Rosenthal, R. (1984). *Meta-analytic procedures for social research.* Beverly Hills, CA: Sage.

Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74,* 166-169.

Schneider, A. L., & Darcy, R. E. (1984). Policy implications of using significance tests in evaluation research. *Evaluation Review, 8,* 573-582.

Subkoviak, M. J., & Levin, J. R. (1977). Fallibility of measurement and the power of a statistical test. *Journal of Educational Measurement, 14,* 47-52.

Yeaton, W. H., & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: Strength, integrity, and effectiveness. *Journal of Consulting and Clinical Psychology, 49,* 156-167.

3

Designing a Qualitative Study

Joseph A. Maxwell

Contrary to what you may have heard,
qualitative research designs do exist.
—Miles and Huberman
Qualitative Data Analysis, 1994

As the above quote suggests, "qualitative research design" has often been treated as an oxymoron. One reason for this is that the dominant, quantitatively oriented models of research design presented in textbooks fit poorly with the ways that most qualitative researchers go about their work (Lincoln & Guba, 1985). These models usually treat "design" in one of two ways. Some take designs to be fixed, standard arrangements of research conditions and methods that have their own coherence and logic, as possible answers to the question, What research design are you using? For example, a randomized, double-blind experiment is one research design; an interrupted time series design is another. Qualitative research lacks any such elaborate typology into which studies can be pigeonholed.

Other models present design as a logical progression of stages or tasks, from problem formulation to the generation of conclusions or theory, that are necessary in planning or carrying out a study. Although some versions of this approach are circular or iterative (see, for example, Bickman, Rog, & Hedrick, Chapter 1, this volume), so that later steps connect back to earlier ones, all such models are linear in the sense that they are made up of one-directional

sequences of steps that represent what is seen as the optimal order for conceptualizing or conducting the different components or activities of a study.

This view of design does not adequately represent the logic and process of qualitative research. In a qualitative study, the activities of collecting and analyzing data, developing and modifying theory, elaborating or refocusing the research questions, and identifying and dealing with validity threats are usually going on more or less simultaneously, each influencing all of the others. In addition, the researcher may need to reconsider or modify any design decision during the study in response to new developments or to changes in some other aspect of the design. Grady and Wallston (1988) argue that applied research in general requires a flexible, nonsequential approach and "an entirely different model of the research process than the traditional one offered in most textbooks" (p. 10).

This does not mean that qualitative research lacks design; as Yin (1994) says, "Every type of empirical research has an implicit, if not explicit, research design" (p. 19). Qualitative research simply requires a broader and less restrictive concept of "design" than the traditional ones described above. Thus Becker, Geer, Hughes, and Strauss (1961), authors of a classic qualitative study of medical students, begin their chapter titled "Design of the Study" by stating:

> In one sense, our study had no design. That is, we had no well-worked-out set of hypotheses to be tested, no data-gathering instruments purposely designed to secure information relevant to these hypotheses, no set of analytic procedures specified in advance. Insofar as the term "design" implies these features of elaborate prior planning, our study had none.
>
> If we take the idea of design in a larger and looser sense, using it to identify those elements of order, system, and consistency our procedures did exhibit, our study had a design. We can say what this was by describing our original view of the problem, our theoretical and methodological commitments, and the way these affected our research and were affected by it as we proceeded. (p. 17)

For these reasons, the model of design that I present here, which I call an *interactive* model, consists of the components of a research study and the ways in which these components may affect and be affected by one another. It does not presuppose any particular order for these components, or any necessary directionality of influence; as with qualitative research in general, "it depends." One of my goals in this chapter is to try to point out the things that I think these influences depend *on*.

The model thus resembles the more general definition of *design* employed outside of research: "an underlying scheme that governs functioning, developing, or unfolding" and "the arrangement of elements or details in a product or work of art" (*Merriam-Webster's Collegiate Dictionary*, 1993). A good design, one in which the components work harmoniously together, promotes efficient and successful functioning; a flawed design leads to poor operation or failure.

This model has five components, each of which addresses a different set of issues that are essential to the coherence of your study:

1. *Purposes:* What are the ultimate goals of this study? What issues is it intended to illuminate, and what practices will it influence? Why do you want to conduct it, and why should we care about the results? Why is the study worth doing?
2. *Conceptual context:* What do you think is going on with the things you plan to study? What theories, findings, and conceptual frameworks relating to these will guide or inform your study, and what literature, preliminary research, and personal experience will you draw on? This component of the design contains the *theory* that you already have or are developing about the setting or issues that you are studying.
3. *Research questions:* What, specifically, do you want to understand by doing this study? What do you *not* know about the things you are studying that you want to learn? What questions will your research attempt to answer, and how are these questions related to one another?
4. *Methods:* What will you actually do in conducting this study? What approaches and techniques will you use to collect and analyze your data, and how do these constitute an integrated strategy?
5. *Validity:* How might you be wrong? What are the plausible alternative explanations and validity threats to the potential conclusions of your study, and how will you deal with these? Why should we believe your results?

These components are not radically different from the ones presented in many other discussions of qualitative or applied research design (e.g., LeCompte & Preissle, 1993; Lincoln & Guba, 1985; Miles & Huberman, 1994; Robson, 1993). What is distinctive in this model are the relationships among the components. The components form an integrated and interacting whole, with each component closely tied to several others, rather than being linked in a linear or cyclic sequence. The lines between the components in Figure 3.1 represent two-way connections of influence or implication. Although there are also connections other than those emphasized here (for example, between purposes and methods, and between conceptual context and validity), those shown in the figure are usually the most important.

The upper triangle of this model should be a closely integrated unit. Your research questions should have a clear relationship to the purposes of your study, and should be informed by what is already known about the things you are studying and the theoretical tools that can be applied to these. In addition, the purposes of the study should be informed both by current theory and knowledge and by what questions you can actually answer, and your choices of relevant theory and knowledge depend on the purposes and questions.

Similarly, the bottom triangle of the model should also be closely integrated. The methods you use must enable you to answer your research questions, and also to deal with plausible validity threats to these answers. The questions, in turn, need to be framed so as to take the feasibility of the methods

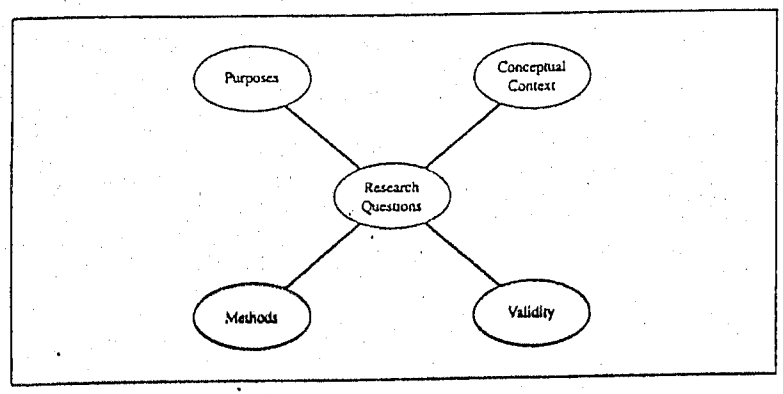

Figure 3.1. An Interactive Model of Research Design

and the seriousness of particular validity threats into account; in addition, the plausibility and relevance of particular validity threats depend on the questions and methods chosen. The research questions are the center or hub of the model; they connect the other four components into a coherent whole, and need to inform, and be responsive to, all of the other components.

There are many other factors besides these five components that should influence the design of your study; these include your research abilities, the available resources, perceived problems, ethical standards, the research setting, and the data and preliminary conclusions of the study. In my view, these are not part of the *design* of a study; rather, they either belong to the *environment* within which the research and its design exist or are *products* of the research. Figure 3.2 presents some of the environmental factors that can influence the design and conduct of a study.

I do not believe that there is one right model for qualitative or applied research design. However, I think that the model I present here is a useful one, for three main reasons:

1. It explicitly identifies as *components* of design the key issues about which decisions need to be made. These issues are therefore less likely to be ignored, and can be dealt with in a systematic manner.
2. It emphasizes the *interactive* nature of design decisions in qualitative and applied research, and the multiple connections among the design components.
3. It provides a model for the structure of a *proposal* for a qualitative study, one that clearly communicates and justifies the major design decisions and the connections among these (see Maxwell, 1996a).

Because a design for your study always exists, explicitly or implicitly, it is important to *make* this design explicit, to get it out in the open, where its strengths, limitations, and implications can be clearly understood. In the re-

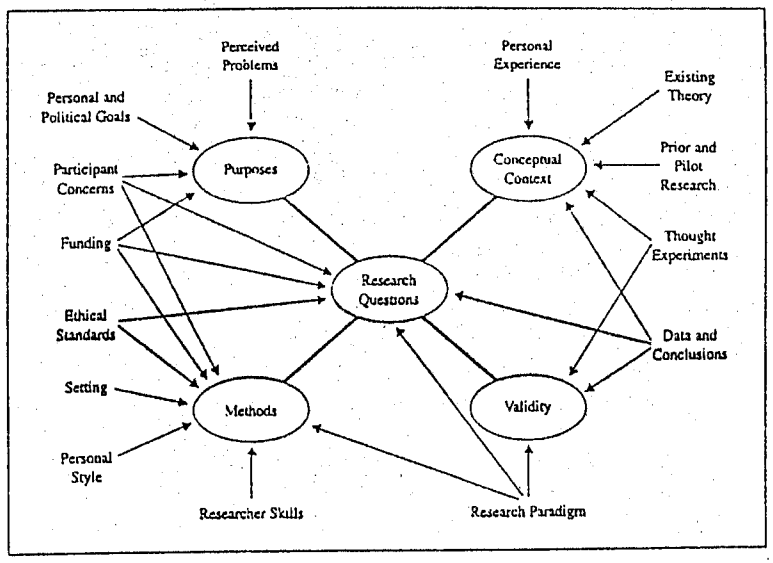

Figure 3.2. Contextual Factors Influencing Research Design

mainder of this chapter, I present the main design issues involved in each of the five components of my model, and the implications of each component for the others. I do not discuss in detail how to actually *do* qualitative research, or deal in depth with the theoretical and philosophical views that have informed this approach. For additional guidance on these topics, see the contributions of Fetterman (Chapter 16) and Stewart and Shamdasani (Chapter 17) to this handbook; the more extensive treatments by Patton (1990), Eisner and Peshkin (1990), LeCompte and Preissle (1993), Glesne and Peshkin (1992), Weiss (1994), Miles and Huberman (1994), and Wolcott (1995); and the encyclopedic handbooks edited by LeCompte, Millroy, and Preissle (1992) and by Denzin and Lincoln (1994). My focus here is on how to design a qualitative study that arrives at valid conclusions and successfully and efficiently achieves its goals.

Purposes: Why Are You Doing This Study?

Without a clear sense of the purposes of your research, you are apt to lose your focus and spend your time and effort doing things that won't contribute to these purposes. (I use *purpose* here in a broad sense, to include motives, desires, and goals—anything that leads you to do the study or that you hope to accomplish by doing it.) Your purposes help to guide your other design decisions, to ensure that your study is worth doing, that you get out of it what you want.

It is useful to distinguish among three kinds of purposes for doing a study: personal purposes, practical purposes, and research purposes. Personal purposes are those that motivate *you* to do this study; they can include a desire to change some existing situation, a curiosity about a specific phenomenon or event, or simply the need to advance your career. These personal purposes often overlap with your practical or research purposes, but they may also include deeply rooted individual desires and needs that bear little relationship to your "official" reasons for doing the study.

It is important that you recognize and take account of the personal purposes that drive and inform your research. Eradicating or submerging your personal goals and concerns is impossible, and attempting to do so is unnecessary. What *is* necessary, in qualitative design, is that you be *aware* of these concerns and how they may be shaping your research, and that you think about how best to deal with their consequences. To the extent that your design decisions and data analyses are based on personal desires and you have *not* made a careful assessment of the implications of these for your methods and results, you are in danger of arriving at invalid conclusions.

However, your personal reasons for wanting to conduct a study, and the experiences and perspectives in which these are grounded, are not simply a source of "bias" (see the later discussion of this issue in the section on validity); they can also provide you with a valuable source of insight, theory, and data about the phenomena you are studying (Marshall & Rossman, 1995, pp. 17-22; Strauss & Corbin, 1990, pp. 42-43). This source is discussed in the next section, in the subsection on experiential knowledge.

One personal purpose in particular that deserves thought is your motivation for choosing a qualitative approach. Qualitative research is *not* easier than quantitative research, and seeking to avoid statistics bears little relationship to having the personal interests and skills required for the conduct of qualitative inquiry (Locke, Spirduso, & Silverman, 1993, pp. 107-110). Your reasons for adopting a qualitative approach need to be compatible with your other purposes, your research questions, and the requirements of carrying out qualitative research.

Besides your personal purposes, there are two other, more public kinds of purposes that I want to distinguish and discuss: practical purposes (including administrative or policy purposes) and research purposes. Practical purposes are focused on *accomplishing* something—meeting some need, changing some situation, or achieving some goal. Research purposes, on the other hand, are focused on *understanding* something, gaining some insight into what is going on and why it is happening. Although applied research design places much more emphasis on practical purposes than does basic research, you still need to address the issue of what you want to *understand* by doing the study, and how this understanding will contribute to your accomplishing your practical purposes. (The issue of what you want to understand is discussed in more detail below, in the section on research questions.)

There are five particular research purposes for which qualitative studies are especially useful:

1. Understanding the *meaning*, for participants in the study, of the events, situations, and actions they are involved with, and of the accounts that they give of their lives and experiences. In a qualitative study, you are interested not only in the physical events and behavior taking place, but also in how the participants in your study make sense of these, and how their understandings influence their behavior. The perspectives on events and actions held by the people involved in them are not simply their accounts of these events and actions, to be assessed in terms of truth or falsity; they are *part of* the reality that you are trying to understand (Maxwell, 1992; Menzel, 1978). This focus on meaning is central to what is known as the "interpretive" approach to social science (Bredo & Feinberg, 1982; Geertz, 1973; Rabinow & Sullivan, 1979).

2. Understanding the particular *context* within which the participants act, and the influence this context has on their actions. Qualitative researchers typically study a relatively small number of individuals or situations and preserve the individuality of each of these in their analyses, rather than collecting data from large samples and aggregating the data across individuals or situations. Thus they are able to understand how events, actions, and meanings are shaped by the unique circumstances in which these occur.

3. Identifying unanticipated phenomena and influences, and generating new, "grounded" theories about the latter. Qualitative research has long been used for this purpose by survey and experimental researchers, who often conduct "exploratory" qualitative studies to help them design their questionnaires and identify variables for experimental investigation. Although qualitative research is not restricted to this exploratory role, it is still an important strength of qualitative methods.

4. Understanding the *processes* by which events and actions take place. Although qualitative research is not unconcerned with outcomes, a major strength of qualitative studies is their ability to get at the processes that lead to these outcomes, processes that experimental and survey research are often poor at identifying (Britan, 1978; Patton, 1990, pp. 94ff.).

5. Developing causal explanations. The traditional view that qualitative research cannot identify causal relationships has long been disputed by some qualitative researchers (Britan, 1978; Denzin, 1978), and both qualitative and quantitative researchers are increasingly accepting the legitimacy of using qualitative methods for causal inference (e.g., Cook & Shadish, 1985; Erickson, 1986/1990, p. 82; Maxwell, 1996b; Miles & Huberman, 1994, pp. 144-148; Mohr, 1995, pp. 261-273, 1996; Rossi & Berk, 1991, p. 226; Sayer, 1992). Deriving causal explanations from a qualitative study is not an easy or straightforward task, but qualitative research is no different from quantitative research in this respect. Both approaches need to identify and deal with the plausible validity threats to any proposed causal explanation, as discussed below.

These research purposes, and the inductive, open-ended strategy that they require, give qualitative research an advantage in addressing numerous practical purposes, including the following.

Generating results and theories that are understandable and experientially credible, both to the people being studied and to others (Bolster, 1983). Although quantitative data may have greater credibility for some purposes and audiences, the specific detail and personal immediacy of qualitative data can lead to their greater influence in other situations. For example, I was involved in one evaluation, of how teaching rounds in one hospital department could be improved, that relied primarily on participant observation of rounds and open-ended interviews with staff physicians and residents (Maxwell, Cohen, & Reinhard, 1983). The evaluation led to decisive departmental action, in part because department members felt that the report, which contained detailed descriptions of activities during rounds and numerous quotes from interviews to support the analysis of the problems with rounds, "told it like it really was" rather than simply presented numbers and generalizations to back up its recommendations.

Conducting formative studies, ones that are intended to help improve existing practice rather than simply to determine the outcomes of the program or practice being studied (Scriven, 1967, 1991). In such studies, which are particularly useful for applied research, it is more important to understand the *process* by which things happen in a particular situation than to measure outcomes rigorously or to compare a given situation with others.

Engaging in collaborative, action, or "empowerment" research with practitioners or research participants (e.g., Cousins & Earl, 1995; Fetterman, Kaftarian, & Wandersman, 1996; Oja & Smulyan, 1989; Whyte, 1991). The focus of qualitative research on particular contexts and their meaning for the participants in these contexts, and on the processes occurring in these contexts, makes it especially suitable for collaborations with practitioners or with members of the community being studied (Patton, 1990, pp. 129-130; Reason, 1994).

A useful way of sorting out and formulating the purposes of your study is to write memos in which you reflect on your goals and motives, as well as the implications of these for your design decisions (for more information on such memos, see Maxwell, 1996a; Mills, 1959, pp. 197-198; Strauss & Corbin, 1990, chap. 12). I regularly use such memos as assignments in my methods courses; one doctoral student, Isabel Londoño, said that "writing memos for classes was key, having to put things to paper," in figuring out her purposes in choosing a dissertation topic (see Maxwell, 1996a, pp. 22-23).

Conceptual Context: What Do You Think Is Going On?

The conceptual context of your study is the system of concepts, assumptions, expectations, beliefs, and theories that supports and informs your research. This context, or a diagrammatic representation of it, is often called a "conceptual framework" (Hedrick, Bickman, & Rog, 1993, p. 19; Miles & Huberman, 1994; Robson, 1993). Miles and Huberman (1994) state that a conceptual framework "explains, either graphically or in narrative form, the main things to be studied—the key factors, concepts, or variables—and the presumed relationships among them" (p. 18).

Thus your conceptual context is a formulation of what you think is *going on* with the phenomena you are studying—a tentative *theory* of what is happening. Theory provides a model or map of *why* the world is the way it is (Strauss, 1995). It is a simplification of the world, but a simplification aimed at clarifying and explaining some aspect of how it works. It is not simply a "framework," although it can provide that, but a *story* about what you think is happening and why. A useful theory is one that tells an enlightening story about some phenomenon, one that gives you new insights and broadens your understanding of that phenomenon. The function of theory in your design is to inform the rest of the design—to help you to assess your purposes, develop and select realistic and relevant research questions and methods, and identify potential validity threats to your conclusions.

Some writers label this part of a research design or proposal as the "literature review." This can be a dangerously misleading term, for three reasons. First, it can lead you to focus narrowly on "literature," ignoring other conceptual resources that may be of equal or greater importance for your study, including your own experience. Second, it tends to generate a strategy of "covering the field" rather than focusing specifically on those studies and theories that are particularly *relevant* to your research. Third, it can make you think that your task is simply descriptive—to tell what previous researchers have found or what theories have been proposed. In developing a conceptual context, your purpose is not only descriptive, but also critical; you need to treat "the literature" not as an *authority* to be deferred to, but as a useful but fallible source of *ideas* about what's going on, and to attempt to see alternative ways of framing the issues.

Another way of putting this is that the conceptual context for your research study is something that is *constructed*, not found. It incorporates pieces that are borrowed from elsewhere, but the structure, the overall coherence, is something that *you* build, not something that exists ready-made. Becker (1986, pp. 141ff.) systematically develops the idea that prior work provides *modules* that you can use in building your conceptual context, modules that you need to examine critically to make sure they work effectively with the rest of your design. There are four main sources for these modules: your own experiential

knowledge, existing theory and research, pilot and exploratory studies, and thought experiments.

☐ *Experiential Knowledge*

Traditionally, what you bring to the research from your background and identity has been treated as "bias," something whose influence needs to be *eliminated* from the design, rather than a valuable component of it. However, the explicit incorporation of your identity and experience (what Strauss, 1987, calls "experiential data") in your research has recently gained much wider theoretical and philosophical support (e.g., Berg & Smith, 1988; Jansen & Peshkin, 1992). Using this experience in your research can provide you with a major source of insights, hypotheses, and validity checks. For example, Grady and Wallston (1988, p. 41) describe how one health care researcher used insights from her own experience to design a study of why many women don't do breast self-examination.

This is not a license to impose your assumptions and values uncritically on the research. Reason (1988) uses the term "critical subjectivity" to refer to "a quality of awareness in which we do not suppress our primary experience; nor do we allow ourselves to be swept away and overwhelmed by it; rather we raise it to consciousness and use it as part of the inquiry process" (p. 12). However, there are few well-developed and explicit strategies for doing this. One technique that I use in my qualitative methods course and in my own research is what I call a "researcher experience memo"; I have given suggestions for this elsewhere (Maxwell, 1996a). Basically, this involves reflecting on, and writing down, the different aspects of your experience and identity that are potentially relevant to your study. Doing this can generate unexpected insights and connections, as well as create a valuable record of these.

☐ *Existing Theory and Research*

The second major source of modules for your conceptual context is existing theory and research. This can be found not only in published work, but in unpublished papers and dissertations, in conference presentations, and in the heads of active researchers in your field (Locke et al., 1993, pp. 48-49).

Using existing theory in qualitative research has both advantages and dangers. A useful theory helps you to organize what you see. Particular pieces of data that otherwise might seem unconnected or irrelevant to one another or to your research questions can be related if you can fit them into the theory. A useful theory also illuminates what you are seeing in your research. It draws your attention to particular events or phenomena, and sheds light on relationships that might otherwise go unnoticed or misunderstood.

However, Becker (1986) warns that the existing literature, and the assumptions embedded in it, can deform the way you frame your research, causing you to overlook important ways of conceptualizing your study or key implications of your results. The literature has the advantage of what he calls "ideological hegemony," making it difficult for you to see any phenomenon in ways that are different from those that are prevalent in the literature. Trying to fit your insights into this established framework can deform your argument, weakening its logic and making it harder for you to see what this new way of framing the phenomenon might contribute. Becker describes how existing theory and perspectives deformed his early research on marijuana use, leading him to focus on the dominant question in the literature and to ignore the most interesting implications and possibilities of his study.

Becker (1986) argues that there is no way to be sure when the established approach is wrong or misleading or when your alternative is superior. All you can do is try to identify the ideological component of the established approach, and see what happens when you abandon these assumptions. He asserts that "a serious scholar ought routinely to inspect competing ways of talking about the same subject matter," and warns, "Use the literature, don't let it use you" (p. 149; see also Mills, 1959).

A review of relevant prior research can serve several other purposes in your design besides providing you with existing theory (see Strauss, 1987, pp. 48-56). First, you can use it to develop a justification for your study—to show how your work will address an important need or unanswered question (Marshall & Rossman, 1995, pp. 22-26). Second, it can inform your decisions about methods, suggesting alternative approaches or revealing potential problems with your plans. Third, it can be a source of *data* that you can use to test or modify your theories. You can see if existing theory, the results of your pilot research, or your experiential understanding is supported or challenged by previous studies. Finally, you can use prior research to help you *generate* theory. For example, I have used a wide range of empirical studies, as well as modules derived from existing theory, to develop a radically different theory of the relationships among diversity, social solidarity, and community from that prevalent in the literature (Maxwell, in press), and I am currently applying this theory in an attempt to explain the success of a systemic educational reform initiative in a multiracial and multiethnic urban school district.

☐ *Pilot and Exploratory Studies*

Pilot studies serve some of the same functions as prior research, but they can be focused more precisely on your own concerns and theories. You can design pilot studies specifically to test your ideas or methods and explore their implications, or to inductively develop *grounded* theory. One particular use that pilot studies have in qualitative research is to generate an understanding of the concepts and theories held by the people you are studying—what I have called "interpretation" (Maxwell, 1992). This is not simply a source of additional concepts for your theory; instead, it provides you with an understanding

of the *meaning* that these phenomena and events have for the actors who are involved in them, and the perspectives that inform their actions. In a qualitative study, these meanings and perspectives should constitute an important focus of your theory; as discussed earlier, they are one of the things your theory is *about*, not simply a source of theoretical insights and building blocks for the latter.

☐ *Thought Experiments*

Thought experiments have a long and respected tradition in the physical sciences (much of Einstein's work was based on thought experiments), but have received little attention in discussions of research design, particularly qualitative research design. Thought experiments draw on both theory and experience to answer "what if" questions, to seek out the logical implications of various properties of the phenomena you want to study. They can be used both to test your current theory for logical problems and to generate new theoretical insights. They encourage creativity and a sense of exploration, and can help you to make explicit the experiential knowledge that you already possess. Finally, they are easy to do, once you develop the skill. Valuable discussions of thought experiments in the social sciences are presented by Mills (1959) and Lave and March (1975).

Experience, prior theory and research, pilot studies, and thought experiments are the four major sources of the conceptual context for your study. The ways in which you can put together a useful and valid conceptual context from these sources are particular to each study, and not something for which any cookbook exists. The main thing to keep in mind is the need for integration of these components with one another, and with your purposes and research questions. A particularly valuable tool for generating and understanding these connections in your research is a technique known as concept mapping (Novak & Gowin, 1984); I have provided guidance for using concept maps in qualitative research design elsewhere (Maxwell, 1996a).

■ *Research Questions: What Do You Want to Understand?*

Your research questions—what you specifically want to understand by doing your study—are at the heart of your research design. They are the one component that directly connects to all of the other components of the design. More than any other aspect of your design, your research questions will have an influence on, and should be responsive to, every other part of your study.

This is different from seeing research questions as the starting point or primary determinant of the design. Models of design that place the formulation of research questions at the beginning of the design process, and that see these questions as determining the other aspects of the design, don't do justice to the interactive and inductive nature of qualitative research. The research questions in a qualitative study should not be formulated in detail until the purposes and context (and sometimes general aspects of the sampling and data collection) of the design are clarified, and should remain sensitive and adaptable to the implications of other parts of the design. Often you will need to do a significant part of the research before it is clear to you what specific research questions it makes sense to try to answer.

This does not mean that qualitative researchers should, or usually do, begin studies with *no* questions, simply going into the field with "open minds" and seeing what is there to be investigated. Every researcher begins with a substantial base of experience and theoretical knowledge, and these inevitably generate certain questions about the phenomena studied. These initial questions frame the study in important ways, influence decisions about methods, and are one basis for further focusing and development of more specific questions. However, these specific questions are generally the *result* of an interactive design process, rather than the starting point for that process. For example, Suman Bhattacharjea (1994; see Maxwell, 1996a, p. 50) spent a year doing field research on women's roles in a Pakistani educational district office before she was able to focus on two specific research questions and submit her dissertation proposal; at that point, she had also developed several hypotheses as tentative answers to these questions.

☐ *The Functions of Research Questions*

In your research design, the research questions serve two main functions: to help you to focus the study (the questions' relationship to your purposes and conceptual context) and to give you guidance for how to conduct it (their relationship to methods and validity). A design in which the research questions are too general or too diffuse creates difficulties both for conducting the study—in knowing what site or informants to choose, what data to collect, and how to analyze these data—and for clearly connecting what you learn to your purposes and existing knowledge (Miles & Huberman, 1994, pp. 22-25). Research questions that are precisely framed too early in the study, on the other hand, may lead you to overlook areas of theory or prior experience that are relevant to your understanding of what is going on, or cause you to pay too little attention to a wide range of data early in the study, data that can reveal important and unanticipated phenomena and relationships.

A third problem is that you may be smuggling unexamined assumptions into the research questions themselves, imposing a conceptual framework that doesn't fit the reality you are studying. A research question such as "How do teachers deal with the experience of isolation from their colleagues in their classrooms?" assumes that teachers *do* experience such isolation. Such an assumption needs to be carefully examined and justified, and without this justification it might be better to frame such a question as a tentative subquestion

to broader questions about the nature of classroom teachers' experience of their work and their relations with colleagues.

For all of these reasons, there is real danger to your study if you do not carefully formulate your research questions in connection with the other components of your design. Your research questions need to take account of what you want to accomplish by doing the study (your purposes), and of what is already known about the things you want to study and your tentative theories about these phenomena (your conceptual context). There is no reason to pose research questions for which the answers are already available, that don't clearly connect to what you think is actually going on, or that would have no direct relevance to your goals in doing the research.

Likewise, your research questions need to be ones that are answerable by the kind of study you can actually conduct. There is no value to posing questions that no feasible study could answer, either because the data that could answer them could not be obtained or because any conclusions you might draw from these data would be subject to serious validity threats.

A common problem in the development of research questions is confusion between research issues (what you want to *understand* by doing the study) and practical issues (what you want to *accomplish*). Your research questions need to *connect* clearly to your practical concerns, but in general an empirical study cannot directly answer practical questions such as, "How can I improve this program?" or "What is the best way to increase medical students' knowledge of science?" In order to address such practical questions, you need to focus on what you don't *understand* about the phenomena you are studying, and to investigate what is really going on with these phenomena. For example, the practical goal of Martha Regan-Smith's (1992) dissertation research was to improve the teaching of the basic sciences in medical school (see Maxwell, 1996a, pp. 116ff.). However, her research questions focused not on this goal, but on what exceptional teachers in her school did that helped students to learn science—something she had realized that she didn't know, and that ought to have important implications for how to improve such teaching overall. Unless you frame research questions that your study can clearly address, you run the risk of either designing a study with unanswerable questions or smuggling your goals into the answers to the questions themselves, destroying the credibility of your study.

A second confusion, one that can create problems for interview studies, is that between research questions and interview questions. Your research questions identify the things that you want to understand; your interview questions generate the data that you need to understand these things. This distinction is discussed in more detail below, in the section on methods.

There are three issues that you should keep in mind in formulating research questions for applied social research. First, research questions may legitimately be framed in particular as well as general terms. There is a strong tendency in basic research to state research questions in general terms, such as, "How do students deal with racial and ethnic difference in multiracial schools?" and then to "operationalize" these questions by selecting a particular sample or site. This tendency can be counterproductive when the purpose of your study is to understand and improve some particular program, situation, or practice. In applied research, it is often more appropriate to formulate research questions in particular terms, such as, "How do students at North High School deal with racial and ethnic difference?"

Second, some researchers believe that questions should be stated in terms of what the respondents report or what can be directly observed, rather than in terms of inferred behavior, beliefs, or causal influences. This is what I call an instrumentalist or positivist, rather than a realist, approach to research questions (Maxwell, 1992; Norris, 1983). Instrumentalists formulate their questions in terms of observable or measurable data, and are suspicious of inferences to things that cannot be defined in terms of such data. For example, instrumentalists would reject a question such as, "How do exemplary teachers help medical students learn science?" and replace it with questions like "How do medical students *report that* exemplary teachers help them learn science?" or "How are exemplary teachers *observed to teach* basic science?"

Realists, in contrast, don't assume that research questions about feelings, beliefs, intentions, prior behavior, effects, and so on need to be reduced to, or reframed as, questions about the actual data that one uses. Instead, they treat their data as fallible *evidence* about these phenomena, to be used critically to develop and test ideas about what is going on (Campbell, 1988; Cook & Campbell, 1979; Maxwell, 1992).

The main risk of using instrumentalist questions is that you will lose sight of what you are really interested in, and define your study in ways that obscure the actual phenomena you want to investigate, ending up with a rigorous but uninteresting conclusion. As in the joke about the man who was looking for his keys under the streetlight (rather than where he dropped them) because the light was better there, you may never find what you started out to look for. An instrumentalist approach to your research questions may also make it more difficult for your study to address important purposes of your study directly, and can inhibit your theorizing about phenomena that are not directly observable.

My own preference is to use realist questions, and to address as systematically and rigorously as possible the validity threats that this approach involves. The seriousness of these validity threats (such as self-report bias) needs to be assessed in the context of a particular study; these threats are often not as serious as instrumentalists imply. There are also effective ways to address these threats in a qualitative design, which I discuss below in the section on validity. The risk of trivializing your study by restricting your questions to what can be directly observed is usually more serious than the risk of drawing invalid conclusions. As the statistician John Tukey (1962) put it, "Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise" (p. 13).

One issue that is not entirely a matter of realism versus instrumentalism is whether research questions in interview studies should be framed in terms of

the respondents' perceptions or beliefs rather than the actual state of affairs. You should base this decision not simply on the seriousness of the validity threats, but also on what you actually want to understand. In many qualitative studies, the real interest is in how participants make sense of what has happened, and how this perspective informs their actions, rather than determining precisely what took place.

Finally, many researchers (consciously or unconsciously) focus their questions on variance rather than process (Maxwell, 1996a; Mohr, 1982). Variance questions deal with difference and correlation; they often begin with "Is there," "Does," "How much," or "To what extent." For example, a variance approach to Martha Regan-Smith's (1992) study would ask questions like "Do exemplary medical school teachers differ from others in their teaching of basic science?" or "Is there a relationship between teachers' behavior and students' learning?" and attempt to measure these differences and relationships. Process questions, in contrast, focus on *how* and *why* things happen, rather than *whether* there is a particular difference or relationship or how much it is explained by other variables. Regan-Smith's actual questions focused on *how* these teachers helped students learn, that is, the process by which their teaching helped the students to learn.

In a qualitative study, it can be dangerous for you to frame your research questions in a way that focuses on differences and their explanation. This may lead you to begin thinking in variance terms, to try to identify the variables that will account for observed or hypothesized differences, and to overlook the real strength of a qualitative approach, which is in understanding the process by which phenomena take place. Variance questions are often best answered by quantitative approaches, which are powerful ways of determining *whether* a particular result is causally related to one or another variable, and to *what extent* these are related. However, qualitative research is often better at showing *how* this occurred. Variance questions are legitimate in qualitative research, but they are often best grounded in the answers to prior process questions.

Qualitative researchers thus tend to generate two kinds of questions that are much better suited to process theory than to variance theory: questions about the *meaning* of events and activities to the people involved in them and questions about the influence of the physical and social *context* on these events and activities. (See the earlier discussion of meaning and context as research purposes.) Because both of these types of questions involve situation-specific phenomena, they do not lend themselves to the kinds of comparison and control that variance theory requires. Instead, they generally involve an open-ended, inductive approach, in order to discover what these meanings and influences are, and *how* they are involved in these events and activities—an inherently processual orientation.

Developing relevant, focused, answerable research questions takes time; such questions cannot be thrown together quickly, nor in most studies can they be definitively formulated before data collection and analysis begin. Generating good questions requires that you pay attention not just to the questions themselves, but to their connections with all of the other design components: the purposes that answering the questions might serve, the implications for your questions of your conceptual context, the methods you could use to answer the questions, and the validity threats you will need to address. As is true with the other components of your design, writing memos about these issues is an extremely useful tool for developing your questions (Maxwell, 1996a, pp. 61-62).

Methods: What Will You Actually Do?

There is no "cookbook" for doing qualitative research. The appropriate answer to almost any question about the use of qualitative methods is "It depends." The value and feasibility of your research methods cannot be guaranteed by your adhering to methodological rules; rather, they depend on the specific setting and phenomena you are studying and the actual consequences of your strategy for studying it.

Prestructuring a Qualitative Study

One of the most important issues in designing a qualitative study is how much you should attempt to prestructure your methods. Structured approaches can help to ensure the comparability of data across sources and researchers, and are thus particularly useful in answering variance questions, questions that deal with *differences* between things and the explanation for these differences. Unstructured approaches, in contrast, allow the researcher to focus on the *particular* phenomena studied; they trade generalizability and comparability for internal validity and contextual understanding, and are particularly useful for understanding the processes that led to specific outcomes, what Huberman and Miles (1988) call "local causality." Sayer (1992, pp. 241ff.) refers to these two approaches as "extensive" and "intensive" research designs, respectively.

However, Miles and Huberman (1994) warn that

> highly inductive, loosely designed studies make good sense when experienced researchers have plenty of time and are exploring exotic cultures, understudied phenomena, or very complex social phenomena. But if you're new to qualitative studies and are looking at a better understood phenomenon within a familiar culture or subculture, a loose, inductive design is a waste of time. Months of fieldwork and voluminous case studies may yield only a few banalities. (p. 17)

They also point out that prestructuring reduces the amount of data that you have to deal with, functioning as a form of preanalysis that simplifies the analytic work required.

Unfortunately, most discussions of this issue treat prestructuring as a single dimension, and view it in terms of metaphors such as hard versus soft and tight

versus loose. Such metaphors have powerful connotations (although they are different for different people) that can lead you to overlook or ignore the numerous ways in which studies can vary, not just in the *amount* of prestructuring, but in *how* prestructuring is used. For example, you could employ an extremely open approach to data collection, but use these data for a confirmatory test of explicit hypotheses based on a prior theory (e.g., Festinger, Riecker, & Schachter, 1956). In contrast, the approach often known as ethnoscience or cognitive anthropology (Werner & Schoepfle, 1987a, 1987b) employs highly structured data collection techniques, but interprets these data in a largely inductive manner, with very few preestablished categories. Thus the decision you face is not primarily *whether* or *to what extent* you prestructure your study, but *in what ways* you do this, and *why*.

Finally, it is worth keeping in mind that you can lay out a *tentative* plan for some aspects of your study in considerable detail, but leave open the possibility of substantially revising this if necessary. Emergent insights may require new sampling plans, different kinds of data, and different analytic strategies.

I distinguish four main components of qualitative methods:

1. The research relationship that you establish with those you study
2. Sampling: what times, settings, or individuals you select to observe or interview, and what other sources of information you decide to use
3. Data collection: how you gather the information you will use
4. Data analysis: what you do with this information in order to make sense of it

It is useful to think of all of these components as involving *design* decisions—key issues that you should consider in planning your study, and that you should rethink as you are engaged in it.

☐ *Negotiating a Research Relationship*

Your relationships with the people in your study can be complex and changeable, and these relationships will necessarily affect you as the "research instrument," as well as have implications for other components of your research design. My changing relationships with the people in the Inuit community in which I conducted my dissertation research (Maxwell, 1986) had a profound effect not only on my own state of mind, but on who I was able to interview, my opportunities for observation of social life, the quality of the data I collected, the research questions I was able to answer, and my ability to test my conclusions. The term *reflexivity* (Hammersley & Atkinson, 1983) is often used for this unavoidable mutual influence of the research participants and the researcher on each other.

There are also philosophical, ethical, and political issues that should inform the kind of relationship that you want to establish. In recent years, there has been a growing interest in alternatives to the traditional style of research, including participatory action research, collaborative research, feminist research, critical ethnography, and empowerment research (see Denzin & Lincoln, 1994; Fetterman et al., 1996; Oja & Smulyan, 1989; Whyte, 1991). Each of these modes of research involves different sorts of relationships between the researcher and the participants in the research, and has different implications for the rest of the research design.

Thus it is important that you think about the kinds of relationships you want to have with the people whom you study, and what you need to do to establish such relationships. I see these as *design decisions*, not simply as external factors that may affect your design. Although they are not completely under your control and cannot be defined precisely in advance, they are still matters that require systematic planning and reflection if your design is to be as coherent as possible.

☐ *Decisions About Sampling: Where, When, Who, and What*

Whenever you have a choice about when and where to observe, whom to talk to, or what information sources to focus on, you are faced with a sampling decision. Even a single case study involves a choice of this case rather than others, as well as requiring sampling decisions *within* the case itself. Miles and Huberman (1994, pp. 27-34) and LeCompte and Preissle (1993, pp. 56-85) provide valuable discussions of particular sampling issues; here, I want to talk more generally about the nature and purposes of sampling in qualitative research.

Works on quantitative research generally treat anything other than probability sampling as "convenience sampling," and strongly discourage the latter. For qualitative research, this ignores the fact that most sampling in qualitative research is neither probability sampling nor convenience sampling, but falls into a third category: purposeful sampling (Patton, 1990, pp. 169ff.). This is a strategy in which particular settings, persons, or events are deliberately selected for the important information they can provide that cannot be gotten as well from other choices.

There are several important uses for purposeful sampling. First, it can be used to achieve representativeness or typicality of the settings, individuals, or activities selected. A small sample that has been systematically selected for typicality and relative homogeneity provides far more confidence that the conclusions adequately represent the average members of the population than does a sample of the same size that incorporates substantial random or accidental variation. Second, purposeful sampling can be used to capture adequately the heterogeneity in the population. The goal here is to ensure that the conclusions adequately represent the entire *range* of variation, rather than only the typical members or some subset of this range. Third, a sample can be purposefully selected to allow for the examination of cases that are critical for the theories the study began with, or that have subsequently been developed. Finally, pur-

poseful sampling can be used to establish particular comparisons to illuminate the reasons for differences between settings or individuals, a common strategy in multicase qualitative studies.

You should not make sampling decisions in isolation from the rest of your design. They should take into account your research relationship with study participants, the feasibility of data collection and analysis, and validity concerns, as well as your purposes and conceptual context. In addition, feasible sampling decisions often require considerable knowledge of the setting studied, and you will need to alter them as you learn more about what decisions will work best to give you the data you need.

☐ *Decisions About Data Collection*

Most qualitative methods texts devote considerable space to the strengths and limitations of particular data collection methods (see particularly Bogdan & Biklen, 1992; Patton, 1990; Weiss, 1994), so I won't deal with these issues here. Instead, I want to address two key design issues in selecting and using data collection methods: the relationship between research questions and data collection methods, and the triangulation of different methods.

Although researchers often talk about "operationalizing" their research questions, or of "translating" the research questions into interview questions, this language is a vestigial remnant of logical positivism that bears little relationship to qualitative research practice. There is no way to convert research questions into useful methods decisions; your methods are the *means* to answering your research questions, not a logical transformation of the latter. Their selection depends not only on your research questions, but on the actual research situation and what will work most effectively in that situation to give you the data you need. For example, your interview questions should be judged not by whether they can be logically derived from your research questions, but by whether they provide the *data* that will contribute to answering these questions, an issue that may require pilot testing a variety of questions or actually conducting a significant part of the interviews. You need to anticipate, as best you can, how particular interview questions or other data collection strategies will actually work in practice. In addition, your interview questions and observational strategies will generally be far more focused, context-specific, and diverse than the broad, general research questions that define what you seek to understand in conducting the study. The development of a good data collection plan requires creativity and insight, not a mechanical translation of your research questions into methods.

In addition, qualitative studies generally rely on the integration of data from a variety of methods and sources of information, a general principle known as triangulation (Denzin, 1978). This reduces the risk that your conclusions will reflect only the systematic biases or limitations of a specific method, and allows you to gain a better assessment of the validity and generality of the explanations that you develop. Triangulation is also discussed below, in the section on validity.

☐ *Decisions About Data Analysis*

Analysis is often conceptually separated from design, especially by writers who see design as what happens *before* the data are actually collected. Here, I treat analysis as a part of design, and as something that must itself be designed. Every qualitative study requires decisions about how the analysis will be done, and these decisions should influence, and be influenced by, the rest of the design.

One of the most common problems qualitative researchers have is that they let their unanalyzed field notes and transcripts pile up, making the task of final analysis much more difficult and discouraging than it needs to be. In my dissertation research on Inuit kinship, if I had not analyzed my data as I collected it, I would have missed the insights that enabled me to collect many of the data I eventually used to support my conclusions. You should begin data analysis immediately after finishing the first interview or observation, and continue to analyze the data as long as you are working on the research. This allows you to progressively focus your interviews and observations, and to decide how to test your emerging conclusions.

Strategies for qualitative analysis fall into three main groups: categorizing strategies (such as coding and thematic analysis), contextualizing strategies (such as narrative analysis and individual case studies), and memos and displays. These strategies are discussed in more detail by Coffey and Atkinson (1996) and Dey (1993). These methods can, and generally should, be combined, but I will begin by discussing them separately.

The main categorizing strategy in qualitative research is coding. This is rather different from coding in quantitative research, which consists of applying a preestablished set of categories to the data according to explicit, unambiguous rules, with the primary goal being to generate frequency counts of the items in each category. In qualitative research, in contrast, the goal of coding is not to produce counts of things, but to "fracture" (Strauss, 1987, p. 29) the data and rearrange it into categories that facilitate comparison between things in the same category and between categories. These categories may be derived from existing theory, inductively generated during the research (the basis for what Glaser & Strauss, 1967, term "grounded theory"), or drawn from the categories of the people studied (what anthropologists call "emic" categories). Such categorizing makes it much easier for you to develop a general understanding of what is going on, to generate themes and theoretical concepts, and to organize and retrieve your data to test and support these general ideas. (An excellent practical source on coding is Bogdan & Biklen, 1992; for more elaborate treatment, see Dey, 1993.)

However, fracturing and categorizing your data can lead to the neglect of contextual relationships among these data, relationships based on contiguity

rather than similarity (Maxwell & Miller, n.d.), and can create analytic blinders, preventing you from seeing alternative ways of understanding your data. Atkinson (1992) describes how his initial categorizing analysis of data on the teaching of general medicine affected his subsequent analysis of the teaching of surgery: "On rereading the surgery notes, I initially found it difficult to *escape* those categories I had initially established [for medicine]. Understandably, they furnished a powerful conceptual grid. . . . The notes as I confronted them had been fragmented into the constituent themes" (pp. 458-459).

What I call contextualizing strategies (Maxwell & Miller, n.d.) were developed in part to deal with these problems. Instead of fracturing the initial text into discrete elements and re-sorting it into categories, contextualizing analysis attempts to understand the data (usually, but not necessarily, an interview transcript or other textual material) in context, using various methods to identify the relationships among the different elements of the text. Such strategies include some forms of case studies (Patton, 1990), profiles (Seidman, 1991), some types of narrative analysis (Coffey & Atkinson, 1996), and ethnographic microanalysis (Erickson, 1992). What all of these strategies have in common is that they look for relationships that connect statements and events within a particular context into a coherent whole. Atkinson (1992) states:

> I am now much less inclined to fragment the notes into relatively small segments. Instead, I am just as interested in reading episodes and passages at greater length, with a correspondingly different attitude toward the act of reading and hence of analysis: Rather than constructing my account like a patchwork quilt, I feel more like working with the whole cloth. . . . To be more precise, what now concerns me is the nature of these products as *texts*. (p. 460)

The distinction between categorizing and contextualizing strategies has important implications for your research questions. A research question that asks about the way events in a specific context are connected cannot be answered by an exclusively categorizing analysis (Agar, 1991). Conversely, a question about similarities and differences across settings or individuals, or about general themes in your data, cannot be answered by an exclusively contextualizing analysis. Your analysis strategies have to be compatible with the questions you are asking. Both categorizing and contextualizing strategies are legitimate and valuable tools in qualitative analysis, and a study that relies on only one of these runs the risk of missing important insights.

The third category of analytic tools, memos and displays, is also a key part of qualitative analysis (Miles & Huberman, 1994, pp. 72-75; Strauss & Corbin, 1990, pp. 197-223). As discussed above, memos can perform functions not related to data analysis, such as reflection on methods, theory, or purposes. However, displays and memos are valuable *analytic* techniques for the same reasons they are useful for other purposes: They facilitate your thinking about relationships in your data and make your ideas and analyses visible and retrievable. You should write memos frequently while you are doing data analysis, in order to stimulate and capture your ideas about your data. Displays (Miles & Huberman, 1994), which include matrices or tables, networks or concept maps, and various other forms, also serve two other purposes: data reduction and the presentation of data or analysis in a form that allows you to see it as a whole.

There are now a substantial number of computer programs available for analyzing qualitative data and a number of recent books comparing and evaluating these (e.g., Tesch, 1990; Weitzman & Miles, 1995). Although none of these programs eliminates the need to read your data and create your own concepts and relationships, they can enormously simplify the task of coding and retrieving data in a large project. However, most of these programs are designed primarily for categorizing analysis, and may distort your analytic strategy toward such approaches. For example, one group of researchers, employing a widely used qualitative analysis program to analyze interviews with historians about how they worked, produced a report that identified common themes and provided examples of how individual historians talked about these, but completely failed to answer the funder's key questions, which had to do with how individual historians thought about the connections among these different issues in their own work (Agar, 1991). So-called hypertext programs (Coffey & Atkinson, 1996, pp. 181-186) allow you to create electronic links, representing any sort of connection you want, among data within a particular context, but the openness of such programs can make them difficult for less experienced researchers to use effectively. A few of the more structured programs, such as ATLAS/ti, enable you not only to create links among data chunks, codes, and memos, but also to display the resulting networks (Weitzman & Miles, 1995, pp. 222-224).

Validity: How Might You Be Wrong?

Quantitative and experimental researchers generally attempt to design, in advance, controls that will deal with both anticipated and unanticipated threats to validity. Qualitative researchers, on the other hand, rarely have the benefit of formal comparisons, sampling strategies, or statistical manipulations that "control for" the effect of particular variables, and must try to rule out most validity threats after the research has begun, using evidence collected during the research itself to make these "alternative hypotheses" implausible. This approach requires you to identify the *specific* threat in question and to develop ways to attempt to rule out that particular threat. It is clearly impossible to list here all, or even the most important, validity threats to the conclusions of a qualitative study, but I want to discuss two broad types of threats to validity that are often raised in relation to qualitative studies: researcher bias and the effect of the researcher on the setting or individuals studied, generally known as reactivity.

Bias refers to ways in which data collection or analysis are distorted by the researcher's theory, values, or preconceptions. It is clearly impossible to deal with these problems by eliminating these theories, preconceptions, or values, as discussed earlier. Nor is it usually appropriate to try to "standardize" the researcher to achieve reliability; in qualitative research, the main concern is not with eliminating *variance* between researchers in the values and expectations they bring to the study, but with understanding how a *particular* researcher's values influence the conduct and conclusions of the study. As one qualitative researcher, Fred Hess, has phrased it, validity in qualitative research is the result not of indifference, but of integrity (personal communication). Strategies that are useful in achieving this are discussed below (and in more detail in Maxwell, 1996a).

Reactivity is a second problem that is often raised about qualitative studies. The approach to reactivity of most quantitative research, of trying to "control for" the effect of the researcher, is appropriate to a "variance theory" perspective, in which the goal is to prevent researcher *variability* from being an unwanted cause of variability in the outcome variables. However, eliminating the *actual* influence of the researcher is impossible (Hammersley & Atkinson, 1983), and the goal in a qualitative study is not to eliminate this influence but to understand it and to use it productively.

For participant observation studies, reactivity is generally *not* as serious a validity threat as many people believe. Becker (1970, pp. 45ff.) points out that in natural settings, an observer is generally much less of an influence on participants' behavior than is the setting itself (though there are clearly exceptions to this, such as settings in which illegal behavior occurs). For all types of interviews, in contrast, the interviewer has a powerful and inescapable influence on the data collected; what the interviewee says is *always* a function of the interviewer and the interview situation (Briggs, 1986; Mishler, 1986). Although there are some things you can do to prevent the more undesirable consequences of this (such as avoiding leading questions), trying to "minimize" your effect on the interviewee is an impossible goal. As discussed above for "bias," what is important is to understand *how* you are influencing what the interviewee says, and how this affects the validity of the inferences you can draw from the interview.

☐ *Validity Tests: A Checklist*

I discuss below some of the most important strategies you can use in a qualitative study to deal with particular validity threats and thereby increase the credibility of your conclusions. Miles and Huberman (1994, pp. 262ff.) include a more extensive list, having some overlap with mine, and other lists are given by Becker (1970), Kidder (1981), Guba and Lincoln (1989), and Patton (1990). Most of these strategies operate primarily not by *verifying* your conclusions, but by *testing* the validity of your conclusions and the existence of potential threats to those conclusions (Campbell, 1988). The idea is to look for evidence that challenges your conclusion, or that makes the potential threat implausible.

The modus operandi approach. One strategy often used for testing explanations in qualitative research, which differs significantly from those prevalent in quantitative research, has been called the "modus operandi method" by Scriven (1974). It resembles the approach of a detective trying to solve a crime, an FAA inspector trying to determine the cause of an airplane crash, a physician attempting to diagnose a patient's illness, or a historian, geologist, or evolutionary biologist trying to account for a particular sequence of events. However, its logic has received little formal explication (recent exceptions are found in Gould, 1989; Maxwell, 1996b; Mohr, 1995; Ragin, 1987), and has not been clearly understood even by many qualitative researchers. Basically, rather than trying to deal with alternative possible causes or validity threats as *variables*, by either holding them constant or comparing the result of differences in their values in order to determine their effect, the modus operandi method deals with them as *events*, by searching for clues as to whether they took place and were involved in the outcome in question. Thus a researcher who is concerned about whether some of her interviews with teachers were being influenced by their principal's well-known views on the topics being investigated, rather than eliminating teachers with this principal from her sample or comparing interviews of teachers with different principals to detect this influence, would look for internal evidence of this influence in her interviews or other data, or would try to find ways of investigating this influence directly through her interviews.

Searching for discrepant evidence and negative cases. Looking for and analyzing discrepant data and negative cases is an important way of testing a proposed conclusion. There is a strong and often unconscious tendency for researchers to notice supporting instances and ignore ones that don't fit their preestablished conclusions (Miles & Huberman, 1994, p. 263; Shweder, 1980). Thus you need to develop explicit and systematic strategies for making sure that you don't overlook data that could point out flaws in your reasoning or conclusions. However, discrepant evidence can itself be flawed; you need to examine both the supporting and discrepant evidence to determine whether the conclusion in question is more plausible than the potential alternatives.

Triangulation. Triangulation, as discussed above, reduces the risk of systematic distortions inherent in the use of only one method, because no single method is completely free from all possible validity threats. The most extensive discussion of triangulation as a validity-testing strategy in qualitative research is offered by Fielding and Fielding (1986), who emphasize the fallibility of *any* particular method and the need to design triangulation strategies to deal with specific validity threats. For example, interviews, questionnaires, and docu-

ments may all be vulnerable to self-report bias or ideological distortion; effective triangulation would require an additional method that is *not* subject to this particular threat, though it might well have other threats that would be dealt with by the former methods.

Feedback. Soliciting feedback from others is an extremely useful strategy for identifying validity threats, your own biases and assumptions, and flaws in your logic or methods. You should try to get such feedback from a variety of people, both those familiar with the phenomena or settings you are studying and those who are strangers to them. These two groups of individuals will give you different sorts of comments, but both are valuable.

Member checks. One particular sort of feedback deserves special attention: the systematic solicitation of the views of participants in your study about your data and conclusions, a process known as "member checks" (Guba & Lincoln, 1989, pp. 238-241; Miles & Huberman, 1994, pp. 275-277). This is the single most important way of ruling out the possibility of your misinterpreting the meaning of what the participants say and the perspective they have on what is going on. However, it is important that you not assume that participants' reactions are themselves necessarily valid (Bloor, 1983); their responses should be taken simply as *evidence* regarding the validity of your account (see Hammersley & Atkinson, 1983).

Rich data. "Rich" data are data that are detailed and complete enough that they provide a full and revealing picture of what is going on. In interview studies, such data generally require verbatim transcripts of the interviews, rather than simply notes on what you noticed or felt was significant. For observation, rich data are the product of detailed, descriptive note taking about the specific, concrete events that you observe. Becker (1970, pp. 51ff.) argues that such data "counter the twin dangers of respondent duplicity and observer bias by making it difficult for respondents to produce data that uniformly support a mistaken conclusion, just as they make it difficult for the observer to restrict his observations so that he sees only what supports his prejudices and expectations" (p. 52). The key function of rich data is to provide a *test* of your developing theories, rather than simply a source of supporting instances.

Quasi-statistics. Many of the conclusions of qualitative studies have an implicit quantitative component. Any claim that a particular phenomenon is typical, rare, or prevalent in the setting or population studied is an inherently quantitative claim, and requires some quantitative support. Becker (1970, p. 31) has coined the term "quasi-statistics" to refer to the use of simple numerical results that can be readily derived from the data. Quasi-statistics not only allow you to test and support claims that are inherently quantitative, they also enable you to assess the *amount* of evidence in your data that bears on a particular conclusion or threat, such as how many discrepant instances exist and from how many different sources they were obtained. For example, Becker et al. (1961), in their study of medical students, present more than 50 tables and graphs of the amount and distribution of their observation and interview data to support their conclusions.

Comparison. Although explicit comparisons (such as control groups) for the purpose of assessing validity threats are mainly associated with quantitative, variance-theory research, there are valid uses for comparison in qualitative studies, particularly multisite studies (e.g., Miles & Huberman, 1994, p. 237). In addition, single case studies often incorporate implicit comparisons that contribute to the interpretability of the case. For example, Martha Regan-Smith (1992), in her "uncontrolled" study of how exemplary medical school teachers helped students learn, used both the existing literature on "typical" medical school teaching and her own extensive knowledge of this topic to determine what was distinctive about the teachers she studied. Furthermore, the students she interviewed explicitly contrasted these teachers with others whom they felt were not as helpful to them, explaining not only what the exemplary teachers did that increased their learning, but *why* this was helpful.

☐ *Generalization in Qualitative Research*

Qualitative researchers often study only a single setting or a small number of individuals or sites, using theoretical or purposeful rather than probability sampling, and rarely make explicit claims about the generalizability of their accounts. Indeed, the value of a qualitative study may depend on its *lack* of generalizability in the sense of being representative of a larger population; it may provide an account of a setting or population that is illuminating as an extreme case or "ideal type." Freidson (1975), for his study of social controls on work in a medical group practice, deliberately selected an atypical practice, one in which the physicians were better trained and more "progressive" than usual and that was structured precisely to deal with the problems he was studying. He argues that the documented failure of social controls in this case provides a far stronger argument for the generalizability of his conclusions than would the study of a "typical" practice.

The generalizability of qualitative studies is usually based not on explicit sampling of some defined population to which the results can be extended, but on the development of a theory that can be extended to other cases (Becker, 1991; Ragin, 1987; Yin, 1994). For this reason, Guba and Lincoln (1989) prefer to talk of "transferability" rather than "generalizability" in qualitative research. Hammersley (1992, pp. 189-191) and Weiss (1994, pp. 26-29) list a number of features that lend credibility to generalizations made from case studies or nonrandom samples, including respondents' own assessments of generalizability, the similarity of dynamics and constraints to other situations, the presumed

ersality of the phenomenon studied, and corroboration from other
ever, none of these permits the kind of precise extrapolation of
ined populations that probability sampling allows.

ion

olcott (1990) provides a useful metaphor for research design:
best advice I've ever seen for writers happened to be included
tions I found for assembling a new wheelbarrow: *Make sure all*
erly in place before tightening" (p. 47). Like a wheelbarrow, your
gn not only needs to have all the required parts, it has to *work*—to
othly and accomplish its tasks. This requires attention to the con-
ng the different parts of the design—what I call *coherence*. There
ht Way to create a coherent qualitative design; in this chapter I
give you the tools that will enable you to put together *a* way that
and your research.

es

). The right brain strikes back. In N. G. Fielding & R. M. Lee (Eds.), *Using
n qualitative research* (pp. 181-194). Newbury Park, CA: Sage.
992). The ethnography of a medical setting: Reading, writing, and rhetoric.
Health Research, 2, 451-474.
1970). *Sociological work: Method and substance.* New Brunswick, NJ:
).
986). *Writing for social scientists: How to start and finish your thesis, book,*
Chicago: University of Chicago Press.
991). Generalizing from case studies. In E. W. Eisner & A. Peshkin (Eds.),
inquiry in education: The continuing debate (pp. 233-242). New York:
ollege Press.
Geer, B., Hughes, E. C., & Strauss, A. L. (1961). *Boys in white: Student
edical school.* Chicago: University of Chicago Press.
mith, K. K. (Eds.). (1988). *The self in social inquiry: Researching methods.*
ark, CA: Sage.
. (1994). *Reconciling "public" and "private": Women in the educational
y in "Sinjabistan" Province, Pakistan.* Unpublished doctoral dissertation,
aduate School of Education.
83). Notes on member validation. In R. M. Emerson (Ed.), *Contemporary
ch: A collection of readings* (pp. 156-172). Prospect Heights, IL: Wave-

Biklen, S. K. (1992). *Qualitative research for education: An introduction
d methods* (2nd ed.). Boston: Allyn & Bacon.
983). Toward a more effective model of research on teaching. *Harvard
Review, 53,* 294-308.

Bredo, E., & Feinberg, W. (1982). *Knowledge and values in social and educational research.* Philadelphia: Temple University Press.

Briggs, C. L. (1986). *Learning how to ask: A sociolinguistic appraisal of the role of the interview in social science research.* Cambridge: Cambridge University Press.

Britan, G. M. (1978). Experimental and contextual models of program evaluation. *Evaluation and Program Planning, 1,* 229-234.

Campbell, D. T. (1988). *Methodology and epistemology for social science: Selected papers.* Chicago: University of Chicago Press.

Coffey, A., & Atkinson, P. (1996). *Making sense of qualitative data: Complementary research strategies.* Thousand Oaks, CA: Sage.

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

Cook, T. D., & Shadish, W. R. (1985). Program evaluation: The worldly science. *Annual Review of Psychology, 37,* 193-232.

Cousins, J. B., & Earl, L. M. (Eds.). (1995). *Participatory evaluation in education: Studies in evaluation use and organizational learning.* London: Falmer.

Denzin, N. K. (1978). *The research act* (2nd ed.). New York: McGraw-Hill.

Denzin, N. K., & Lincoln, Y. S. (1994). *Handbook of qualitative research.* Thousand Oaks, CA: Sage.

Dey, I. (1993). *Qualitative data analysis: A user-friendly guide for social scientists.* London: Routledge.

Eisner, E. W., & Peshkin, A. (Eds.). (1990). *Qualitative inquiry in education: The continuing debate.* New York: Teachers College Press.

Erickson, F. (1990). Qualitative methods. In *Research in teaching and learning* (Vol. 2, pp. 75-194). New York: Macmillan. (Reprinted from *Handbook of research on teaching,* 3rd ed., pp. 119-161, by M. C. Wittrock, Ed., 1986, New York: Macmillan)

Erickson, F. (1992). Ethnographic microanalysis of interaction. In M. D. LeCompte, W. L. Millroy, & J. Preissle (Eds.), *The handbook of qualitative research in education* (pp. 201-225). San Diego, CA: Academic Press.

Festinger, L., Riecker, H. W., & Schachter, S. (1956). *When prophecy fails.* Minneapolis: University of Minnesota Press.

Fetterman, D. M., Kaftarian, S. J., & Wandersman, A. (Eds.). (1996). *Empowerment evaluation: Knowledge and tools for self-assessment and accountability.* Thousand Oaks, CA: Sage.

Fielding, N. G., & Fielding, J. L. (1986). *Linking data.* Beverly Hills, CA: Sage.

Freidson, E. (1975). *Doctoring together: A study of professional social control.* Chicago: University of Chicago Press.

Geertz, C. (1973). *The interpretation of cultures: Selected essays.* New York: Basic Books.

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

Glesne, C., & Peshkin, A. (1992). *Becoming qualitative researchers: An introduction.* White Plains, NY: Longman.

Gould, S. J. (1989). *Wonderful life: The Burgess shale and the nature of history.* New York: W. W. Norton.

Grady, K. E., & Wallston, B. S. (1988). *Research in health care settings.* Newbury Park, CA: Sage.

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation.* Newbury Park, CA: Sage.

Hammersley, M. (1992). *What's wrong with ethnography? Methodological explorations.* London: Routledge.

Hammersley, M., & Atkinson, P. (1983). *Ethnography: Principles in practice.* London: Tavistock.

Hedrick, T. E., Bickman, L., & Rog, D. J. (1993). *Applied research design: A practical guide.* Newbury Park, CA: Sage.

Huberman, A. M., & Miles, M. B. (1988). Assessing local causality in qualitative research. In D. N. Berg & K. K. Smith (Eds.), *The self in social inquiry: Researching methods* (pp. 351-381). Newbury Park, CA: Sage.

Jansen, G., & Peshkin, A. (1992). Subjectivity in qualitative research. In M. D. LeCompte, W. L. Millroy, & J. Preissle (Eds.), *The handbook of qualitative research in education* (pp. 681-725). San Diego, CA: Academic Press.

Kidder, L. H. (1981). Qualitative research and quasi-experimental frameworks. In M. B. Brewer & B. E. Collins (Eds.), *Scientific inquiry and the social sciences.* San Francisco: Jossey-Bass.

Lave, C. A., & March, J. G. (1975). *An introduction to models in the social sciences.* New York: Harper & Row.

LeCompte, M. D., Millroy, W. L., & Preissle, J. (Eds.). (1992). *The handbook of qualitative research in education.* San Diego, CA: Academic Press.

LeCompte, M. D., & Preissle, J., with Tesch, R. (1993). *Ethnography and qualitative design in educational research* (2nd ed.). San Diego: Academic Press.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry.* Beverly Hills, CA: Sage.

Locke, L., Spirduso, W. W., & Silverman, S. J. (1993). *Proposals that work* (3rd ed.). Newbury Park, CA: Sage.

Marshall, C., & Rossman, G. (1995). *Designing qualitative research* (2nd ed.). Thousand Oaks, CA: Sage.

Maxwell, J. A. (1986). *The conceptualization of kinship in an Inuit community.* Unpublished doctoral dissertation, University of Chicago.

Maxwell, J. A. (1992). Understanding and validity in qualitative research. *Harvard Educational Review, 62,* 279-300.

Maxwell, J. A. (1996a). *Qualitative research design: An interactive approach.* Thousand Oaks, CA: Sage.

Maxwell, J. A. (1996b). *Using qualitative research to develop causal explanations.* Paper presented to the Task Force on Evaluating New Educational Initiatives, Committee on Schooling and Children, Harvard University.

Maxwell, J. A. (in press). Diversity, solidarity, and community. *Educational Theory.*

Maxwell, J. A., Cohen, R. M., & Reinhard, J. D. (1983). A qualitative study of teaching rounds in a department of medicine. In *Proceedings of the Twenty-second Annual Conference on Research in Medical Education.* Washington, DC: Association of American Medical Colleges.

Maxwell, J. A., & Miller, B. A. (n.d.). Categorization and contextualization in qualitative data analysis. Unpublished paper.

Menzel, H. (1978). Meaning: Who needs it? In M. Brenner, P. Marsh, & M. Brenner (Eds.), *The social contexts of method.* New York: St. Martin's.

Merriam-Webster's Collegiate Dictionary (10th ed.). (1993). Springfield MA: Merriam-Webster.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.

Mills, C. W. (1959). *The sociological imagination.* New York: Oxford University Press.

Mishler, E. G. (1986). *Research interviewing: Context and narrative.* Cambridge, MA: Harvard University Press.

Mohr, L. (1982). *Explaining organizational behavior.* San Francisco: Jossey-Bass.

Mohr, L. (1995). *Impact analysis for program evaluation* (2nd ed.). Thousand Oaks, CA: Sage.

Mohr, L. (1996). *The causes of human behavior: Implications for theory and method in the social sciences.* Ann Arbor: University of Michigan Press.

Norris, S. P. (1983). The inconsistencies at the foundation of construct validation theory. In E. R. House (Ed.), *Philosophy of evaluation* (pp. 53-74). San Francisco: Jossey-Bass.

Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn.* Cambridge: Cambridge University Press.

Oja, S. N., & Smulyan, L. (1989). *Collaborative action research: A developmental approach.* London: Falmer.

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: Sage.

Rabinow, P., & Sullivan, W. M. (1979). *Interpretive social science: A reader.* Berkeley: University of California Press.

Ragin, C. C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies.* Berkeley: University of California Press.

Reason, P. (1988). Introduction. In P. Reason (Ed.), *Human inquiry in action: Developments in new paradigm research* (pp. 1-17). Newbury Park, CA: Sage.

Reason, P. (1994). Three approaches to participative inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 324-339). Thousand Oaks, CA: Sage.

Regan-Smith, M. G. (1992). *The teaching of basic science in medical school: The students' perspective.* Unpublished doctoral dissertation, Harvard Graduate School of Education.

Robson, C. (1993). *Real world research: A resource for social scientists and practitioner-researchers.* London: Blackwell.

Rossi, P. H., & Berk, R. A. (1991). A guide to evaluation research theory and practice. In A. Fisher, M. Pavlova, & V. Covello (Eds.), *Evaluation and effective risk communications: Workshop proceedings* (pp. 201-254). Washington, DC: Interagency Task Force on Environmental Cancer and Heart and Lung Disease.

Sayer, A. (1992). *Method in social science: A realist approach* (2nd ed.). London: Routledge.

Scriven, M. (1967). The methodology of evaluation. In R. E. Stake (Ed.), *Perspectives on curriculum evaluation* (pp. 39-83). Chicago: Rand McNally.

Scriven, M. (1974). Maximizing the power of causal investigations: The modus operandi method. In W. J. Popham (Ed.), *Evaluation in education: Current applications* (pp. 68-84). Berkeley, CA: McCutchan.

Scriven, M. (1991). Beyond formative and summative evaluation. In M. W. McLaughlin & D. C. Phillips (Eds.), *Evaluation and education at quarter century* (pp. 19-64). Chicago: National Society for the Study of Education.

Seidman, I. E. (1991). *Interviewing as qualitative research.* New York: Teachers College Press.

Shweder, R. A. (Ed.). (1980). *Fallible judgment in behavioral research.* San Francisco: Jossey-Bass.

Strauss, A. L. (1987). *Qualitative analysis for social scientists.* New York: Cambridge University Press.

Strauss, A. L. (1995). Notes on the nature and development of general theories. *Qualitative Inquiry, 1,* 7-18.

Strauss, A. L., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* Newbury Park, CA: Sage.