on the other hand, differ from systematic approaches in that they question a researcher's ability to maintain objectivity, particularly in the ethnographic context where the ethnographer is often the instrument of measurement. A variety of names are used in the lexicon of social scientists that can be associated to varying degrees with an interpretive strategy. Phenomenology, hermeneutics, symbolic anthropology, interpretive anthropology, interpretive interactionism, deconstructionism, postmodernism, and constructivism, to name a few, question, in one way or another, some or all of the ontology, epistemology, and methodology of systematic approaches. Although some of the older interpretive strategies that emerged from the scientific tradition in the social sciences, such as early interpretive anthropology, still adhered to some logical empiricist methodology and maintained a degree of belief in ethnographic authority, more recent approaches, such as postmodernism and constructivism, are more radical in their sweeping rejection of scientific method and design logic (see Schwandt 1994). In contrasting Geertz and early interpretive anthropology with some of the later postmodern turns of such ethnographic writers as James Clifford, Rabinow (1986) observes:

> At first glance James Clifford's work, like that of others in this volume, seems to follow naturally in the wake of Geertz's interpretive turn. There is, however, a major difference. Geertz (like the other anthropologists) is still directing his efforts to reinvent an anthropological science with the help of textual mediations. The core activity is still social description of the other, however modified by new conceptions of discourse, author, or text. The other for Clifford is the anthropological representation of the other. This means that Clifford is simultaneously more firmly in control of his project and more parasitical. He can invent his questions with few constraints; he must constantly feed off others' texts. (p. 242)

There is a fundamental belief that the intersubjective, everyday meanings and how they are produced, maintained, and changed in any given context often defy objective study and explanation. Practitioners of almost all interpretive paradigms are searching in one way or another for some understanding (*verstehen*) rather than for some explanation of social phenomena. However, some interpretive work is more similar in nature to the exploratory or descriptive strategies found under the systematic side of Figure 1 than to some of the more radical forays into, for example, postmodernism. Thus, the rather simple characterization of research strategies found in Figure 1 attempts to recognize the variation inherent in the range of work found in contemporary anthropology by placing "interpretive anthropology" adjacent to "exploratory/descriptive" (see, for example, the work of Zabusky 1995). Discussions about this debate can be found in Seidman (1994), on the one hand, and Faia (1993), on the other, and, more specifically for anthropology, by Kuznar (1997).

An important implication here is that scholars who follow this line of inquiry are searching for local rationales rather than nomothetic theory or universal foundations and may be more interested in conveying a moral tale of some type rather than a value-free account (Seidman 1994). Further, the purpose of research strategies under these interpretive paradigms is more focused on the production of a believable or plausible account or story rather than a single depiction of the truth, since it is thought that there are a multitude of plausible accounts rather than just a single true story. Epistemologically, interpretive paradigms are subjective, with findings that are value mediated or even created. Thus, there is less focus on the means of research, such as methods of data collection and analysis as found in the systematic strategies, and more on the ends of research—the ethnographic or literary product. In contrast to the field-worker-as-writer, we find the writer-as-field-worker (Denzin and Lincoln 1994).

For scholars like Geertz, analysis of ethnography has less to do with the methods of observation and description than the inscriptions and writings concerning the meaning of human action. In many ways, this blurs the distinction between what is anthropological and what is literary. More extreme forays into experimental ethnography have blurred this distinction even further, and there is more of a focus on writing strategies that include such approaches as montages, evocative representations, polyvocal texts, and even ethnographic fictions (Denzin and Lincoln 1994). While systematic analytical paradigms are primarily concerned with threats to validity, recent interpretive paradigms are focused more on threats to believability —as in "Do you believe my story?" (Tyler 1991:85)—or, in critical theory, threats to trustworthiness (Kincheloe and McLaren 1994). If we talk of an interpretive method, particularly with regard to postmodernism, it more than likely involves both the researcher's immersion into the cultural context of the actor(s) and some means, usually literary, for conveying the understanding gained from such an immersion.

As stated, many interpretive studies are closer in character to exploratory and descriptive research in the systematic mode than to some of the more extreme postmodern studies. A good example of this is Zabusky's (1995) ethnographic study of cooperation in European space science that she admits "took the form of mutual exploration rather than unidirectional examination" (p. 46). She contrasts her study with research on cooperation by "experimental" psychologists, emphasizing the cultural and social orientation of her work and the importance of considering context (social, cultural, political, etc.) in her analysis. Following in the "thick description" tradition of Geertz, Zabusky clearly believes in some kind of ethnographic authority. In a short methodology section, she discusses the challenge of conducting participant observation research in this rather complex, geographically dispersed, cross-cultural setting. She also discusses the rationales for selecting the site and the group she studied, problems of working in a linguistically and technically diverse social milieu, the use of semistructured and unstructured interviews, and the effect of her role as ethnographer on informant relations and data quality. Although Zabusky doesn't talk specifically about design or about concerns for potential threats to validity, there is implicit concern for such issues throughout the ethnography.

In contrast to Zabusky, there is a body of interpretive work in anthropology that is more extreme in its rejection of systematic design issues. Ramos (1995), for example, has recently published an ethnography based on a rewrite of her 1972 dissertation, with additional ethnographic insights. She rejects the "anthropological austerity" of her original work in favor of an "intersubjective understanding" that captures the "flavor" of her ethnographic encounter with the Yanomami. To her, the original work was "old-fashioned and theoretically unsophisticated" and had to be replaced by a more reflexive work. This contrast between the old and the new reflects the increased variation in epistemological emphasis in the field that has developed over the last 30 years. As Ramos sees it, "I found myself making forays into the self-conscious meanderings of reflexive anthropology in order to shift the axis of analysis from the skeletonlike dissertation to the flesh and blood of ethnography" (p. 6).

Along with this shift came the freedom not to be concerned with issues of bias and validity or with the need for working systematically, thus allowing for a less restrictive ethnographic narrative. Although Ramos discusses informant interviewing and various sources of data, her introduction is largely devoted to discussions of her reliance on her own memory in writing the ethnography and the shift in the narrative between synchrony and diachrony. Thus, there is little discussion of research design and methods of data collection as might be found in work in the systematic tradition. Instead, Ramos emphasizes the emergent and reflexive nature of data and the literary strategies used in producing the ethnographic product. Other examples in this vein include Panourgia's (1995) use of *we* and *they* in her "Athenian Anthropography" and Behar's (1993) use of montage in her collaboration with a single woman in the telling of that woman's life story. Behar discusses the multiplexity of roles, in that she was variously involved as "priest, interviewer, collector, transcriber, translator, analyst, academic, connoisseur, editor, and peddler" (p. 12).

The idea of a montage as an organizing principle was also central to Taussig's (1987) historical and ethnographic account of shamanism, colonialism, and terror in South America. This work is important in at least two ways. First, it is representative of the genre that rejects explanation in favor of conveying a moral tale. Its purpose is not a traditional attempt at explanation where facts are considered real, but political interpretation and representation of facts, independent of their "realness." Second, Taussig uses the "principle of montage" as a means, at least in his view, for better relating the lessons of history. As he states:

> As against the magic of academic rituals of explanation which, their alchemical promise of yielding system from chaos, do nothing to ruffle the placid surface of this natural order, I choose to work with a different conflation of modernism and the primitivism it conjures into life—namely the carrying over into history of the principle of montage, as I learned that principle not only from terror, but from Putumayo shamanism with its adroit, albeit unconscious, use of the magic of history and its healing power. (p. xiv)

These examples offer only a brief glimpse of the range of possible strategies in use by interpretivists in anthropology. For some, interpretive work is an exploratory enterprise with an implicit concern for methodological issues. For others, interpretive work is concerned more with the strategies and methods of ethnographic presentation and with the reflexive character of the ethnographic enterprise. Thus, traditional methods sections are replaced by discussions on how to read the work or on the particular methods used in writing the ethnography itself (see, for example, Panourgia's discussion on the use of the *parerga*).

In the following pages, I focus primarily on research designs in systematic research. For further discussion of research strategies in the interpretive mode, see Fernandez and Herzfeld (this volume).

Research Design in Systematic Research: The Challenge of Making a Case

In some social science disciplines, like psychology, the design of research is driven by features of the analysis. Analysis-of-variance models and multigroup comparisons (factorial designs) may dictate the whos, whats, and wheres of a given project. In sociology, multiple regression models, structural equation models, and path analytic models (all related analytical techniques) have influenced the design of survey research. Ethnography, referred to as the anthropological method by William Foote Whyte (1984), has influenced the nature of design in anthropology, but in profoundly different ways.

Whereas the analytical techniques most often used in psychology, sociology, and economics often led to rather standard designs, in anthropology the eclectic nature of ethnography leaves the design of research more open ended. There are generally no ethnographic "analytical techniques" driving the design, although ethnography has been variously associated with a number of qualitative methods. The good news is that ethnographic research is amenable to a wide range of research designs, including the use of multiple designs within a single ethnographic context. This allows for flexibility, multiple tests of a theory, increased chances for various types of validity, triangulation, and the potential for high levels of innovation and creativity. The bad news is that the open-ended character of ethnography contributes to a less well-focused discussion of research design issues in ethnographic approaches.

Part of the confusion stems from a lack of consensus on what ethnography really is (Johnson 1990). To some, it is both a process and a product (Van Maanen 1988). Although this process might be equated to a method, it's better to think of ethnography as a strategy in which a variety of methods can be used in the quest for knowledge (Pelto and Pelto 1978). Thus, ethnography should involve multiple methods, both qualitative and quantitative, and may involve applying more than one research design. This is particularly true today, given the large number of computer analytical packages available for analyzing text (see Bernard and Ryan, this volume). Currently, the qualitative analysis of text and discourse is no longer restricted to either interpretive or exploratory approaches, but can also be used in hypothesis testing and explanatory research.

Figure 2 illustrates the relationship between exploratory and explanatory approaches within the ethnographic context. This contrast between explanatory and descriptive or exploratory approaches is commonly made in nonexperimental disciplines in both the natural and social sciences. Community ecologists, for example, similarly distinguish between exploratory or descriptive studies that seek to describe and determine patterns in ecological data and those studies that specifically seek to predict or test hypotheses. As with research in community ecology, ethnographic research can be purely exploratory or descriptive—involving a research process focused on producing better theory—or purely explanatory, although this is usually not the case. Rather, the most common model has exploratory research informing and complementing explanatory research. As we will see in the examples to come, exploratory research is often an essential component of the explanatory research process. Exploratory research may contribute to the production of reliable and valid measures, provide information essential for constructing comparison groups, facilitate construction of structured questions or questionnaires, or provide information necessary for producing a sound probability or nonprobability sample.

The figure shows that the overall research process is more than just a matter of study design. There is no substitute for a good theory, and there is a critical need to link theory, design, data collection, analysis, and interpretation in a coherent fashion. Design, however, is the foundation of good research. No amount of sophisticated statistics, computer intensive text analysis, or elegant writing can salvage a poorly designed study. Hurlbert (1984) emphasizes this in a classic paper on the design of field experiments in ecology. "Statistical analysis and interpretation," he says, "are the least critical aspects of experimentation, in that if purely statistical or interpretive errors are made, the data can be reanalyzed. On the other hand, the only complete remedy for design or execution errors is repetition of the experiment" (p. 189). Redoing an experiment because of fundamental design errors is one matter; redoing a year-long ethnographic field study because of such errors is quite another.

Figure 2 shows that the research process involves a simultaneous concern for the development of empirical statements from theory (for example, hypotheses), the operationalization of theoretical concepts (for example, meaningful and reliable measures), design (for example, groups to be studied), data collection (for example, qualitative versus quantitative), and data analysis (for example, multiple regression and text analysis). Theoretical knowledge is derived either from earlier studies or from exploratory work. The levels at which theoretical concepts are measured (for example, nominal or ordinal), the types of sampling strategies used, and the application of appropriate types of analysis must all be considered as a part of the design. For example, the particular structure of an empirical statement or hypothesis will partially determine the manner in which theoretical concepts are operationalized and eventually analyzed. (Stinchcombe [1987] provides an excellent discussion of how empirical statements are derived from theory.)

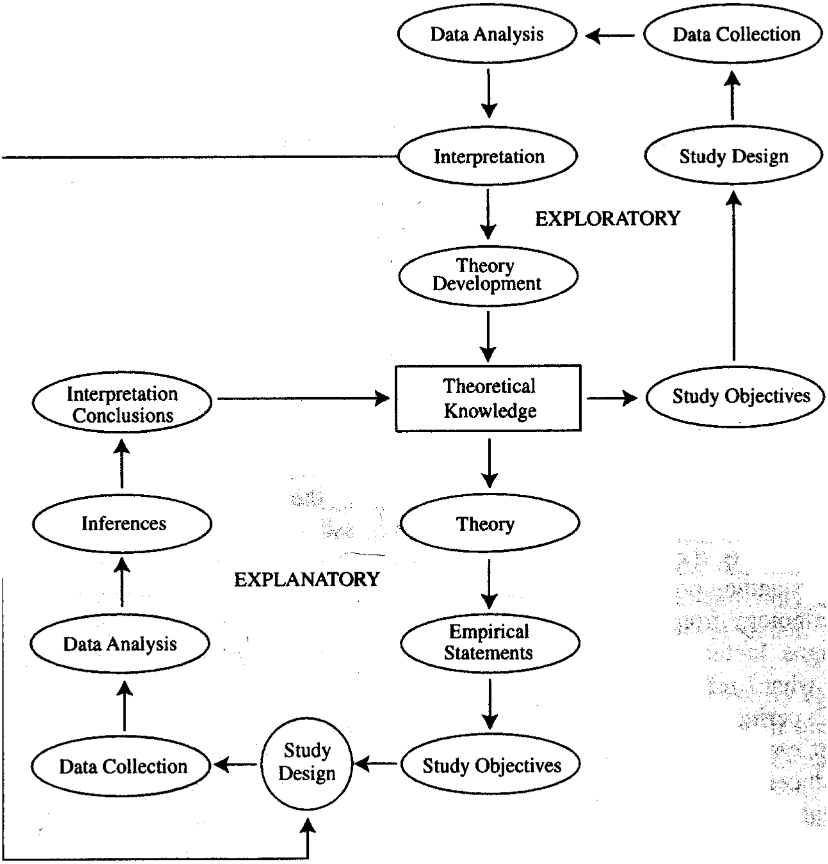


Figure 2. Relationship between exploratory and explanatory approaches within the overall ethnographic research process.

Thus, research design is more than just methods of data collection and analysis. It involves constructing a logical plan that links all the elements of research together so as to produce the most valid assessment possible of some theory, given some set of realistic constraints (for example, cost, scope, geographical setting, etc.). The purpose of research design is to ward off as many threats to validity as possible and to help one eliminate competing hypotheses. It requires careful attention to detail and, often, an admission concerning the potential weakness of a given design. Outside the laboratory, a multitude of influences can threaten the validity of any conclusions. In natural settings, particularly fieldwork, there is no perfect design that can control for all possible extraneous effects at once. A recognition of limitations doesn't invalidate a study's results. Rather it creates an open forum that can contribute much to important theoretical and methodological debates. Without such attention to good design and methodological detail, researchers leave themselves open to one of the worst criticisms of all—of being "not even wrong" (Orans 1996). In other words, a lack of design and methodological detail makes it next to impossible to fairly and adequately assess the validity of any study's conclusions such that "rightness" or "wrongness" may not even be debatable.

True experiments involve random assignment and afford the best chances for controlling for things like: the effects of extraneous factors (that is, unmeasured variables that might affect the dependent variable); the effects of selection (that is, comparison groups differ because of the way they were selected and not due to the treatment); the effects of reactive measurement (that is, the measurement procedure itself caused a change in the dependent variable); or interaction effects involving selection (that is, when selection interacts with other factors to create erroneous findings). These and other sources of error are all potential rival hypotheses and randomized experiments are best at eliminating the threats of rival explanations. Designs of this type, however, are often impossible in anthropological fieldwork. Nevertheless, the principles of experimentation are instructive and are a guide for understanding potential sources of error, even in a nonlaboratory setting. I borrow terminology from Kleinbaum et al. (1982) in constructing a typology of research designs. Included are *experiments, quasi-experiments*, *observational study designs,* and what I refer to as *natural experiments.*

Experiments involve the random allocation of subjects to groups and afford the most control over distorting effects from extraneous factors. Random allocation produces equivalent comparison groups, and artificial manipulation of independent variables (also known as explanatory variables or study factors), with all other variables or factors controlled for, allows for the most valid assessment of the causal relationship between the independent and dependent variables or response variables. What separates quasi-experiments from true experiments is the lack of random assignment of group members. Random assignment maximizes the probability that experimental groups are equivalent on key variables prior to the introduction of an intervention. Nonrandom assignment lays an experiment open to validity threats and reduces our ability to make causal inferences. Observational studies involve neither random assignment of members to comparison groups nor the manipulation by the observer of independent variables.

This distinction between experimental and observational approaches is similar to one in ecological field studies. Hurlbert (1984) distinguishes between two classes of experiments. He terms the first *manipulative experiments.* These are basically true experiments involving random assignment, multiple comparisons (for example, treatment versus control), and the manipulation of independent variables. He refers to the second as *mensurative experiments,* which involve simply the measurement of variables in space and time and among a number of comparison groups, without random allocation and the manipulation of experimental factors.

The primary distinction lies between that of *sampling* versus *allocation.* In manipulative experiments, analytical units are randomly allocated to comparative groups, whereas in mensurative experiments selection of units is based on some probability or nonprobability sampling scheme. While random assignment aids in controlling for confounding variables by producing homogeneous comparative groups, random sampling of units produces comparison groups that are representative of such groups. Random sampling meets the restrictions of some statistical tests, but it does not afford the same protection as does random assignment of group members against the potential effects of extraneous factors. Mensurative designs, then, are observational and characteristic of the types of comparative designs found in field studies in anthropology.

Finally, natural experiments are similar to quasi-experiments except that the manipulation of independent variables occurs naturally or is unplanned rather than artificial or directed. Thus, comparison groups may be chosen on the basis of different levels of exposure to some naturally occurring or human-induced phenomena (for example, natural disaster, war, or the building of a dam). Cook and Campbell (1979) make a similar distinction but refer to these kinds of natural experiments as "passive-observational studies." Anthropologists involved in development and evaluation research are most likely to use this design.

True experiments are, of course, rare in anthropology (but see Harris et al. [1993] for an example of a true experiment in a field setting). Even in quasi-experiments, it's often difficult to manipulate independent variables directly. Howevert, with careful attention to design and ethnographic context, quasi-experimental and natural experimental designs can be applied to anthropological field settings, particularly in evaluation research and development research. Johnson and Murray (1997), for example, used a quasi-experimental design to evaluate the use of fish aggregation devices (FADS) in small-scale fisheries development projects. Two fixed fishing structures (piers) were pretested for differences in catch rates. Then, FADS, umbrella-like units suspended in the water column, were alternately placed at the piers and individual fishers were interviewed simultaneously during randomly selected times at both the treatment (the pier with the FADS) and the control (the pier without the FADS) piers. Johnson and Murray compared and determined catch rates.

From a statistical standpoint, designs that don't involve random assignment—including quasi-experiments—are considered observational (Cook and Campbell 1979). It is important, though, to contrast quasi-experiments to what Kleinbaum et al. (1982) refer to as observational studies. The most common designs used traditionally by anthropologists have been observational in nature. Designs of this type lack direct control over independent variables and, thus, have more potential problems with various types of internal validity and with the ability to assess time order effects and causality. However, if done properly, such designs can have increased external validity and generalizability.

Due to their predominance in anthropology, the examples that follow are comparative observational designs. Most research designs in the explanatory mode, like true experimental designs, are comparative (for example, control versus treatment). Table 1 describes examples from observational and quasi-experimental study designs discussed by Kleinbaum et al. (1982) and Cook and Campbell (1979). More details can be found in these and other sources (for example, Robson 1993). In anthropological fieldwork, these designs and others can be used in tandem to test or explore components of a theory (such as combinations of time series and repeated measures designs particularly applicable to long-term fieldwork). For example, in their study of preschool children, Johnson et al. (1997) used a *cross-sequential design*, which involved cross-sectional research on a cohort of children carried out over time.

When one is interested in explanation, the importance of comparative thinking in ethnographic work cannot be overemphasized. Discussing "common sense knowing" in evaluation research, Campbell (1988) gives an important critique of ethnography. His idea is that "to know is to compare" is fundamental to explanatory work in anthropology:

> The anthropologists have never studied a school system before. They have been hired after (or just as) the experimental program has got under way, and are inevitably studying a mixture of the old and the new under conditions in which it is easy to make the mistake of attributing to the program results which would have been there anyway. It would help in this if the anthropologists were to spend half of their time studying another school that was similar, except for the new experimental program. This has apparently not been considered. It would also help if the anthropologists were to study the school for a year or two prior to the program evaluation. (This would be hard to schedule, but we might regard the current school ethnographies as prestudies for new innovations still to come.)
>
> *All knowing is comparative*, however phenomenally absolute it appears, and an anthropologist is usually in a very poor position for valid comparison, as their own student experience and their secondhand knowledge of schools involve such different perspectives as to be of little comparative use. (p. 372; emphasis added)

While the purpose of experimental design is to ward off as threats to validity, there are several types of validity—face, construct, statistical conclusion, internal,

TABLE 1

Examples of Basic Research Designs Relevant to Anthropologists

<u>Observational Designs</u>

Cohort Study

Design: Often referred to as a panel study, this is a longitudinal design where individuals are followed through time. May involve comparison groups subjected to different treatments or exposed to different conditions.

Cross-Sectional Study

Design: Often referred to as a survey study, it generally involves a random sample of a target population. Stratified sampling is often used to ensure adequate sampling of comparison groups. Although study factors are not controlled directly, designs of this type allow for the statistical control of variables during analysis.

Case-Control Study

Design: For some study factor (like an outcome variable), compares a group of cases in which members have some characteristic of interest with one or more groups in which the characteristic of interest is absent. It is assumed that both groups come from the same underlying population. Often, members of the groups are matched on one or more variables.

Static-Group Comparison

Design: A variant of the cross-sectional design in which a treatment group(s) (that is, members exposed to some variable of interest) is compared with a comparison or control group whose members are not exposed to the variable of interest.

<u>Quasi-Experimental Designs</u>

One group posttest only design

Design: Pretest observations are made on a single group. The group receives a treatment of some type and posttest observations are made.

Posttest only nonequivalent groups design

Design: Experimental and comparison or control group are determined without random allocation of group members. Experimental group receives treatment while the control group does not. Posttest observations are made and groups are compared.

Pretest/posttest nonequivalent groups design

Design: Experimental and comparison or control group is determined without random allocation of group members. Pretest observations are made on both groups. Experimental group gets the treatment while control group does not. Posttest observations are made and groups are compared.

Interrupted time series design

Design: One experimental group in which a series of observations is made both prior to some treatment and after the treatment.

external, etc. In one way or another, various study designs, in combination with other considerations such as the operationalization of theoretical constructs and sampling, are better or worse at dealing with each. Here, I stress the importance of thinking through how validity threats have influenced and will influence observations or data (for a more in-depth discussion of how these types of validity can impact study conclusions, see Cook and Campbell 1979). Potential errors and bias creep in at various steps in the research process. It's your job to contain these errors. In research design, forewarned is forearmed.

Tables 2 and 3 give examples of threats to internal and external validity as discussed in Cook and Campbell (1979) for quasi-experimental designs. Internal validity is concerned with the approximation to the truth within the research setting. External validity is concerned with the approximation to the truth as expanded to other settings—that is, with the generalizability of research findings. The threats in Table 2 deal with extraneous factors that may account for the presence or absence of a hypothesized effect (that is, contrast validity with invalidity). In the quasi-experimental case, this means changes between pre- and posttest, but this way of thinking can be expanded to include hypothesized effects dealing with differences, similarities, or associations whether diachronic or synchronic.

TABLE 2
Threats to Internal Validity in Quasi-Experimental Designs

History—Change due to unmeasured or unobserved factors
Testing—Change resulting from experience gained by subjects as a consequence of measurement
Instrumentation—Change resulting from varying the way study participants are tested
Regression—When selection of participants are atypical or extreme on a given measure, subsequent measures will become less extreme and there will be regression toward the mean
Mortality—Changes due to participants dropping out of the study
Maturation—Change in study participants over time due to factors unrelated to expected effects
Selection—Observed effects due to nonrandom assignment of members and nonequivalence of groups
Selection by Maturation Interaction—Predisposition of selected group members to grow apart
Ambiguity about Causal Direction—When time-order and causal direction is ambiguous
Diffusion of Treatment—Change due to one group receiving all or a portion of treatment meant for another group
Compensatory Equalization of Treatments—Tendency toward giving all groups the same treatment
Compensatory Rivalry—Participants' perceptions (for example, threats) that affects performance not a part of the treatment

TABLE 3
Threats to External Validity

Selection—Problems with generalizing due to the selection process for study subjects (e.g., nonrepresentative)
Setting—Problems with generalizing due to the nature of the study setting (e.g., setting atypical)
History—Problems with generalizing to either the past or the future

Cook and Cambell (1979) detail how each of the quasi-experimental designs in Table 1 are better or worse at dealing with each of the threats to validity that are found in Tables 2 and 3. For example, the pretest/posttest nonequivalent groups design controls for some internal threats to validity, but it's problematic with respect to controlling for changes due to how groups members were selected (selection maturation), changes due to how individuals were tested (instrumentation), changes due to the selection of individuals with extreme pretest measures leading to regression toward the mean (regression), and changes due to local events not a part of the study (history). Each of these threats may hamper a researcher's ability to assess the contribution of a hypothesized effect to any changes observed. Similarly, threats to external validity, such as problems stemming from biased samples or research in atypical or unique settings, can hamper the generalizability of one's findings. Kleinbaum et al. (1982) offer a similar discussion of the strengths and weaknesses of observational designs in terms of controlling for threats to both internal and external validity.

Other sources of potential bias include sampling error (that is, chance), non-response, the use of imprecise measures, data recording errors, informant inaccuracies, and interviewer effects (see Pelto and Pelto 1978; Bernard 1994). Careful attention to sampling, whether probabilistic (Babbie 1990) or nonprobabilistic (Johnson 1990), is essential. Measurement, operationalization of theoretical concepts, and type of analysis used are other important factors. How reliable are your measures in terms of precision, sensitivity, resolution, and consistency? Are they valid, particularly with respect to accuracy and specificity, in that they are actually measuring what they are intended to measure? Attention and concern with all the potential sources of error, whether stemming from how the study was designed, how the data were collected (for example, face-to-face interviews or mail-out surveys), or how the data were analyzed (for example, statistical conclusion validity), will help lead to the production of solid evidence.

Some Comments on Sampling

Many probability and nonprobability sampling designs are available for any given research problem. These include systematic sampling, stratified random sampling, cluster sampling, and multistage sampling. The selection of any of these designs or the development of some hybrid design depends on the overall design of the research itself. The nature of the groups or characteristics to be compared—in terms of such things as the size of the comparison groups in the overall population, the frequency of characteristics of interest in the population, the availability of a sampling frame, the ability to identify members of the population (for example, hidden or clandestine populations)—all influence the choice of a sample design. But it's not always easy to know who or what you want to sample and to know enough about these sampling units to derive a valid sample.

The selection of units of analysis, whether settings, events, times, households, or people, is important for understanding a variety of internal and external threats to validity, but it is particularly important for increasing external validity. We mostly think of selection in terms of some type of sample units. To generalize to a target population, the sample has to be representative of the population of interest. This is essential if we are to generalize to a whole population and is generally, though not always, a requirement for classical statistical tests.

When generalization to a target population is the objective, you should strive to define a sampling universe or frame using a selection procedure with known error limits and one that represents the population of interest. This usually entails a random sample of some kind. There is a vast literature on sampling theory and random sampling procedures, including discussions of sample sizes (see, for example, Bernard [1994] for a summary and Babbie [1990] for detailed discussion of sampling issues).

Cook and Campbell (1979) discuss two sampling models for increasing external validity in quasi-experiments. These models don't necessarily involve random selection and are consequently less powerful than are random samples. In one approach, the *model of deliberate sampling for heterogeneity,* target classes of units, whether classes or categories of persons, places, times, or events, are deliberately chosen to represent the range of such classes found in the population. Thus, testing for a treatment effect across a wide range of classes in the set of all possible classes (including both extremes and the modal class) in the population allows the researcher to say something about how the effect holds in a variety of settings. While this might not be generalized to the population as a whole, it does inform the researcher if an effect holds across wide ranging classes within the population. The logic behind this model can be extended beyond the quasi-experimental case to observational studies. Kempton et al. (1996) used a static-group comparative design sampling across a range of groups that varied with respect to their values on environmental issues. Kempton et al. interviewed members of Earth First (a radical environmentalist group) and dry cleaning shop owners (who depend on toxic chemicals for their business).

For some populations, it may be impossible to develop a sampling frame from which to draw a sample. In these cases, there are a variety of solutions, including intercept sampling, snowball sampling, random walks, quota sampling, and purposive sampling. Each of these approaches has potential problems, and most do not allow for generalizations about a population since they involve elements of unknown error even if the method involves some form of random selection criteria (for example, random selection of locations in which to intercept respondents).

Nonprobability sampling methods have come to be associated with qualitative approaches or for the selection of ethnographic informants, particularly key informants or consultants (Werner and Schoepfle 1987; Johnson 1990; Miles and

Huberman 1994). In some cases, a researcher may not be interested in generalizing to a population but may just want to know whether two subgroups obtained from a snowball sample differ with respect to some variable of interest. In that case, much of the bias in the sample is a matter of the logic used in the original selection of sample seeds and any statistical analysis of the data must be concerned about violations of assumptions for the particular statistical test to be employed (for example, independence of observations or random sample from a population). Such matters are particularly germane for observational designs using various social network approaches (see Johnson [1994] for a review).

How samples are chosen is an important element of any research design. If you are interested in generalizing to a given population, random sampling of some kind is essential. If generalization is not a primary goal, then sampling requirements may be relaxed. In most cases, if you can use a random sample, do it! No matter what the sampling method, you should be *explicit* about how you chose the sampling units. This increases the chances of detecting potential bias and also makes replication feasible. Replication is extremely important to external and other types of validity, such as construct validity. Random sampling has been a primary requirement in the proper application of parametric statistics. If you don't use random sampling, pay careful consideration to possible violations of assumptions for a given statistical test.

Recent developments in randomization and computer-intensive methods of statistical analysis involve less restrictive assumptions concerning the data (for example, assumption of a random sample from a population or skewed, sparse, or small sample sizes), opening the way for the development of new test statistics particularly suited for the problem at hand (Noreen 1989; Johnson and Murray 1997). These new approaches seem particularly well suited for the imperfect world of ethnographic research, where the rather restrictive assumptions of parametric analysis are often difficult to meet. But it is critical to remember the connection between theory, design (including sampling), and data analysis from the beginning, because how the data were collected, both in terms of measurement and sampling, is directly related to how they can be analyzed. The next section shows how concern for the elimination of potential errors and bias through design and attention to methodological detail applies to discussions about the findings of Margaret Mead and Derek Freeman in Samoa.

Mead Versus Freeman: Research Design as Mediator

Derek Freeman's (1983) criticism of Margaret Mead's work and her findings in Samoa has led to reactions from anthropologists who come from different epistemological traditions. Some have defended Mead (Shankman 1996); others have pointed to the biases and flaws in Freeman's argument (Marcus 1983; Ember 1985).

The criticisms and counter-criticisms are difficult to assess, given the time between Mead's and Freeman's studies, the differences in locations of their work, and the differences in their ideological positions (Ember 1985). Freeman contended that some of Mead's informants lied to her and that Mead's commitment to a particular ideological position caused her to evaluate evidence incorrectly. We certainly cannot hold Mead to the design standards available today. Still, it is instructive to review her work through a contemporary design lens, noting how slight modifications in design and method could have thwarted later criticisms.

Mead used what can be referred to as a static group comparison design with a conjectural treatment group. The comparison group, Samoan adolescent girls, was compared to a conjectural treatment group, American adolescent girls, to test the proposition that exposure to Western civilization increases adolescent trauma. Implicit in this proposition is the overall theoretical notion that culture is the major factor contributing to human behavior. Brim and Spain (1974) recognized several problems in the design that could have affected Mead's ability to draw valid conclusions.

> There were no equivalent measurement procedures for the two groups. In her use of a conjectural treatment group, Mead assumed some things about American adolescents without collecting comparable data. Mead relied mostly on herself as an instrument to measure the variables of interest.
> There were possible problems with interaction between selection and the effects of extraneous variables. That is, any observed difference between the two groups with respect to the dependent variable, adolescent trauma, might have been due to one or several extraneous (unmeasured) factors and might have had nothing to do with the independent variable, exposure to Western culture.
> In lieu of the between-culture comparisons, Mead could have made a within-case comparison that would have suffered less from problems with possible sources of error. She could have chosen comparison groups that were as similar as possible in order to rule out the effects of unmeasured variables as much as possible. For example, Mead could have compared girls living in the households of native pastors to those who did not. She could then have tested the proposition that exposure to competing standards of sexual morality leads to higher levels of emotional distress in adolescents.

More recently, Martin Orans did fieldwork in Samoa. Some of his experiences were incongruent with Mead's descriptions. Orans (1996) reanalyzed Mead's field notes and correspondence and once again found that her depiction of Samoa as a halcyon society was at odds with his own impression of Samoa as much more agonistic. If Samoa was not, during Mead's day, a halcyon society, then her conclusions might have been flawed. Orans's work was, of course, many years after Mead's, and he worked at different field sites than did Mead. But, in common with Brim and Spain, Orans found itemized problems with Mead's research design.

- There was a lack of comparisons between various sources of data that were crucial to Mead's argument. For example, Mead made an assertion concerning the relationship between the size of residential units and adolescent troubles. She did not, however, make any systematic comparisons among the different units. Similar to the observation by Brim and Spain, Orans points out that Mead made no comparison of sexual behavior between girls living in a native pastor's household and girls living with their own family.
- There is a lack of well-defined samples for both people and events. Mead makes assertions about the rarity of events without any knowledge of the frequency distribution of all such events. In addition, she had a tendency to understate the population and overstate the proportion of girls in her study.
- There is a lack of specificity in the development and operationalization of key concepts. For example, there was no measurement on which to compare differences in stress experienced by adolescents in Samoa and the United States.

For Brim and Spain, and for Orans, Mead's research design limited her ability to draw the conclusions she did. More attention to issues of research design and methods would have improved her chances to make valid claims and possibly limited later criticism of her work. In ethnographic research, no matter the mix of methods, the design of the study should allow for an ethnographers' hypotheses or hunches to be rejected as well as confirmed (Campbell 1975).

Research Design in Anthropological Practice: Systematic Research Strategies

The following examples illustrate some of the issues discussed so far. Several examples are reviews of studies that incorporate comparative designs of various types in which nonequivalent groups are constructed in order to control for as many extraneous factors as possible, and the manipulation of independent variables is a function of how comparative groups were chosen. These examples show how, even for less powerful designs, the interplay of exploratory and explanatory approaches can aid in guarding against threats to validity (Robson 1993).

Field Experiments

In field experiments, the experimenter has little control over all possible extraneous factors and the experiment may not involve random assignment of subjects to groups. Nevertheless, field experiments can be quite informative and, if carefully constructed, can provide formal tests of hypotheses derived from and complementary with ethnography. In his work on "colonizing the night," Melbin (1987) theorized that the night was a frontier, not unlike the western United States in the nineteenth century. Frontiers have certain features in common. Among other things, they provide escape and opportunity, tolerate a wider range of behaviors, consist of isolated settlements, have fewer status distinctions, involve novel hardships, have decentralized authority, involve lawlessness and peril, have a reputation for helpfulness and sociability, lag in the development of policies to exploit and regulate, and involve a variety of interest group conflicts.

Melbin conducted four tests of the feature relating to helpfulness and sociability. He designed a clever experiment in which keys were placed at similar locations during each two-hour field visit over a 24-hour period covering day and night. The idea was to see if there were a difference in key-returning behaviors among the different times. According to Melbin, "To find a key is to come across an implied need for help" (p. 75). The hypothesis was that residents of the night would return keys on average more often than those of the day. The keys had the request to "Please Return," with an address encased in plastic (keys dropped in the mailbox were delivered by the U.S. Postal Service to the address on the keys with postage due). Each of the keys were coded so they could be identified as to what time of day they were picked up and from what location.

In all, 326 keys were picked up, of which 220 were returned. Returned keys were also scored for the manner in which they were sent. One point was given for keys dropped unwrapped in the mail, two points for keys returned wrapped in an envelope, and three points if the envelope contained a personal note. Contrary to expectations, night-timers were not more amiable than day-timers in their key-returning behaviors; in fact, they were the least "helpful." However, Melbin's three other tests supported the hypothesis of more sociability and helpfulness at night.

Melbin speculates that the variation in results may have been due to the fact that the other three experiments involved direct personal contact among the subjects, while the key experiment involved no such interactions. This example illustrates nicely the importance of not relying on a single test, but having multiple tests and measures (Stinchcombe 1987). Had Melbin conducted only the key experiment, he may have come to very different conclusions regarding the helpfulness and sociability of night-timers. This example also shows how readily multiple tests can be incorporated into a research design within a field setting.

In all research, but particularly in field experiments like the one described above, there should be a concern for ethics and the well-being of experimental participants. Unlike studies where informed consent is obtained prior to participation, in experiments like Melbin's, individuals often participate without knowing about it. The ramifications and consequences of experimental outcomes must be considered thoroughly before any experimental design is implemented.

Control and Treatment in a Two-Community Comparative Design

One of the central concerns of medical anthropologists has been to better understand the relationship between health-related behaviors and native perceptions about illness. Young and Garro's (1982) investigation of treatment choice in two Mexican communities is an example of a static-group comparison where the presence or absence of the treatment is based on selection criteria not directly under the control of the researchers. One of the primary purposes of the research design was the elimination of competing hypotheses—the hallmark of good research design—and the testing of the primary hypothesis is an example of descriptive inference, as opposed to statistical inference. Descriptive inference is an approach highly suited for much anthropological research.

An important issue in this area of research concerns the factors influencing the use of Western treatments among non-Western populations. One explanation views use tied to congruence between a client's medical beliefs and scientific medical theory: the higher the congruence, the more likely the client will choose a physician's treatment. Termed the "conceptual-incompatibility" hypothesis, a number of studies have suggested that such a congruence was the primary determinant of treatment choice among Third World peoples. Young and Garro took a different stance, stressing physician accessibility as the most important determinant of physician use. An important element of this position is that traditional medical beliefs are not a barrier to choice of physician treatment.

The research design included the comparison of two Mexican communities that were similar in terms of cultural traditions and economies but varied in terms of access to Western medical services. The town of Pichátaro had restricted access (a 20-minute bus ride from Uricho), while the town of Uricho had easy access. From a random sample of approximately 10% of the households in each of the towns, Young and Garro collected data on the number of illnesses that had occurred during the previous two months and the treatment each had received. Later, the researchers collected triads data and what they call term-frame data on informants' perceived similarity of illnesses.

Young and Garro tested the two main hypotheses in sequence. They had to establish differences in treatment choice behavior in the two communities before they could assess any hypotheses concerning differences in beliefs. Using a standard chi-square test, the authors found a significant difference in the frequency distribution of treatment alternatives between the two towns, with the exception of folk curers. Thus, the two communities seemed to differ in their use of Western medical services. This established, Young and Garro could then test the second hypothesis relating to the similarity in beliefs between the two communities. Ironically, in statistical terms, the authors have more interest in the null hypothesis of no difference in beliefs than in the alternative hypothesis of a difference in beliefs between the two communities. Using multidimensional scaling, Young and Garro (1982) compared the belief data and found striking similarities in the medical beliefs of communities. They conclude:

> On the basis of the data from the triads study and the term-frame interviews, we see little reason to reject the "null hypothesis" of no significant differences between the responses of the two groups of informants. This leads us to the conclusion that the substantial variation apparent in the use of a physician's treatment between the two samples, a consequence of differential access to such treatment, occurs without corresponding degrees of variation in resident's attitudes and beliefs about illness. (p. 1462)

The authors' careful attention to research design and analytical issues contributed to the production of impressive evidence that casts doubt on the validity of the "conceptual-incompatibility" hypothesis. Note that the analysis used to test the hypothesis concerning similarities in beliefs involved descriptive inference, not statistical inference. Despite the authors' claims of finding no "significant difference," there was no real way, at least when the study was conducted, to assess the extent to which any differences were significant in the sense of statistical probability. Recent developments in statistical procedures allow us to assess the similarities in aggregated judged-similarity matrices between the two communities (see Handwerker and Borgatti, this volume, and Hubert 1987). In Young and Garro's case, a visual inspection of the graphical representations of the data could lead to no other conclusion than that there was little or no difference in beliefs between the two communities (see Figure 3). This distinction is important, particularly with regard to anthropological research, in that hypothesis-testing research can be done without narrowly restricting it to analytical methods using statistical inference.

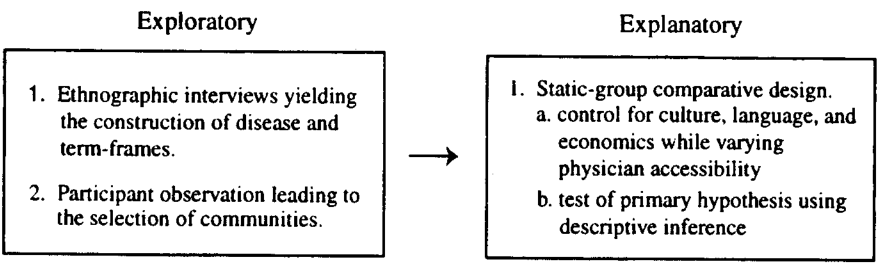


Figure 3. Overall design framework for the Young and Garro study.

There are, of course, threats to validity in this study. Because respondents weren't randomly assigned into comparison groups, it's difficult to know the influences of confounding variables on physician utilization and beliefs about illness. It is unrealistic to suppose that Young and Garro could have randomly assigned community members to the different comparison groups in order to control for confounding variables and then subject their informants to the treatments of interest. Given a lack of pretest observations, we can only assume that beliefs were similar prior to the availability of physicians in Uricho. In lieu of equalization through randomization, Young and Garro, through extensive ethnographic background research, produced groups that, although nonequivalent in the quasi-experimental sense, shared similarities with regard to a number of important characteristics. This isn't perfect, but a greater in-depth exploratory understanding and an explicit discussion of design can enhance our chances for the production of valid explanations.

Comparative Design and Ethnobiology

Ethnobiologists have long debated whether folk biological classifiers are natural historians who compare animals on the basis of their morphological characteristics or pragmatists who compare on the basis of the utility of organisms. Boster and Johnson (1989) explored this issue in an ethnobiological study of fish. Were individual informants classifying organisms on the basis of form or function? Boster and Johnson used a static group comparison design to compare several groups of expert fishermen with a group of novice fishermen. This is analogous to treatment and control groups without the random assignment of subjects to experimental units and where the treatment is implied rather than researcher directed (that is, natural differences in experience with fish). In the comparison, both culture and language were held constant while experience with fish was varied. Four groups—from North Carolina, East Florida, West Florida, and Texas—were sampled to examine the effects of different kinds of experience since there are regional variation in species abundance.

To ensure that experts were, in fact, experienced recreational fishermen, the rosters of sport fishing clubs in each region were sampled at random. The selection of control group subjects, by contrast, involved a purposeful selection procedure in which potential subjects were screened for recreational fishing experience. Using a questionnaire to gain background information, 15 college undergraduates who had the least amount of recreational fishing experience were selected from two introductory anthropology classes. These students were the control group. Each of the four expert groups comprised 15 subjects chosen at random from a larger sample of recreational fishermen. Thus the groups to be compared consisted of five groups of 15 subjects, four consisting of experts and one of novices.

All the groups were shown cards with artists' renderings and the common names of 43 marine species commonly found from North Carolina to Texas. Individuals were asked to perform an unconstrained judged similarity of the fish—a free pile sort (see Weller, this volume, and Weller and Romney 1988). Further, beliefs about the use and functional characteristics of the fish obtained from extensive ethnographic interviews were turned into a sentence-frame completion task described by Weller and Romney (1988). Finally, a measure of morphological similarities was determined, using taxonomic distances between pairs of fish. Boster and Johnson used statistical and graphical methods to evaluate whether experts' and novices' judgments of fish, at the aggregate and individual levels, were closer to the morphological characteristics of fish (taxonomic distance) or the uses of fish (beliefs about use). Using statistical and descriptive inference, the authors concluded that whether informants use form or function for classification depends on the knowledge base of the informants and the methods used to test their knowledge (see Figure 4).

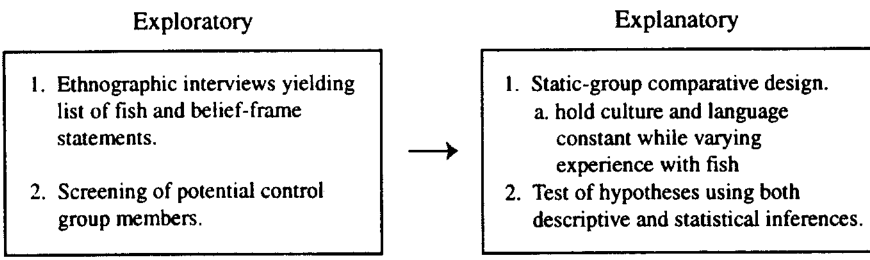


Figure 4. Overall design framework for the Boster and Johnson study.

Some of the criticisms of the Young and Garro study apply to this example as well. Lack of random assignment of subjects to treatment and control groups and pretest observations limit the ability to make causal inferences. But the in-depth ethnographic background research, the particular structure of the hypothesis, and the overwhelming reliability of informant responses make for more confidence in the possible validity of the study's conclusions.

Exploratory Research and the Development of Cultural Models

Often the primary objective of research design is more a matter of discovery and exploration than the testing of hypotheses. Although such designs are less driven by an established theoretical framework, there still is need to pay careful attention to a number of design details in the proper development of new theories and models.

An example of research in this mode is Naomi Quinn's (1996) development of Americans' cultural models concerning marriage.

There is a body of literature that views the interaction of culture with the individual as so deeply unique and personal as to not be researchable in terms of cultural universals, coherence, or even sharing. In contrast, Quinn views culture as being shared—that there are cultural models for a variety of domains that are widely held in common, and that these models can be developed from the discourses of cultural members.

Based on in-depth interviews with 22 informants, Quinn (1996) attempted to build a cultural model of Americans' reasoning about marriage. Because she was interested in a model that was shared, it was crucial to interview a wide range of couples who, although of the same culture, were not just from one region of the country or of only one ethnicity, religion, or social class. As she puts it:

> All of my interviewees were residents of the same middle-sized southeastern city or its immediate environs; all were native-born Americans who spoke English as a first language; and all were married during the period of their interviews, all in first marriages. Beyond these constancies of cultural and marital experience, they were selected to maximize diversity with regard to such obvious differences as their occupations and educational backgrounds, religious affiliations and ethnic and racial identities, their neighborhoods and social networks, and the duration of their marriages. (p. 399)

Although not generally representative of either the regional population or of the population of the United States, Quinn claims that her sample of informants represents the regions' population in terms of the high degree of recent in-migration to the area from regions outside the South. Her sample is an attempt to capture the range of diversity found in the region. In my view, the consistency of her findings in this diverse sample of informants makes her case stronger (see Johnson 1990). That is, finding commonality in the face of diversity provides stronger evidence of a shared cultural model (Johnson and Griffith 1996). In principle, this is similar to Cook and Campbell's (1979) model of deliberate sampling for heterogeneity as one of several means for warding off threats to external validity.

Based on an in-depth analysis of informants' discourse about marriage, Quinn produced a cultural model incorporating a number of causal links in informants' reasoning as to a "lasting marriage" (see Figure 5). Although the model appeared to be widely shared among informants from Quinn's sample and data collected from other studies on marriage, research still has to be designed to test this model across settings and researchers.

Issues of validity in this case are not as overriding as they would be in a purely explanatory study. Quinn was careful and diligent in her selection of informants, and her diligence certainly contributes to the potential validity of her model. However, further research in the explanatory mode is now warranted.

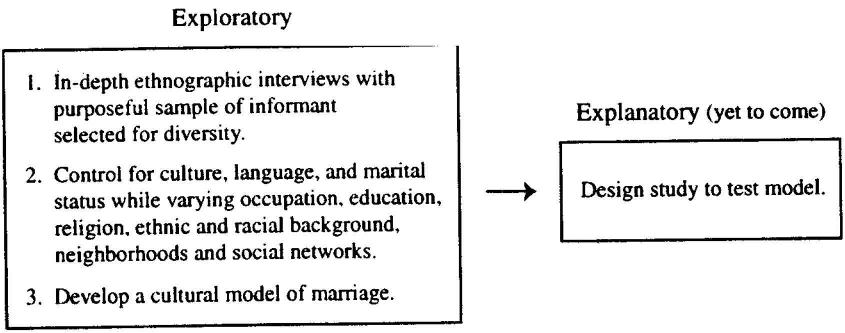


Figure 5. Overall design framework for the Quinn study.

Participant Observation and the Search for Validity

As seen in Figure 2, exploratory and descriptive research are often essential components of an overall explanatory research design. In a series of papers, Koester (1996) and his colleagues (Koester et al. 1996) offer excellent examples of the role of participant observation in more clearly defining the set of HIV risk behaviors surrounding injection drug use. In most earlier research on injection drug users (IDUs) and HIV risk, the primary risk factor was viewed in terms of direct needle sharing. Thus, most large epidemiological studies of IDUs focused mainly on direct sharing behaviors in attempts to understand seroconversion rates and other risk factors.

Based on participant observation among IDUs, Koester (1996) identified nine other behaviors that were outside the realm of the direct sharing of a single syringe by two or more IDUs. Termed "indirect sharing," these nine behaviors can promote the transmission of HIV among IDUs who, although not sharing needles directly, often share water for mixing of drugs or for rinsing syringes, share drug-mixing containers (cookers and spoons), share cottons for filtering, and share the actual drug solution itself. These findings are undeniably important for larger epidemiological work that examines elements of IDUs' behaviors and such things as producing valid models of seroconversion.

In a subsequent study, Koester et al. (1996) used these additional distinctions in sharing to look at the prevalence of injection-related HIV risk behaviors among several subpopulations of injection drug users (see Figure 6). A major component of the study was the comparison of IDUs who engaged in both direct sharing and indirect sharing with IDUs who engaged in indirect sharing only and those who several subpopulations of injection drug users (see Figure 6). A major component of the study was the comparison of IDUs who engaged in both direct sharing and indirect sharing with IDUs who engaged in indirect sharing only and those who neither shared directly nor indirectly. Statistical tests of group differences provided a greater understanding of the risk factors associated with the different types of behavior. This is a good example of the application of exploratory research in the production of better measures of potentially important explanatory variables.

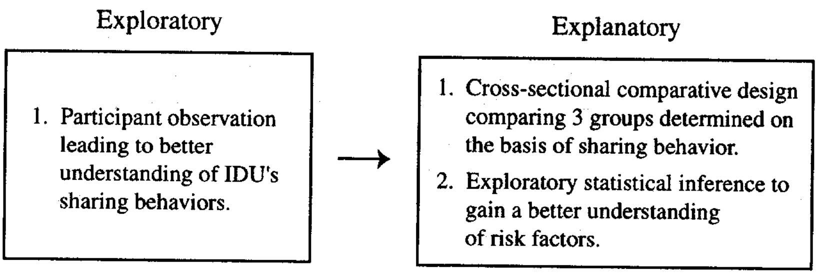


Figure 6. Overall design framework for the Koester and Koester et al. studies.

Case-Control Study Design: *Susto*, A Folk Illness

Here we look at an example of a study design used to investigate the extent to which disease is molded by culture. Rubel et al. (1985) report on a study of a folk illness known as *susto*, found in many cultural groups throughout North and South America. Folk beliefs surrounding *susto* attribute loss of a critical substance or force due to a frightening experience. The authors were interested in three primary hypotheses relating to role performance and the presence of the illness, psychiatric impairment and *susto*, and relationship between organic disease and *susto*. The ultimate aim of the study was to show the relationship between various social forces and *susto* susceptibility.

The design involved three communities that differed in history, language, and culture but had similarities in social, demographic, and economic factors. Rubel et al. carefully selected communities that were as similar as possible in terms of forms of government and gender-specific role expectations. The proposed design could have been conducted in a single community, but the authors felt that the generalizability of the results would be enhanced with a multiple case-control study design.

One subsample was of individuals who complained of *susto* during the fieldwork or who had admitted their condition to relatives or curers. Selection depended on the condition by men in the community and may also have affected the reporting of the condition by women. Another problem involved the existence of more social stratification in one community than expected, leading to a lower incidence of reported cases of the illness (higher-income people recognized the condition but felt that belief in it was more superstitious than real). However, Rubel et al. felt comfortable with the comparability among the *susto* subsamples from the communities. These groups will be referred to as the *asustados* groups.

The researchers were careful to make the control group as comparable to the *asustados* groups as possible. Because the *asustados* were "sick," control group members must also be sick. Thus, sick people were compared to sick people and control group members were selected from the pool of patients at the project clinics in each of the communities. Patient records provided the information on which to make the final selection. In addition to the control group being sick, males were matched with males and females with females and *asustados* and controls were matched in terms of age. Matched pairs were made within communities only. This design allowed for a variety of comparisons, including comparisons by controls and *asustados*, by gender, and by matched pairs both within and between cultural groups (see Figure 7).

Symptomology and health problems were operationalized using a panel of physicians. Psychiatric impairment was operationalized using the 22-item Screening Score for Psychiatric Impairment. Based on earlier ethnographic research, social stress, an important component for understanding an individual's inability to perform social roles, was operationalized using the Social Stress Gauge developed by one of the researchers.

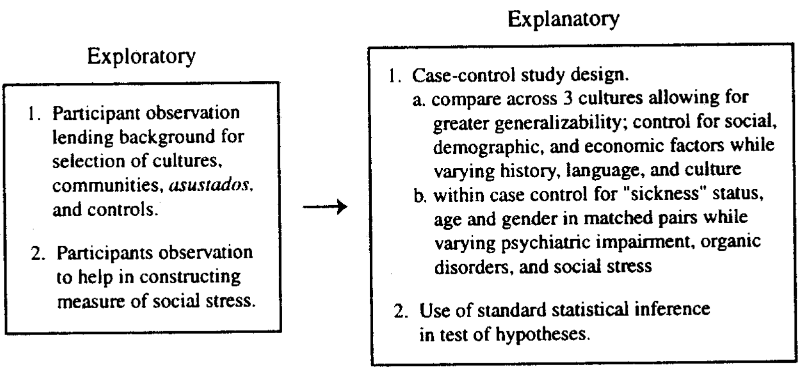


Figure 7. Overall design framework for the Rubel et al. study.

Using standard methods of statistical inference, Rubel et al. found that there was, in fact, an association between *susto* and an individual's perception of the adequacy of his or her performance of critical social roles. Although there was no association between *susto* and psychiatric impairment, there was a relationship between *susto* and the suffering of more organic disease signs.

This study is important because of the authors' candor about the potential threats to validity they encountered in conducting the research. The stigma of *susto* among males and the greater social stratification encountered in one of the communities are possible threats to the validity of their conclusions. But the researchers' awareness of the problems, combined with the strength of their multiple case-control study design, increases our confidence in their conclusions. This is an excellent example of a study design that incorporates within-study replication or multiple tests of a theory. Multiple tests are always much more convincing than a single test (Stinchcombe 1987).

Multimethod Ethnography and the Comparison of Models

Many peasant societies in Central America are experiencing dramatic economic and cultural change. One consequence of these changes is increasing economic differentiation. Parallel to these economic changes, there has been a shift in religious preference over the last 70 years. For some researchers, the shift from Catholicism to Protestantism helps account for economic change, as Protestantism is more compatible with capitalist ideology and the accumulation of wealth. Goldin (1996) wanted to understand the relationships among religious affiliation, economic ideology, occupation, and economic status in a Guatemalan township (Almolonga). Her study design incorporates quantitative and qualitative methods in the overall ethnographic enterprise.

Based on extensive participant observation, Goldin constructed four plausible models that might account for what she observed while in the field. Using her experience as a participant observer, Goldin developed a survey which she applied to a random sample of 10% of the heads of households in the township ($n = 57$). She made an earnest attempt to control for as many biases as possible and, using the data collected during the survey, conducted statistical tests of the four competing models. This provided for an evaluation of the explanatory power of each. Her selection of variables allowed a comparison of different levels (for example, Catholic versus Protestant) across the four variables. Using path analytic modeling, she applied different statistical controls in each of the competing models. As she describes the process leading to the selection of the best model:

> The results of my study, of course, must be interpreted within the constraints of the data collection methods. First, qualitative approaches were used to suggest different mechanisms and relationships that might be operating within Almolonga. Then, a

> survey approach was used to evaluate the viability of these mechanisms in terms of characterizing general trends within Almolonga. My conclusions must be interpreted in terms of these general trends. I don't doubt that there are exceptions to them. Indeed, I interacted with several individuals who had life histories that were inconsistent with my general characterization and who were the basis for suggesting the competing models discussed above. However, when a large representative sample of the township was aggressively pursued, the different data sets tended to support model C as the one that characterizes the general tendencies within the township. (p. 72)

This study is an example of multimethod ethnography in which there was a combination of exploratory and explanatory approaches—that is, qualitative data and tests of models with data collected using a cross-sectional design. The combination helped Goldin in the specification of appropriate variables, in the development of a sound survey instrument, and in the specification and assessment of the four competing models. The study shows how the use of multiple methods fosters triangulation that contributes to the production of valid conclusions (see Figure 8). Her research design illustrates the danger in relying on a single method without attention to sampling. Had Goldin relied exclusively on, say, the life histories of a nonprobabilistic sample of informants without specified selection criteria (Johnson 1990), she might have arrived at a very different, and possibly erroneous, conclusion.

Exploratory		Explanatory
1. Participant observation leading to the development of four competing models. 2. Participating observation aiding in the development of survey instrument and sample design.	→	1. Cross-sectional survey design comparing across different levels of religious affiliation, occupation, economic status, and economic ideology. 2. Applications of statistical procedure that allows for statistical controls in the evaluation of the models.

Figure 8. Overall design framework for the Goldin study.

Summary

This review of research design and strategies in cultural anthropology only scratches the surface of the research designs, hybrid designs, and combinations of designs possible within an ethnographic context. The newer forays into experimental and other ethnographic forms of presentation are more reflexive in character and more concerned with believable and moving representations rather than the production of valid accounts or conclusions.

With advances in computer technology, qualitative data analysis can now be a powerful mode to test theories. Similarly, advances in computer-intensive methods for testing hypotheses have the potential to expand the range of designs possible, particularly in the imperfect world of fieldwork (Johnson and Murray 1997). The strength of the ethnographic approach is its ability to incorporate a wide range of methods, strategies, and designs within a single enterprise, all combining in ways to improve the chances for credible results.

As anthropologists, we should take full advantage of both our current understanding of research design and these new developments to produce a "powerful mode of argumentation." It is mostly through attention to these concerns that anthropology and anthropologists will have the opportunity to, as Agar says, "move the world."

REFERENCES

Agar, M. 1980. *The Professional Stranger: An Informal Introduction to Ethnography.* New York: Academic Press.

Agar, M. 1996. *AAA Newsletter.* January.

Babbie, E. 1990. *Survey Research Methods*, 2d ed. Belmont, CA: Wadsworth.

Behar, R. 1993. *Translated Woman.* Boston: Beacon Press.

Bernard, H. R. 1994. *Research Methods in Anthropology: Qualitative and Quantitative Approaches*, 2d ed. Walnut Creek, CA: AltaMira Press.

Boas, F. 1920. The Methods of Ethnology. *American Anthropologist 22*(4):311–321.

Boster, J. S., and J. C. Johnson. 1989. Form or Function: A Comparison of Expert and Novice Judgments of Similarity Among Fish. *American Anthropologist 91*(4): 866–889.

Brim, J. A., and D. H. Spain. 1974. *Research Design in Anthropology: Paradigms and Pragmatics in the Testing of Hypotheses.* New York: Holt, Rinehart and Winston.

Campbell, D. T. 1975. "Degrees of Freedom" in the Case Study. *Comparative Political Studies 8*(2):178–213.

Campbell, D. T. 1988. Qualitative Knowing in Action Research. In *Methodology and Epistomology for Social Science: Selected Papers* E. S. Overman, ed. Pp. 360–376. Chicago: University of Chicago Press.

Cook, T. D. 1994. *Criteria of Social Scientific Knowledge: Interpretation, Prediction, Praxis.* Lanham, MD: Rowman and Littlefield.

Cook, T. D., and D. T. Campbell. 1979. *Quasi-Experimentation: Design and Analysis for Field Settings.* Chicago: Rand McNally.

Denzin, N. K., and Y. S. Lincoln. 1994. Entering the Field of Qualitative Research. In *Handbook of Qualitative Research.* N. K. Denzin and Y. S. Lincoln, eds. Pp. 1–19. Thousand Oaks, CA: Sage Publications.

Ellen, R. F. 1984. Introduction. In *Ethnographic Research: A Guide to General Conduct.* R. F. Ellen, ed. Pp. 1–12. London: Academic Press.

Ember, M. 1985. Evidence and Science in Ethnography: Reflections on the Mead-Freeman Controversy. *American Anthropologist 87*(4):906–910.

Faia, M. A. 1993. *What's Wrong with the Social Sciences? The Perils of the Postmodern.* Lanham, MD: University Press of America.

Freeman, D. 1983. *Margaret Mead and Samoa: The Making and Unmaking of an Anthropological Myth.* Cambridge: Harvard University Press.

Glaser, B. G., and A. L. Strauss. 1967. *The Discovery of Grounded Theory: Strategies for Qualitative Research.* Chicago: Aldine.

Goldin, L. R. 1996. Models of Economic Differentiation and Cultural Change. *Journal of Quantitative Anthropology 1–2*(6):49–74.

Harris, M., J. G. Consorte, J. Lang, and B. Byrne. 1993. Who Are the Whites: Imposed Census Categories and the Racial Demography of Brazil. *Social Forces 72*(2): 451–462.

Hubert, L. J. 1987. *Assignment Methods in Combinational Data Analysis.* New York: Marcel Dekker.

Hurlbert, S. H. 1984. Pseudoreplication and Design of Ecological Field Experiments. *Ecological Monographs 54*(2):187–211.

Johnson, J. C. 1990. *Selecting Ethnographic Informants.* Qualitative Research Methods Series, Vol. 22. Thousand Oaks, CA: Sage Publications.

Johnson J. C. 1994. Anthropological Contributions to the Study of Social Networks: A Review. In *Advances in Social Network Analysis.* S. Wasserman and J. Galaskiewicz, eds. Pp. 113–151. Thousand Oaks, CA: Sage Publications.

Johnson, J. C., and D. C. Griffith. 1996. Pollution, Food Safety, and the Distribution of Knowledge. *Human Ecology 24*(1):87–110.

Johnson, J. C., M. Ironsmith, A. L. Whitcher, G. M. Poteat, and C. W. Snow. 1997. The Development of Social Networks in Preschool Children. *Early Education and Development 8*(4):389–406.

Johnson, J. C., and J. D. Murray. 1997. Evaluating FAD Effectiveness in Development Projects: Theory and Praxis. In *Fish Aggregation Devices in Developing Fisheries: Potential and Pitfalls.* R. Pollnac and J. Poggie, eds. Pp. 143–158. Kingston: ICMRD.

Johnson, J. C., and R. Pollnac, eds. 1989. *Managing Marine Conflicts.* Special issue of *Ocean and Shoreline Management 12*(3).

Kempton, W., J. S. Boster, and J. A. Hartley. 1996. *Environmental Values in American Culture.* Cambridge: M.I.T. Press.

Kincheloe, J. L., and P. L. McLaren. 1994. Rethinking Critical Theory and Qualitative Research. In *Handbook of Qualitative Research*. N. K. Denzin and Y. S. Lincoln, eds. Pp. 138–158. Thousand Oaks, CA: Sage Publications.

Kirk, J., and M. L. Miller. 1986. *Reliability and Validity in Quantitative Research.* Qualitative Research Methods Series, Vol. 1. Thousand Oaks, CA: Sage Publications.

Kleinbaum, D. G., L. L. Kupper, and H. Morgenstern. 1982. *Epidemiologic Research: Principles and Quantitative Methods.* Belmont, CA: Lifetime Learning Publications.

Koester, S. 1996. The Process of Drug Injection: Applying Ethnography to the Study of HIV Risk Among IDU's. In *AIDS, Drugs and Prevention: Perspectives on Individual and Community Action.* T. Rhodes and R. Hartnoll, eds. Pp. 133–148. London: Routledge Press.

Koester, S., R. E. Booth, and Y. Zhang. 1996. The Prevalence of Additional Injection-Relation HIV Risk Behaviors Among Injection Drug Users. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology 12*:202–207.

Kruskal, J. B., and M. Wish. 1978. *Multidimensional Scaling.* Beverly Hills, CA: Sage Publications.

Kuznar, L. A. 1997. *Reclaiming a Scientific Anthropology.* Walnut Creek, CA: AltaMira Press.

LeVine, R. A. 1973. Research Design in Anthropological Field Work. In *A Handbook of Methods in Cultural Anthropology.* R. Naroll and R. Cohen, eds. Pp. 183–195. New York: Columbia University Press.

Lowie, R. H. 1937. *The History of Ethnological Theory.* New York: Rinehart.

Marcus, G. 1983. One Man's Head. *New York Times Book Review, March 27*:3, 22–23.

Melbin, M. 1987. *Night as Frontier: Colonizing the World After Dark.* New York: The Free Press.

Miles, M. B., and A. M. Huberman. 1994. *Qualitative Data Analysis*, 2d ed. Thousand Oaks, CA: Sage Publications.

Moran, E. F., ed. 1995. *The Comparative Analysis of Human Societies: Toward Common Standards for Data Collection and Reporting.* Boulder: Lynne Rienner.

Naroll, R., and R. Cohen, eds. 1973. *A Handbook of Method in Cultural Anthropology.* New York: Columbia University Press.

Noreen, E. W. 1989. *Computer-Intensive Methods for Testing Hypotheses.* New York: John Wiley.

Orans, M. 1996. *Not Even Wrong: Margaret Mead, Derek Freeman, and the Samoans.* Novato, CA: Chandler and Sharp.

Panourgia, N. 1995. *Fragments of Death, Fables of Identity.* Madison: University of Wisconsin Press.

Pelto, P. J., and G. H. Pelto. 1978. *Anthropological Research: The Structure of Inquiry*, 2d ed. Cambridge: Cambridge University Press.

Plattner, S. In press. Scientific Anthropology at the National Science Foundation. In *Anthropology Between Science and the Humanities*. C. Furlow, ed. Walnut Creek, CA: AltaMira Press.

Porter, T. M. 1995. *Trust in Numbers*. Princeton: Princeton University Press.

Quinn, N. 1996. Culture Contradictions: The Case of America's Reasoning about Marriage. *Ethos 24*(3):391–425.

Rabinow, P. 1986. Representations are Social Facts: Modernity and Post-Modernity in Anthropology. In *Writing Culture: The Poetics and Politics of Ethnography*. J. Clifford and G. E. Marcus, eds. Pp. 234–262. Berkeley: University of California Press.

Ramos, A. R. 1995. *Sanuma Memories*. Madison: University of Wisconsin Press.

Robson, C. 1993. *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*. Oxford: Blackwell Publishers.

Rubel, A. J., C. W. O'Nell, and R. Collado-Ardon. 1985. *Susto, A Folk Illness*. Berkeley: University of California Press.

Schneider, D. 1996. Alarming Nets. *Scientific American (September)*:40–42.

Schwandt, T. A. 1994. Constructivist, Interpretivist Approaches to Human Inquiry. In *Handbook of Qualitative Research*. N. K. Denzin and Y. S. Lincoln, eds. Pp. 118–138. Thousand Oaks, CA: Sage Publications.

Sechrest, L. 1973. Experiments in the Field. In *A Handbook of Methods in Cultural Anthropology*. R. Naroll and R. Cohen, eds. Pp. 196–209. New York: Columbia University Press.

Seidman, S. 1994. *The Postmodern Turn: New Perspectives on Social Theory*. Cambridge: Cambridge University Press.

Shankman, P. 1996. The History of Samoan Sexual Conduct and the Mead-Freeman Controversy. *American Anthropologist 98*(3):555–567.

Spindler, G., and W. Goldschmidt. 1973. An Example of Research Design: Experimental Design in the Study of Culture Change. In *A Handbook of Method in Cultural Anthropology*. R. Naroll and R. Cohen, eds. Pp. 210–219. New York: Columbia University Press.

Stinchcombe, A. L. 1987. *Constructing Social Theories*. Chicago: University of Chicago Press.

Taussig, M. 1987. *Shamanism, Colonialism, and the Wild Man: A Study in Terra and Healing*. Chicago: University of Chicago Press.

Tyler, S. A. 1991. A Post-modern In-stance. In *Constructing Knowledge: Authority and Critique in Social Science*. L. Nencel and P. Pels, eds. Pp. 78–95. London: Sage Publications.

Urry, J. 1984. A History of Field Methods. In *Ethnographic Research: A Guide to General Conduct*. R. F. Ellen ed. Pp. 35–62. London: Academic Press.

Van Maanen, J. 1988. *Tales of the Field: On Writing Ethnography*. Chicago: University of Chicago Press.

Weller, S. C., and A. K. Romney. 1988. *Systematic Data Collection*. Qualitative Research Methods Series, Vol. 10. Thousand Oaks, CA: Sage Publications.

Werner O., and G. M. Schoepfle. 1987. *Systematic Fieldwork*, Vol. 2. Thousand Oaks, CA: Sage Publications.

Whyte, W. F. 1984. *Learning from the Field: A Guide from Experience*. Newbury Park, CA: Sage Publications.

Williams, T. R. 1967. *Field Methods in the Study of Culture*. In the series *Studies in Anthropological Method*, George Spindler and Louise Spindler, eds. New York: Holt, Rinehart and Winston.

Young, J. C., and L. Y. Garro. 1982. Variation in the Choice of Treatment in Two Mexican Communities. *Social Science and Medicine 16*:1453–1465.

Zabusky, S. E. 1995. *Launching Europe: An Ethnography of European Cooperation in Space Science*. Princeton: Princeton University Press.