Handbook of Methods in Cultural Anthropology

Second Edition

Edited by
H. Russell Bernard and Clarence C. Gravlee

ROWMAN & LITTLEFIELD
Lanham • Boulder • New York • Toronto • Plymouth, UK

Published by Rowman & Littlefield
4501 Forbes Boulevard, Suite 200, Lanham, Maryland 20706
www.rowman.com

10 Thornbury Road, Plymouth PL6 7PP, United Kingdom

British Library Cataloguing in Publication Information Available

Library of Congress Cataloging-in-Publication Data

Handbook of methods in cultural anthropology / edited by H. Russell Bernard and Clarence C. Gravlee. — Second edition.
pages cm
Includes bibliographical references and index.
ISBN 978-0-7591-2070-9 (cloth : alk. paper) — ISBN 978-0-7591-2071-6 (pbk. : alk. paper) — ISBN 978-0-7591-2072-3 (electronic) 1. Ethnology—Methodology. I. Bernard, H. Russell (Harvey Russell), 1940–
GN345.H37 2014
305.8001—dc23

2014007881

∞™ The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI/NISO Z39.48-1992.

Printed in the United States of America

CHAPTER 3

Research Design and Research Strategies

Jeffrey C. Johnson and Daniel J. Hruschka

> 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)

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, to be viewed as proper evidence for making a case? Although some may be swayed by the elegance of a well-written essay, for many it is crucial to know something about the author, his or her motivations, and the experiences, skills, and methods of investigation 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.

Research design is the careful planning and implementation of this process of knowing. A priori planning of all phases of research (including analysis and writing) can benefit one's research in several ways.

First, planning helps us deal with (and take advantage of) inevitable contingencies in the field. Research projects rarely, if ever, unfold seamlessly, and a plan provides a backbone for making (and justifying) tough decisions about site selection, participant sampling, and data collection and analysis in the face of inevitable bumps and detours.

Second, clearly specifying the plan as well as its imperfect implementation lets other researchers check under the hood, identify potential biases, and judge for themselves whether the results are supported by the data and the methods. This transparency helps the reader determine how robust the findings are when research procedures change in presumably minor ways (reliability), how well the findings might extend to other people or groups (generalizability), and how well the findings really represent what we think they do (validity). Exposing these imperfections can be disconcerting, but it ensures that others can judge a researcher's claims and can build on his or her research.

Third, an idealized plan clarifies how methods of data collection and analysis are linked to specific concepts and processes in one's theory or theories.

Finally, on a practical level, good research design is also essential 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 (e.g., validity, reliability, hypotheses, etc.). Others may require a thorough exploration 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 is sometimes called the laundry-list component of research versus research design. The laundry-list component is important. It involves details about getting into and out of the field situation, making 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 planning that contributes to the credibility, validity, believability, replicability, or plausibility of any study.

In this chapter, we concentrate on elements of design important for generating valid results or a believable account. We describe the benefits of research design, outline how cultural anthropologists have traditionally approached study design in their research, and discuss how cultural anthropologists approach research design today. Then we describe key elements in research designs, focusing on choice of comparison groups and sampling, and present case studies of common designs in anthropology—including exploratory designs, comparing individuals within populations, case-control designs, two-site comparisons, large-scale cross-population comparisons, field experiments, and longitudinal studies.

THE BENEFITS OF DESIGN

Evidence for the powerful benefits of research design is all around us. For example, the invention of the control/treatment design of clinical trials allowed researchers in the twentieth 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 also 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 a matter of who has the most political clout.

The outcry for a ban on nets in tuna fishing is a famous example. Environmental organizations launched campaigns to ban nets in tuna fishing because dolphins are often caught incidentally in that fishery. Media campaigns in the United States 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, in the absence of a research design that addresses a specific problem or goal, the policy emerges merely from interactions between groups of differing political, ideological, social, and economic backgrounds.

There was a similar concern over the unintended but regular 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 nets 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—and subsequent revisions to the 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 (i.e., the effectiveness of using 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, attention to 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, 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), and the first edition of this handbook (Bernard 1998). Bernard (2011) 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.

BOAS, MALINOWSKI, AND RESEARCH DESIGN IN THE SCIENTIFIC TRADITION

Despite anthropologists' long-held interest in attention to detail, the structure for gathering those details has been documented and explicated much less thoroughly than have the details themselves. Franz 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.

Notwithstanding his concern for scientific method, Boas was more explicit about his methods of data analysis than about his methods of fieldwork and data collection (Boas 1920; Ellen 1984). Malinowski, another great contributor to anthropological inquiry, was also concerned with the aims of science and methodological rigor. However, his earliest contributions were more a demonstration of the value of ethnographic writing—his "unusual literary sense" (Lowie 1937, 231)—than of methodological details of 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 we 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 actually attempting to forgo 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. (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 (i.e., 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 discussion of this issue, and see Ember et al., this volume).

Second, Boas's concern for contextual meaning over the statistical analysis of data was prophetic. The concern for understanding human behavior in context is one of anthropology's strengths. However, thinking of quantification as incompatible with an attention to context and meaning has often clouded discussions about research design.

Boas's final sentence in his response to Mead illustrates that even at this early stage the issue of subjectivity of ethnographic research was of concern. There was faith, though, that awareness of the potential biases associated with the investigator's subjectivity meant that it 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 there was a concern (perhaps incorrect) that meaning might be compromised. Mead's position on these various elements of research design provided fuel for the continuing discussions about how successful her original findings were in addressing the questions she posed, how valid her interpretations were, and whether the research design was adequately rigorous (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 research 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 possessed a 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 staunch belief, across the British and U.S. traditions, in the trial-by-fire method of training ethnographers. This belief likely supports the current lack of formal training in methods and research design in cultural anthropology. Agar (1980) and Bernard (2011) relate stories about Kroeber's recommendations regarding how ethnographic research is taught and conducted. One story concerns Charles Wagley's teaching of a field methods course while the other concerns a graduate student at Berkeley asking Kroeber for methodological advice before going to the field. According to Agar, student folklore has it that Kroeber stated tersely to the nervous student "I suggest you buy a notebook and a pencil" (Agar 1980, 2).

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, including the operationalization of concepts, the reliability and generalizability of findings, and the assessment of competing claims through formal model comparison and hypothesis testing (LeVine 1973). A good example of this is a book by Thomas Rhys Williams (1967) published in the Spindlers' 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)

Early on, LeVine (1973) and others (Johnson 1990) made 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 with a description of what is really going on there.

Laboratory and survey researchers have some flexibility to change the problem focus and study populations in light of emerging problems, but fieldworkers 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 ethnographers resist being expected to replicate someone else's work. The "my natives" or "my village" mentality of some and the fact that careers are made by discovering new theories or describing exotic less well-known cultures has certainly inhibited replication efforts (Johnson 1990).

CONTEMPORARY RESEARCH STRATEGIES IN CULTURAL ANTHROPOLOGY

As cultural anthropologists plan and conduct research, there are several distinct goals that they can pursue, a variety of criteria they may choose for judging good work, and numerous research strategies for fulfilling these specific criteria. Among possible goals, researchers can aim to accurately describe a social situation, vividly convey another person's perspective, teach a lesson, or develop and test general explanations for behavior. Debates sometimes erupt about the primacy of these goals, but they need not be at odds.

Cultural anthropologists also draw from a diverse set of criteria for evaluating good work. Here is a list of some of these criteria:

1. Does the account vividly convey the situation?
2. Is the description sufficiently rich?
3. Is the narrative compelling?
4. Do the local actors agree with the findings or interpretation?
5. Is the description accurate by some objective criteria?
6. Is the description consistent with all data?
7. Are rival explanations or interpretations considered?

8. Do the methods appropriately capture the intended concepts and variables?
9. Are the methods reliable?
10. Can the findings be replicated?
11. Are findings generalizable?
12. Do the findings advance or challenge established theory?

To create work that meets these criteria, or at least appropriate subsets of these criteria, researchers have drawn from a number of research strategies. These include extended fieldwork in a setting, crafting morally compelling narratives, triangulating observations and data to minimize bias, using systematic sampling of populations, and planning study designs that can discriminate between different explanations for human behavior.

Although we may be tempted to classify researchers or traditions as preferentially valuing specific criteria, it is an empirical question whether such clean divisions exist. In fact, anthropology may be unique among disciplines in its tolerance for such a diverse set of criteria. That said, solid research design is especially crucial for satisfying the criteria of accurate description (#5), consistency with data (#6), discrimination between alternative explanations (#7), reliability (#8), validity (#9), replicability (#10), and generalizability (#11). These are also some of the key criteria used for judging scientific work more generally.

Research design can involve both qualitative and quantitative data, objective and subjective measurements, and biological and cultural components in the same project. Research design is also an important element of both exploratory/inductive and evaluative/deductive phases of scientific inquiry (see the section on Research Design in Scientific Inquiry, below).

Regardless of whether you are pursuing a more exploratory or a more explanatory agenda, research design requires advanced planning about how you will generate reliable, valid, and generalizable findings or how the study can discriminate between alternative explanations for the phenomenon being studied. This involves meticulous attention to questions about sampling (to accurately represent a population and to improve generalizability), about measurement and triangulation (to ensure reliability and validity), about analyzing the data and reporting results, and about the kinds of data required to rule out or rule in specific hypotheses.

A key principle in designing the process is to ward off as many threats to validity as possible by using appropriate methodological and analytical checks. The vicissitudes of real life in a field site nearly always requires some adjustments will be made to such advanced planning. Nonetheless, decisions about these adjustments still rely on the same concerns about minimizing threats to validity within the constraints of the field site, and being skilled, or at least experienced, in research design will be important in making those adjustments. Similar to approaches critical of science, the design process is based on extreme skepticism about the researcher's ability to minimize bias. For example, the research design process is highly skeptical of meeting the criteria stated above by relying exclusively on a single ethnographer and his or her memory or field notes as the single instrument of measurement. Research design aims to develop and implement measures that minimize bias.

Many ethnographic studies have an exploratory and descriptive aim. But within this large genre, authors vary greatly in the research strategies they follow. For example, Zabusky's (1995) ethnographic study of cooperation in European space science takes "the form of mutual exploration rather than unidirectional examination" (p. 46). She contrasts her study with research on cooperation conducted 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 Clifford 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 semi-structured 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 her ethnography.

In contrast to Zabusky, there is a body of work in anthropology that is more extreme in its rejection of systematic design issues in favor of vivid accounts (#1), rich description (#2), and compelling narratives (#3) as well as novelty in representing the human subjects (including the ethnographer). Ramos (1995), for example, 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 50 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 to be less concerned by the issues of bias and validity or 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.

Another tradition in anthropology that gained prominence in the last 30 years focuses on human behavior, speech, and culture as texts to be analyzed by principles similar to those in literary analysis. A parallel tradition focused on problems of representation, delving into experimental 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).

A different concern is threats to compellingness—whether the story is worth telling or reading (Tsing 2005). The methodological focus is on engagement and representation. How does the researcher obtain suitable immersion into the cultural context of the actor(s) to represent it? How does one write a text that conveys the understanding gained from such an immersion in a believable and compelling way?

A great deal of innovation in descriptive, ethnographic research has involved moving away from place-based inquiries in traditional or small-scale societies toward new populations and social situations and larger processes of globalization. For example, multi-sited ethnography is a research strategy that follows a topic, social problem, or object through field sites that may be geographically or socially distant from one another (Marcus 1995). Nancy Scheper-Hughes (2004) used this approach to study the global traffic in human body parts by "following the bodies" to the diverse places where human organs and tissue are extracted, brokered, transported, and transplanted (p. 32). A key goal of this research was to make public (p. 37) the harvesting, selling, and distribution of human body parts; the research shows the value of ethnography in revealing covert social worlds.

Scheper-Hughes observed practices and social interactions in these diverse sites, conducted open-ended interviews with key informants, and administered structured questionnaires with vulnerable populations. To expose these practices concerning body parts, she identified key players in the trade—surgeons, kidney sellers and buyers, kidney hunters, and kidney brokers—who would grant interviews. To document various views of the organ trade as well as the participation of vulnerable populations in the organ trade, Scheper-Hughes conducted surveys of squatter settlements in South Africa, five villages in Moldova, and residents of a large slum in the Philippines. Each of these pieces of data provides material for describing this particular social world.

Another ethnography in the global framework is one by Anna Tsing (2005). Whereas Scheper-Hughes focuses on key intermediaries in global connections (sellers, buyers, brokers), Tsing examines connections and interactions between local and global actors to tell the story of rainforest exploitation in Indonesia in the 1990s. To do this, she focuses on the awkward interactions of key actors and agencies. Tsing uses observation, interviews, and journalistic and archival data as raw material for these stories. She devotes considerable care to crafting these stories—to make the reader "feel the rawness of the frontier" (p. 28) and to ensure the story is compelling (p. 25). A key strategy in her ethnography is to point out observations and anecdotes that don't fit (i.e., that "interrupt") dominant narratives about capitalism and globalization (p. 270). The observations on which these stories are based presumably derive from Tsing's personal experience. However, the work makes no reference to checks on reliability and validity threats—for example a possible bias toward telling stories in a way that accentuates divergences from dominant narratives to support the overarching thesis about cultural friction.

More recently, anthropologists have expanded their ethnographic inquiries into subject matter and populations that fall well outside of anthropological traditions.

Most notable among these is Tom Boellstorff's (2008) study of virtual worlds in his ethnography of Second Life. He advocates the development of a virtual anthropology, as opposed to just simply virtual ethnography, that would bring a unique anthropological perspective to the study of online and virtual worlds. In this ethnographic study of Second Life, Boellstorff wants to understand social interaction through participant observation. As one might imagine, participant observation in this context poses its own set of unique challenges and opportunities.

Boellstorff carefully frames the bounds of inquiry in terms of interaction among participants in the virtual space, with little concern for the nature of their lives and interactions in their non-virtual worlds. He creates a character and participates in interactions with others engaging in conversations and teleporting to virtual locations (e.g., a friend's house). His data are derived from casual conversations, observations, formal interviews, and even virtual focus groups, being careful to inform all encountered that he is an anthropologist doing a study of Second Life.

Boellstorff's conception of participant observation here involves the discovery of "culture through nonelicited, everyday interactions" (p. 72). This distinction is critical to his methodological approach in that, in his mind, the use of elicitation techniques for interviewing (as might be found in cognitive anthropology) only helps in uncovering the cultural rules for interaction, which he sees as an impoverished model of culture, rather than a deeper understanding of "culture in virtual worlds" (p. 66). In emphasizing this epistemological and methodological tension, he contrasts what he calls *epistome* (science or knowledge) from *techne* (craftsmanship), reflecting an attempt to understand the complex whole rather than just simple knowledge and beliefs.

Karen Ho's (2009) book on financial firms on Wall Street is another example of an ethnographic study in a nontraditional setting. Anthropologists have been notoriously absent in the study of the corporate world, particularly finance. Having worked on Wall Street in several jobs, Ho used her extensive network of friends and acquaintances to make contacts and conduct interviews, in a sense, "snowballing" her way through the financial world of Wall Street. As in some of the earlier ethnographies focusing on the complexity of global connections and processes, her ethnography sought to look simultaneously at both the globalization of capital markets and the strategies of financial actors.

Using participant observation as more of a true participant rather than a simple observer (Johnson et al. 2006), Ho engaged in a study of the powerful or engaged in "studying-up." Participant observation was important in that it provided a referent to better situate the talk of powerful informants. Ironically, the usual means for dealing with the imbalance of power between ethnographer and those studied, giving voice to the native, could be problematic in this case in that giving her informants voice could over-privilege the powerful.

These examples offer only a glimpse of the range of possible research strategies that can be used to do research and write ethnographies. In some cases, producing a picture of human life of a population or for a phenomenon of interest is an exploratory enterprise with an implicit concern for methodological issues. In other cases, anthropological research 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. In some cases, more authoritative ethnographic methods are combined with literary devices to develop a more compelling argument.

This combination of more ethnographic authority and ethnographic crafting has been the more recent trend, as illustrated by the last few ethnographic examples. Again, these innovations in anthropology do not necessarily need to be at odds with systematic inquiry. We now focus primarily on research designs aimed at the scientific concerns laid out earlier, including accurate description (#5), consistency with data (#6), discrimination between alternative explanations (#7), reliability (#8), validity (#9), replicability (#10), and generalizability (#11). For further discussion of research strategies in the interpretive mode, see Fernandez and Herzfeld (this volume).

RESEARCH DESIGN FOR SYSTEMATIC INQUIRY

In some social science disciplines, like psychology, the design of research is driven by the method of analysis: Analysis-of-variance models and multi-group 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 cultural anthropology, but in profoundly different ways.

While 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. There has not even been firm consensus on what ethnography really is (Johnson 1990; Van Maanen 1988), and consensus is even more fleeting today.

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, and increased chances for various types of validity, triangulation, and potential for high levels of innovation and creativity. This is particularly true today, given the large number of tools available for assisting the researcher in managing and analyzing text data (see Wutich et al., 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 3.1 reveals 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

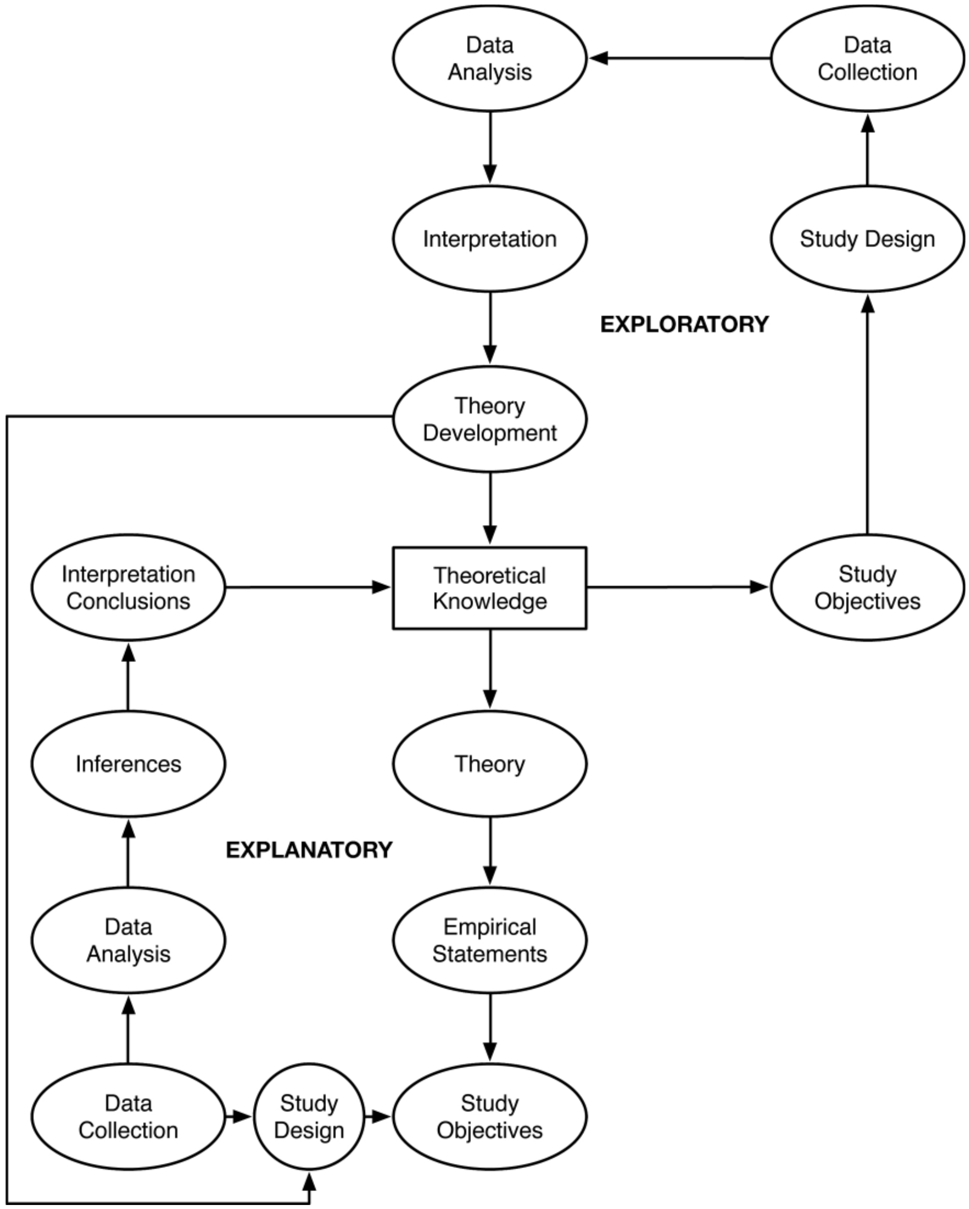

Figure 3.1. Relationship between exploratory and explanatory approaches within the overall ethnographic research process.

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 3.1 also shows that the research process involves a simultaneous concern for the development of empirical statements from theory (e.g., hypotheses), the operationalization of theoretical concepts (e.g., meaningful and reliable measures), design (e.g., groups to be studied), data collection (e.g., qualitative vs. quantitative), and data analysis (e.g., multiple regression and text analysis). Theoretical knowledge is derived either from earlier studies or from exploratory work. Which theoretical concepts are measured and which are bracketed, the levels at which theoretical concepts are measured (e.g., 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).

Most importantly, Figure 3.1 illustrates the complementary relationship between exploratory and explanatory approaches in ethnographic (and more generally scientific) investigations. Theory must, at some point, be derived from observations in the world. Thus, exploratory, inductive inquiry is crucial for theory development. Darwin built his theory of natural selection on his systematic (and frequently qualitative) description of organisms around the world. The modern theory of prion diseases arose in part from systematic exploratory investigation of South Fore horticulturalists in Papua New Guinea, a group afflicted with what is now known to be a prion disease. In addition to developing new hypotheses and constructing theories, 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 non-probability sample. Today, there are many tools for conducting systematic exploratory research in cultural anthropology, including sampling strategies to increase the representativeness of one's descriptions, behavioral observation and interviewing techniques for collecting valid and reliable information, and techniques for analyzing qualitative and quantitative data.

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 possible description of a situation or assessment of a theoretical framework or parts of it, given a set of realistic constraints (e.g., cost, scope, geographical setting, etc.). The purpose of research design is to ward off as many threats to validity as possible. In the case of exploratory descriptive work, this means ensuring that one's findings are reliable and valid. In the case of explanatory designs, this also means considering and assessing alternative explanations for one's findings.

Research design 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. Recognition of these limitations doesn't invalidate a study's results. Rather, it creates an open forum that can contribute a lot to important theoretical and methodological debates. Without 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.

Here we outline two key considerations when planning a study. The first is choosing appropriate comparison groups and interpreting differences. The second is how to sample groups and participants to make one's findings suitably representative and generalizable.

Planning and Interpreting Comparisons

Many questions of interest to cultural anthropologists are comparative. How and why do people around the world differ in how they think and act? What explains diversity in social, economic, and political systems? How does exposure to a specific economic, social, or political system, or one's position within such systems influence thought, behavior, and well-being? How do these systems change over time, what causes these changes, and what consequences do these changes have for people on the ground? These are all questions that require comparison—across groups, over time, across individuals from different groups. Indeed, it is difficult to find research questions in anthropology that do not require some, perhaps implicit, comparison between groups. For this reason, selecting appropriate comparison groups is an important part of many research designs.

Experiments are a relatively extreme example of making comparisons that can help us think about comparisons in general. In a true experiment, a researcher randomly assigns individuals to one of two (or more) treatments (e.g., bariatric surgery or not), and compares an outcome between these two groups (e.g., weight change). Ideally, the random assignment means that the only difference between the two randomly chosen groups is the different treatment they received. Thus, any difference in outcome between these two groups should have resulted from the treatment. For this reason, experiments can help rule out a number of alternative explanations for differences in outcomes between groups, including the effects of extraneous factors (i.e., unmeasured variables that might affect the dependent variable), the effects of selection (i.e., comparison groups differ because of the way they were selected and not due to the treatment), the effects of reactive measurement (i.e., the measurement procedure itself caused a change in the dependent variable), or interaction effects involving selection (i.e., when selection interacts with other factors to create erroneous findings).

These and other sources of error are also potential rival explanations, and randomized experiments are best at eliminating these rival explanations. Although designs of this type are often impossible in anthropological fieldwork, the principles of experimentation are instructive and are a guide for understanding potential sources of error when comparing groups, even in a non-laboratory setting. We borrow terminology from Kleinbaum et al. (1982) in constructing a typology of research designs. Included are experiments, quasi-experiments, observational study designs, and what we refer to as natural experiments.

Experiments involve the random allocation of subjects to different treatments or conditions and afford the most control over distorting effects from extraneous factors. A quasi-experiment compares groups exposed to different conditions or treatments but lacks random assignment of group members. Nonrandom assignment lays any comparison between groups 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.

Table 3.1 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 (e.g., Robson 2002). 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 (e.g., 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

Table 3.1. Examples of Basic Research Designs Relevant to Anthropologists

Observational Designs	
Cohort Study	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	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	For some study factor (like an outcome variable), it 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	A variant of the cross-sectional design in which a treatment group(s) (i.e., 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.
Quasi-Experimental Designs	
One group posttest only 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	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	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	One experimental group in which a series of observations is made both prior to some treatment and after the treatment.

is often difficult to manipulate independent variables directly while still preserving aspects of context that anthropologists hold on to dearly. And, traditionally, the most common designs used by anthropologists have been observational. However, 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, the fish aggregrating structures—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 piers (with the devices) and the control piers (without the devices). Johnson and Murray compared and determined catch rates.

Whether a study is observational, experimental, or quasi-experimental, most research designs in the explanatory mode are comparative. Experimental designs compare treatment and control groups. Longitudinal studies compare individuals and groups at different time points. Cross-cultural studies compare behavior across populations. In anthropological fieldwork, these designs and others can be used in tandem to test or explore components of a theory. For example, in their study of preschool children, Johnson et al. (1997) used a longitudinal cross-sequential design (also known as a panel study or cohort study), which involved periodic interviews and observations of a cohort of preschool children carried out over the course of the year. By doing this, they were able to make comparisons between children at one point in time and comparisons of the same child with him- or herself over time.

The importance of comparative thinking in ethnographic work cannot be overemphasized. Discussing common-sense knowing in evaluation research, Campbell (1988) describes the challenge of assessing change without careful attention to study design and comparative cases:

> 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)

The purpose of experimental design is to ward off threats to validity, although this is not as straightforward as it sounds. There are several types of validity—face, construct, statistical conclusion, internal, external, and so on. In one way or another, various

Table 3.2. Threats to Internal Validity in Quasi-Experimental Designs

History	Change due to unmeasured or unobserved factors (spuriousness)
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 is 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 (e.g., threats) that affect performance not a part of the treatment

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, we 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 3.2 and 3.3 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—requiring study variables to covary but without spurious or unintended causes. 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 3.2 deal with extraneous factors that may account for the presence or absence of a hypothesized effect (thus, contrasting 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 3.3. Threats to External Validity in Quasi-Experimental Designs

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 Campbell (1979) detail how each of the quasi-experimental designs in Table 3.1 are better or worse at dealing with each of the threats to validity that are found in Tables 3.2 and 3.3. For example, the pretest/posttest nonequivalent groups design controls for some internal threats to validity but does not do well at 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 affecting internal validity include sampling error (i.e., chance), nonresponse, the use of imprecise measures, data recording errors, informant inaccuracies, and interviewer effects (see Bernard 2011; Pelto and Pelto 1978). Careful attention to sampling, whether probabilistic (Babbie 1990) or non-probabilistic (Johnson 1990; Guest, this volume), 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 actually measure what they are intended to measure? Attention and concern with the potential sources of error, whether stemming from how the study was designed, how the data were collected (e.g., face-to-face interviews or mail-out surveys), or how the data were analyzed (e.g., statistical conclusion validity) will help lead to the production of solid evidence. However, a particularly important source of errors is due to problems with proper sampling and informant selection.

Sampling

With limited time and resources, a researcher can only interview or observe a select set of events, communities, or individuals. However, the researcher may also hope that the findings from this sample reflect what one would find in a much broader population. Steps taken to ensure that a sample is representative of that larger population are crucial for making such generalizations. Or a researcher may want to catalogue the maximum diversity of views in a population rather than getting a representative portrait. In this case, a researcher would aim to select individuals expected to be maximally diverse in their views. Depending on a researcher's specific goals, there are many sampling strategies that help meet those goals.

When generalization to a target population is the objective, you should strive to define a sampling universe or frame from which you will select individuals, events, or communities and then develop a system for selecting a sample that accurately reflects the broader population. This usually entails a random sample of some kind, but can also involve stratifying strategically on sex, age, or ethnicity to guarantee appropriate representation of different groups. There is a vast literature on sampling theory and random sampling procedures, including discussions of sample sizes (see, e.g., Bernard [2011, 2012] for a summary, and the classic text by 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 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 range 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), for example, used a static-group comparative design sampling across a range of groups that were hypothesized to vary with respect to their environmental values. Kempton et al. interviewed a range of informants from members of Earth First! (a radical environmentalist group) to 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 (e.g., random selection of locations in which to intercept respondents).

Non-probability sampling methods have come to be associated with qualitative approaches or for the selection of ethnographic informants, particularly key informants or consultants (Johnson 1990; Miles and Huberman 1994; Werner and Schoepfle 1987; Guest, this volume). 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 take into account violations of assumptions for the particular statistical test to be employed (e.g., independence of observations or random sample from a population). Such matters are particularly germane for observational designs using various social network approaches (see Borgatti et al. 2013).

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 (e.g., 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 (Edgington and Onghena 2007; Johnson and Murray 1997; Noreen 1989). 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.

RESEARCH DESIGN IN ANTHROPOLOGICAL PRACTICE: SYSTEMATIC RESEARCH STRATEGIES

The following examples illustrate some of the issues discussed so far. These examples show how the interplay of exploratory and explanatory approaches isa crucial for the development and testing of theory and can also aid in guarding against threats to validity (Robson 2002). This is particularly evident in what has more recently been the emphasis on the use of mixed methods in anthropological research, and in the social sciences more generally (see, e.g., *Journal of Mixed Methods Research*). Here we describe seven different kinds of study design used in cultural anthropology today—exploratory designs, comparisons of individuals within societies, two-community comparisons, case control designs, large-scale cross-population comparisons, field experiments, and longitudinal designs.

Using Exploratory Research to Identify Locally Relevant Categories

As seen in Figure 3.1, exploratory research 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 of the prior 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 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 find-

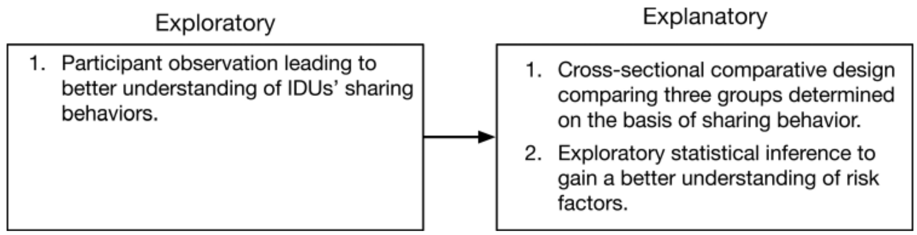

Figure 3.2. Overall design framework for the Koester study.

ings 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 3.2). 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.

Comparing Individuals in a Society: Experts, Novices, 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 (i.e., 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 variations 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.

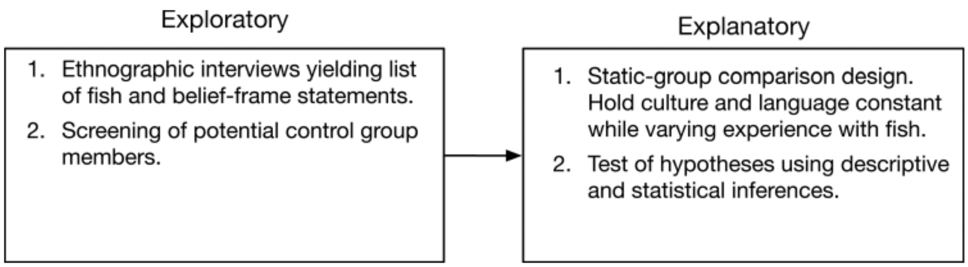

Figure 3.3. Overall design framework for the Boster and Johnson study.

Thus, the groups to be compared consisted of five groups of 15 subjects, four comprised 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 (1989) 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 or the uses of fish. Using statistical and descriptive inference, Boster and Johnson 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 3.3).

Lack of random assignment of subjects to treatment and control groups and pretest observations limit the ability to make causal inferences in this case. 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.

Case Control Design: Folk Illness and Disease Risks

In a recent paper, Baer et al. (2012) used a case-control design to study the relationship between the ethnomedical diagnosis of a folk illness and disease risks. The primary objective of the research was to understand the relationship between the ethnomedical diagnosis of the folk illnesses *susto* (fright) and/or *nervios* (nerves) and the risk for developing type 2 diabetes. Understanding the relationship could be useful in health-care provider screening for diabetes.

The authors selected Guadalajara, Mexico's second-largest city, as the study site, based on their earlier work that showed a widespread belief in *susto* and *nervios* in this region. In addition, type 2 diabetes is becoming an increasing problem in Mexico more generally. The research questions emerged from the earlier work, some of it descriptive and exploratory, stemming from interviews eliciting the causes, symptoms, and treatments for *susto*. In these interviews, respondents kept talking about a perceived

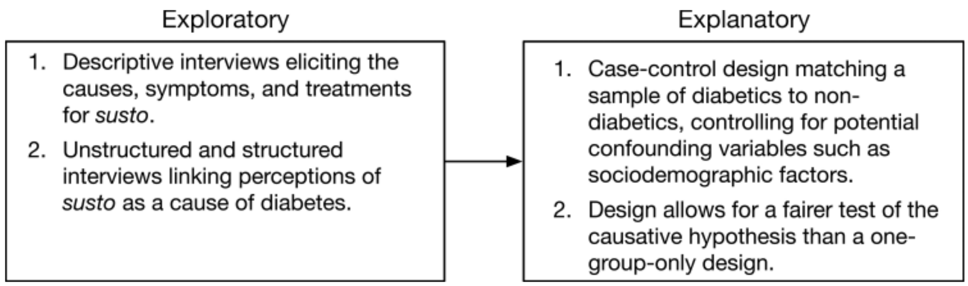

Figure 3.4. Overall design framework for the Baer et al. study.

link between *susto* and diabetes (see Figure 3.4). Further, previous studies using one-group-only designs (looking at only diabetics) had suggested a relationship between folk illness and diabetes.

Respondents were drawn from a family practice clinic serving approximately 110,000 patients with mostly working-class backgrounds. The first sample group involved patients who went to the clinic, had been diagnosed with type 2 diabetes following their thirtieth birthday, and had been diagnosed more than one year ago. Patients with type 1 diabetes were excluded from the study. The comparison group consisted of patients at the clinic who were 30 years old or more who had not been diagnosed with type 2 diabetes at the time of the study. Both groups were asked questions about their perceptions of the link between *susto*/*nervios* and diabetes and whether they have ever suffered from *susto*/*nervios*. The type 2 diabetes group was also asked whether they believed *susto*/*nervios* was the cause of their diabetes. Finally, blood glucose levels were measured for all study participants.

The authors recruited 836 patients, 811 of whom were interviewed. One potential problem was the possibility that the folk illness might have occurred following the development of diabetes rather than before its onset. To minimize this problem, the study limited the period for type 2 diabetes diagnosis to less than five years before the interview. This reduced the diabetic comparison group from 811 to 239 patients. As in the previous *susto* example, to help control for any confounding factors such as educational background and social class, members of the non-diabetic group, who were drawn from the patient population who went to the clinic for general health problems such as colds, colitis, bronchitis, and so on, were matched to the diabetic group on as many sociodemographic factors as possible, although there were some differences (e.g., gender, marital status).

In comparisons of the diabetic and non-diabetic groups, there were no significant differences in the prevalence of *susto*; nor was the prevalence of *susto* significantly higher for those with undiagnosed diabetes. The same was true for *nervios*. The study design, using comparison groups and matched controls, helped cast doubt on the hypothesis that *susto* was an underlying cause of type 2 diabetes. To find that out, we would need either a true experiment or longitudinal design such as a cohort or panel study to determine if the risk factor preceded the outcome. However, the importance of this research is that the study design allowed for a fairer test of the hypothesis concerning *susto* as a risk factor in the development of diabetes than that afforded by a simple one-group-only design.

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) now classic investigation of treatment choice in two Mexican communities is an example of a static-group comparison in which 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 consideration of competing hypotheses—the hallmark of good research design—and the testing of the primary hypothesis as 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 Pichataro 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 triad data and what they call term-frame data on informants' perceived similarity of illnesses.

Young and Garro (1982) 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 groups.

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 con-

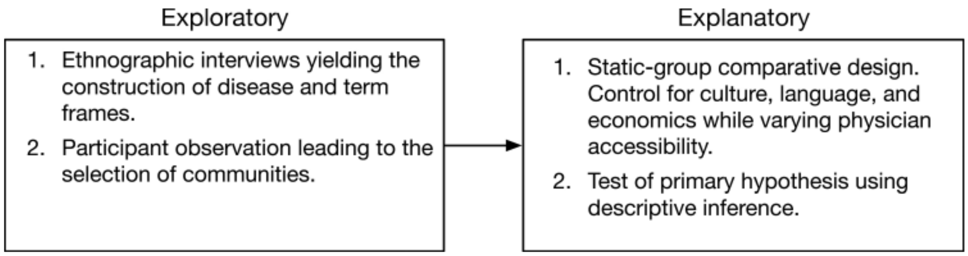

Figure 3.5. Overall design framework for the Young and Garro study.

sequence 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 originally 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 3.5). 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.

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 to control for confounding variables and then subjected their informants to the treatments of interest. That said, in a latter section, we review how Paul Farmer attempted to mitigate such threats to validity through a field experiment in Haiti asking similar questions. 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 (1982), 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.

Large-scale Cross-population Comparison

People around the world differ remarkably in how they think, talk, and behave. Understanding the roots of this striking cross-population variation has been one of the core concerns of anthropology since its inception. Why are some groups more collectivist than others? What explains differences in norms of fairness? Why do different groups have different forms of kinship and social organization? A long line of research in anthropology has used comparative, cross-population analysis to answer these questions, sometimes focusing on population variation in specific regions of the world (e.g., Edgerton 1971) and sometimes using worldwide samples. More recently, a number of researchers have developed novel forms of data collection for cross-population analysis. An exemplar in this emerging tradition is research by Joe Henrich and colleagues (2006) that examines how people make decisions to share with others in different cultural settings.

The research was a response to work in behavioral economics and evolutionary psychology that tried to explain why people share substantial amounts of money with anonymous strangers, even though the recipients will never find out who gave them the cash. To explain this phenomenon, which had largely been observed among U.S., European, and college populations, researchers argued that this was a psychological relic from our evolutionary past. They posited that we evolved in an environment where we rarely encountered strangers, so we give to strangers today because such an anonymous situation doesn't make sense to us. The prediction from this explanation is that all humans should show similar biases toward sharing with strangers.

To test this prediction, Henrich and colleagues (2006) brought the same experiments used by behavioral economists to small-scale societies around the world. People were asked to allocate 10 real dollars (or the local equivalent) to another anonymous person. The researchers found striking cross-population variation in people's offers, all the way from stingy (giving next to nothing) to hyper-fair (giving more than half). These findings raised serious questions about explanations for such giving based on a pan-human–evolved psychology.

The researchers also found in two independent studies (Henrich et al. 2010; Hermann et al. 2008) that the degree to which a population depended on markets (% of calories purchased) accounted for a large part of variation in how much people shared with strangers. Specifically, populations interacting more with markets showed more fair offers on average, a finding more consistent with adaptation to local social and cultural environments (Henrich et al. 2010). Since that time, similar studies have examined cross-cultural differences in willingness to punish others (Henrich et al. 2010; Herrman et al. 2008) and willingness to violate a rule to help friends and community members (Hruschka 2010).

Figure 3.6 shows the overall design framework for this study. Such research is challenging. It requires long-term collaboration across a number of field sites and it can be difficult to identify measures that are meaningful across diverse societies. If you are studying friendship, for example, what word would you use for friend in the language of each society studied (Hruschka 2010), or does the measure used by the researcher have the same meaning (Gelfand et al. 2011)? However, rather than being a reason to avoid such explorations, dealing with these issues can teach us a great deal about where human populations diverge and where they are similar.

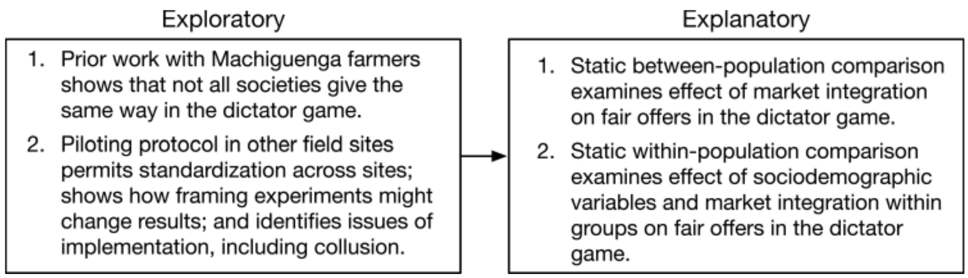

Figure 3.6. Overall design framework for the Henrich et al. studies.

Field Experiments

Observational data can help us understand how people think and act in different social and cultural situations. However, it is sometimes difficult to infer from observational data alone what causes people to do one thing or another. Suppose that we find that poor individuals are less likely to take medical treatment for tuberculosis than are wealthy individuals. We might conclude that lack of resources affects people's behavior. It is also possible, though, that underlying variables, such as cultural beliefs about sorcery, might be the real drivers. One way to address such potential confounding by another variable is to conduct a field experiment. A field experiment randomly assigns people to a treatment condition and a control condition so that they ideally differ only on whatever is included in the treatment. Then, if we see a difference in behavior between the two groups, we can be fairly confident that the difference resulted from the treatment (rather than other unmeasured variables).

Paul Farmer conducted just such an experiment at a clinic in rural Haiti to find out how to improve adherence to tuberculosis treatment (Farmer 1999). Treating tuberculosis requires an extended regime of antibiotics. If not followed completely, the patient's condition can deteriorate and drug-resistant strains of TB can result. Despite the availability of treatment at the clinic, people were still dying of the disease, and practitioners proposed two theories. Community health workers pointed to economic barriers to completing regular treatment as well as problems of treating TB when people are malnourished. Staff doctors and nurses, on the other hand, argued that patients often stopped taking pills in part because they believed TB was caused by sorcery rather than microbes.

To assess which factor would be most important in improving care, Farmer assigned TB patients to two groups of 50 individuals each (Farmer 1999, 219). One group got free treatment. The other got free treatment, but members were also eligible for a stipend of $30 per month for the first three months, $5 in travel expenses for coming to the clinic, nutritional supplements, and daily visits from a community health worker for the first month. To assess any potential impact of sorcery beliefs on treatment, each participant was interviewed about how he or she perceived the causes and treatments for TB (see Figure 3.7).

The two groups differed very little in terms of age and sex, although economic indicators suggested that the treatment group was slightly poorer than the control group. Nearly all participants felt that sorcery played some role in their illness. However, a year after the study began, there were big differences in treatment outcomes. The treatment group had a 100% cure rate, whereas the control group had a cure rate of less than

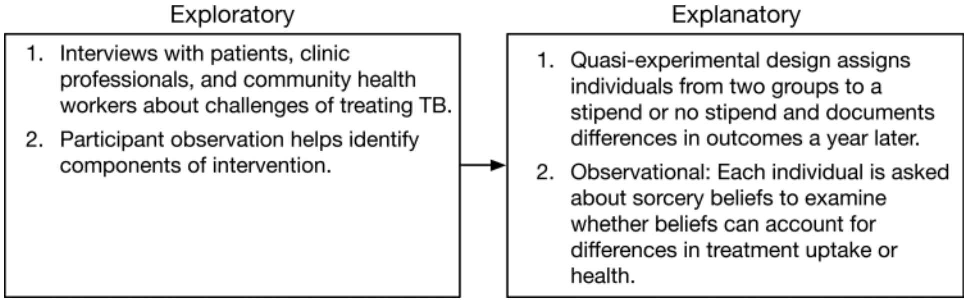

Figure 3.7. Overall design framework for the Farmer study.

50%. The treatment group showed much greater weight gain (10 pounds vs. 2 pounds), and more members of the treatment group returned to work after one year of treatment (92% vs. 48%). After 18 months, one person from the treatment group had died, but not from TB. In the control group, six patients had died.

These dramatic results indicated that small changes in economic barriers (and perhaps social influence from community health workers) could dramatically change treatment behavior and health outcomes. Moreover, sorcery beliefs played little role in the outcomes. Armed with such powerful results, the clinic began treating all patients with the combined package (Kidder 2004).

Such field experiments are always subject to constraints of the field setting. For example, Farmer did not assign patients completely at random. Individuals in the treatment group were selected from a single region, whereas individuals in the control group were selected from patients not from that sector. However, checks on other variables for comparability across groups provide additional guarantees that the assignment was effectively random.

Large-scale field experiments can be incredibly costly and time consuming to implement. However, their benefits in determining causality have led social science researchers to make increasing use of them in areas of critical importance for health and well-being (Banarjee and Duflo 2011). The same principle of randomization can also be applied at shorter time scales to understand how priming individuals with information or resources changes their behavior or responses. How does framing someone as a friend or a stranger change someone's willingness to help or trust that person (Cronk 2007; Hruschka 2010) or how does thinking about a higher power make us more or less likely to share with others (Shariff and Norenzayan 2007)? Such experiments complement the findings and hypotheses generated by observational data.

Repeated Measures Design: The Evolution of Network Structure

As discussed earlier in this chapter, experimental and longitudinal designs are necessary for determining causal relationships. Whereas true experiments are not always possible, or for that matter desirable, longitudinal designs are quite appropriate for testing hypotheses in many ethnographic contexts. Ethnography, by its very nature, involves extended periods of time in one or more given field sites. Thus, they provide the opportunity to collect data over time and, possibly, across multiple groups. In addition, unlike the requisite artificial conditions imposed by experiments (i.e., manipulation of independent variable while controlling for all other variables), longitudinal research in field settings allows for the study of social and behavioral phenomena in situ.

Although anthropologists tend to spend months, sometimes years, in the field collecting data, few of them use panel or cohort designs. Gravlee et al. (2009) carefully document this and point out that anthropologists rarely employ systematic longitudinal designs in their work. The article provides a good discussion of the ins and outs of systematic panel and cross-sectional panel designs, informative discussions of sources of potential errors in such designs, and anthropological examples of work in this vein.

In research that spanned well over eight years, Johnson, Boster, and Palinkas (2003; Johnson, Palinkas, and Boster 2003; Palinkas, Johnson, and Boster 2004; Palinkas et al. 2004) were interested in the influence of the emergence of informal social roles on the evolution of small-group network structure and ultimately on group well-being and performance. In essence, they conducted a kind of natural experiment by studying populations that were quite isolated over an extended period of time so that the emergent properties of the groups could be more readily studied with less interference from extraneous influences. Their basic theoretical proposition was that groups that formed more cohesive networks (i.e., core–periphery networks) would have higher morale and would be more productive. Furthermore, the presence of various informal social roles helps facilitate social cohesion. To test these hypotheses, they went to Antarctica to study small-group dynamics at polar research stations.

The design of the study was quite simple and involved repeated interviews with winter-over crew members at Antarctic stations during the winter (an interview each month over the eight–nine months of the winter). Although a single group at a single station could have been studied over time (i.e., a single case-study design), it would have posed a number of problems for adequately testing the hypotheses of interest. To be able to draw conclusions at the group level of analysis, as opposed to simply the individual-actor level, there needed to be observations on multiple cases (i.e., across multiple years). Therefore, the design involved three observations at a station (i.e., the group level) that included the study of three separate groups per station each of a year's duration. In addition, the researchers were interested in the role of culture in group formation. They therefore included five different cultures in the overall design. These included the Americans at the Amundsen-Scott South Pole Station, the Russians at Vostok Station, the Chinese at the Great Wall Station, the Poles at the Arktowski Station, and the Indians at the Maitri Station. This is similar to the rationale for comparisons discussed in the cross-population comparison section earlier and allowed for the study of variation both within and between the various stations and cultures and facilitated the study of social networks at the group level, the dyadic or tie level, and the individual-actor level (see Borgatti et al. [2013] for a review).

Prior to the data collection at the stations, Johnson et al. (2003) conducted a series of semi-structured interviews with former winter-overs to elicit informal social roles recognized by the winter-overs themselves. In addition, the social network questions were developed in consultation with winter-over crew members early in the study to maximize cultural understanding and appropriateness of the social network questions (see Figure 3.8). These early exploratory interviews were important for producing a condensed survey instrument to be administered during the monthly winter

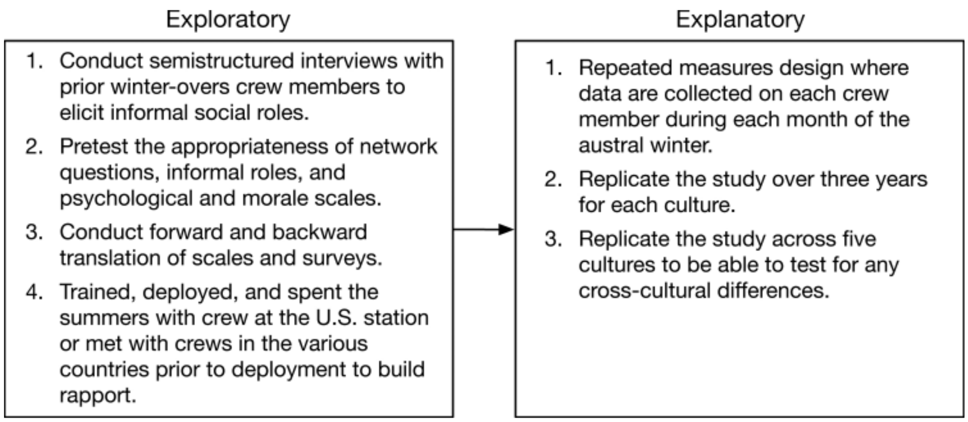

Figure 3.8. Overall design framework for the Johnson et al. study.

interviews. The researchers did not want to risk over-burdening the crew members with questions not central to the testing of their research hypotheses (i.e., not a data-fishing expedition). A shorter, more theoretically relevant survey instrument would help ensure a sustained level of study participation over the course of the winter. A huge threat in repeated measures designs involves individuals dropping out of the study prematurely (Gravlee et al. 2009). Missing data are also critically important for social network studies.

The research found that winter-over groups in which various informal social roles (e.g., clown, expressive leaders) emerged over the course of the winter had more cohesive social networks and higher morale and individual level psychological well-being. In addition, these findings were consistent across the five cultures, suggesting some degree of cultural universality in human group dynamics. Since the researchers were interested in social group dynamics, particularly in aspects of emergent group properties, a longitudinal design was essential for them to be able to draw the conclusions they did.

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 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 and valid 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" (1996, 13).

REFERENCES

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

Agar, M. 1996. Schon Wieder? Science in linguistic anthropology. *Anthropology News* 37(1): 13.*r*. January.

Babbie, E. 1990. *Survey research methods*, 2nd ed. Belmont, CA: Wadsworth.
Baer, R. D., S. C. Weller, J. G. Garcia, and A. L. S. Rocha. 2012. Ethnomedical and biomedical realities: Is there an epidemiological relationship between stress-related folk illnesses and type 2 diabetes? *Human Organization* 71: 339–47.
Banarjee A. V., and E. Duflo. 2011. *Poor economics: A radical rethinking of the way to fight global poverty*. New York: Public Affairs Press.
Behar, R. 1993. *Translated woman*. Boston: Beacon Press.
Bernard, H. R., ed. 1998. *Handbook of methods in cultural anthropology*. Walnut Creek, CA: Sage.
Bernard, H. R. 2011. *Research methods in anthropology: Qualitative and quantitative approaches*, 5th ed. Walnut Creek, CA: AltaMira.
Bernard, H. R. 2012. *Social research methods: Qualitative and quantitative approaches*, 2nd ed. Thousand Oaks, CA: Sage.
Boas, F. 1920. The methods of ethnology. *American Anthropologist* 22: 311–21.
Boellstorff, T. 2008. *Coming of age in Second Life: An anthropologist explores the virtually human*. Princeton, NJ: Princeton University Press.
Borgatti, S., M. Everett, and J. C. Johnson. 2013. *Analyzing social networks*. London: Sage.
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: 866–89.
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. 1988. Qualitative knowing in action research. In *Methodology and epistemology. for social science: Selected papers*, ed. E. S. Overman, 360–76. Chicago: University of Chicago Press.
Cook, T. D., and D. T. Campbell. 1979. *Quasi-experimentation: Design and analysis for field settings*. Chicago: Rand McNally.
Cronk, L. 2007. The influence of cultural framing on play in the trust game: A Maasai example. *Evolution and Human Behavior* 28: 352–58.
Denzin, N. K., and Y. S. Lincoln. 1994. Entering the field of qualitative research. In *Handbook of qualitative research*, ed. N. K. Denzin and Y. S. Lincoln, 1–19. Thousand Oaks, CA: Sage.
Edgerton, R. 1971. *The individual in cultural adaptation: A study of four East African peoples*. Berkeley: University of California Press.
Edgington, E., and P. Onghena. 2007. *Randomization tests*, 4th ed. London: Chapman and Hall/CRC.
Ellen, R. F. 1984. Introduction. In *Ethnographic research: A guide to general conduct*, ed. R. F. Ellen, 1–12. London: Academic Press.
Farmer, P. 1999. *Infections and inequalities*. Berkeley: University of California Press.
Freeman, D. 1983. *Margaret Mead and Samoa: The making and unmaking of an anthropological myth*. Cambridge, MA: Harvard University Press.
Gelfand, M. J., J. L. Raver, L. Nishii et al. 2011. Differences between tight and loose cultures: A 33-nation study. *Science* 332: 1100–104.
Gravlee, C. C., D. P. Kennedy, R. Godoy, and W. R. Leonard. 2009. Methods for collecting panel data: What can cultural anthropology learn from other disciplines? *Journal of Anthropological Research* 69: 453–83.
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: 451–62.
Henrich, J., R. McElreath, A. Barr et al. 2006. Costly punishment across human societies. *Science* 312: 1767–70.
Henrich, J., R. McElreath, A. Barr et al. 2010. Markets, religion, community size, and the evolution of fairness and punishment. *Science* 327: 1480–84.

Hermann, B., C. Thöni, and S. Gächter. 2008. Antisocial punishment across societies. *Science* 319: 1362–67.

Ho, K. 2009. *Liquidated: An ethnography of Wall Street*. Durham, NC: Duke University Press.

Hruschka, D. J. 2010. *Friendship: Development, ecology and evolution of a relationship*. Berkeley: University of California Press.

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: 187–211.

Johnson, J. C. 1990. *Selecting ethnographic informants*. Qualitative Research Methods Series, Vol. 22. Thousand Oaks, CA: Sage.

Johnson, J. C., C. Avenarius, and J. M. Weatherford. 2006. The active participant observer: Applying social role analysis to participant observation. *Field Methods* 18: 111–34.

Johnson, J. C., J. S. Boster, and L. Palinkas. 2003. Social roles and the evolution of networks in isolated and extreme environments. *The Journal of Mathematical Sociology* 27: 89–122.

Johnson, J. C., M. Ironsmith, A. L. Whitcher et al. 1997. The development of social networks in preschool children. *Early Education and Development* 8: 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*, ed. R. Pollnac and J. Poggie, 143–58. Kingston, RI: ICMRD.

Johnson, J. C., L. A. Palinkas, and J. S. Boster. 2003. Informal social roles and the evolution and stability of social networks. In *Dynamic social network modeling and analysis*, ed. R. Brieger, K. Carley, and P. Pattison, 121–32. Washington, DC: The National Academies Press.

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, MA: MIT Press.

Kidder, T. 2004. *Mountains beyond mountains*. New York: Random House.

Kincheloe, J. L., and P. L. McLaren. 1994. Rethinking critical theory and qualitative research. In *Handbook of qualitative research*, ed. N. K. Denzin and Y. S. Lincoln, 138–58. Thousand Oaks, CA: Sage.

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 IDUs. In *AIDS, drugs and prevention: Perspectives on individual and community action*, ed. T. Rhodes and R. Hartnoll, 133–48. 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: 202–7.

LeVine, R. A. 1973. Research design in anthropological field work. In *A handbook of method in cultural anthropology*, ed. R. Naroll and R. Cohen, 183–95. New York: Columbia University Press.

Lowie, R. H. 1937. *The history of ethnological theory*. New York: Rinehart.

Marcus, G. 1995. Ethnography in/of the world system: The emergence of multi-sited ethnography. *Annual Review of Anthropology* 24: 95–117.

Miles, M. B., and A. M. Huberman. 1994. *Qualitative data analysis*, 2nd ed. Thousand Oaks, CA: Sage.

Moran, E. F., ed. 1995. *The comparative analysis of human societies: Toward common standards for data collection and reporting*. Boulder, CO: 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.

Palinkas, L. A., J. C. Johnson, and J. S. Boster. 2004. Social support and depressed mood in isolated and confined environments. *Acta Astronautica* 54: 639–47.

Palinkas, L. A., J. C. Johnson, J. S. Boster et al. 2004. Cross-cultural difference in psychosocial adaptation to isolated and confined environments. *Aviation, Space, and Environmental Medicine* 75: 973–80.

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*, 2nd ed. Cambridge: Cambridge University Press.

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

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

Robson, C. 2002. *Real world research: A resource for social scientists and practitioner-researchers*, 2nd ed. Hoboken, NJ: John Wiley & Sons.

Scheper-Hughes, N. 2004. Parts unknown: Undercover ethnography of the organs-trafficking underworld. *Ethnography* 5: 29–73.

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

Sechrest, L. 1973. Experiments in the field. In *A handbook of method in cultural anthropology*, ed. R. Naroll and R. Cohen, 196–209. New York: Columbia University Press.

Shariff, A. F., and A. Norenzayan. 2007. God is watching you: Priming God concepts increases prosocial behavior in an anonymous economic game. *Psychological Science* 18: 803–9.

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*, ed. R. Naroll and R. Cohen, 210–19. New York: Columbia University Press.

Stinchcombe, A. L. 1987. *Constructing social theories*. Chicago: University of Chicago Press.

Tsing, A. 2005. *Friction: An ethnography of global connection*. Princeton, NJ: Princeton University Press.

Tyler, S. A. 1991. A post-modern instance. In *Constructing knowledge: Authority and critique in social science*, ed. L. Nencel and P. Pels, 78–95. London: Sage.

Urry, J. 1984. A history of field methods. In *Ethnographic research: A guide to general conduct*, ed. R. F. Ellen, 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.

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

Whyte, W. F. 1984. *Learning from the field: A guide from experience*. Newbury Park, CA: Sage.

Williams, T. R. 1967. *Field methods in the study of culture*. 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: 1453–65.

Zabusky, S. E. 1995. *Launching Europe: An ethnography of European cooperation in space science*. Princeton, NJ: Princeton University Press.