Levels of Analysis

Data analysis is the process of making sense out of the data. And making sense out of data involves consolidating, reducing, and interpreting what people have said and what the researcher has seen and read—it is the process of making meaning. Data analysis is a complex process that involves moving back and forth between concrete bits of data and abstract concepts, between inductive and deductive reasoning, between description and interpretation. These meanings or understandings or insights constitute the *findings* of a study. Findings can be in the form of organized descriptive accounts, themes, or categories that cut across the data, or in the form of models and theories that explain the data. Each of these forms reflects different analytical levels, ranging from dealing with the concrete in simple description to high-level abstractions in theory construction.

The first section of this chapter focuses on data analysis related to descriptive accounts, category construction, and theory building. These levels of analysis occur in case study research as well as other types of qualitative research. However, multiple or comparative case study research has its own "levels" in the sense that there is "within-case" data analysis as well as "cross-case" analysis. The last section of this chapter is a discussion of data analysis in case study research.

Levels of Analysis

The most basic presentation of a study's findings is a descriptive account; even that description requires thinking through what will be included and what will be left out from the hundreds of pages of data collected for the study. Data are compressed and linked together

in a narrative that conveys the meaning the resc.. from studying the phenomenon. While description is . component of all forms of qualitative research, few studies a. ited to this level of analysis. LeCompte and Preissle (1993), for example, believe that ethnographers who "simply describe what they saw . . . fail to do justice to their data. By leaving readers to draw their own conclusions, researchers risk misinterpretation. Their results also may be trivialized by readers who are unable to make connections implied, but not made explicit, by the researcher" (p. 267).

Category Construction

Moving beyond basic description to the next level of analysis, the challenge is to construct categories or themes that capture some recurring pattern that cuts across "the preponderance" (Taylor and Bogdan, 1984, p. 139) of the data. These categories or themes are "concepts indicated by the data (and not the data itself). . . . In short, conceptual categories and properties have a life apart from the evidence that gave rise to them" (p. 36). Devising categories is largely an intuitive process, but it is also systematic and informed by the study's purpose, the investigator's orientation and knowledge, and the meanings made explicit by the participants themselves. Typically, guidelines for category construction found in the literature are very general "and their applications are subject to the situational demands of a given study" (Constas, 1992, p. 255).

Categories and subcategories (or properties) are most commonly constructed through the constant comparative method of data analysis. As I explained in Chapter Eight, at the heart of this method is the continuous comparison of incidents, respondents' remarks, and so on, with each other. Units of data—bits of information—are literally sorted into groupings that have something in common. A *unit of data* is any meaningful (or potentially meaningful) segment of data; at the beginning of a study the researcher is uncertain about what will ultimately be meaningful. A unit of data can be as small as a word a participant uses to describe a feeling or phenomenon, or as large as several pages of field notes describing a particular incident. According to Lincoln and Guba (1985) a unit must meet two criteria. First, it should be heuristic—that is, the unit should reveal information relevant to the study and stimulate the reader to think beyond the particular bit of

information. Second, the unit should be "the smallest piece of information about something that can stand by itself—that is, it must be interpretable in the absence of any additional information other than a broad understanding of the context in which the inquiry is carried out" (p. 345).

The task is to compare one unit of information with the next in looking for recurring regularities in the data. The process is one of breaking data down into bits of information and then assigning "these bits to categories or classes which bring these bits together again if in a novel way. . . . In the process we begin to discriminate more clearly between the criteria for allocating data to one category or another. Then some categories may be subdivided, and others subsumed under more abstract categories" (Dey, 1993, p. 44).

For a simple but vivid example of how to take raw data and sort them into categories, consider the task of sorting two hundred food items found in a grocery store. These two hundred food items in a research study would be bits of information or units of data upon which to base an analysis. By comparing one item with another, the two hundred items could be classified into any number of categories. Starting with a box of cereal, for example, you could ask whether the next item, an orange, is like the first. Obviously not. There are now two piles into which the next item may or may not be placed. By this process you can sort all the items into categories of your choice. One scheme may separate the items into the categories of fresh, frozen, canned, or packaged goods. Or you could divide them by color, weight, or price. More likely, you would divide the items into common grocery store categories: meat, dairy, produce, canned goods, and so on. These categories would be fairly comprehensive classes, each of which could be further subdivided. Produce, for example, includes the subcategories of fruits and vegetables. Fruits include citrus and noncitrus, domestic and exotic. All these schemes emerge logically from the "data"—the food items. The names of the categories and the scheme you use to sort the data will reflect the focus of your study.

The Step-by-Step Process

Category construction *is* data analysis, and all of the caveats about this process I discussed in the previous chapter should be kept in mind, the most important being that data analysis is done in con-

junction with data collection. There is, however, a growing preoccupation with analysis in proportion to collection as the study progresses. And once all of the data are in, there is generally a period of intensive analysis when tentative findings are substantiated, revised, and reconfigured.

Category construction begins with reading the first interview transcript, the first set of field notes, the first document collected in the study. As you read down through the transcript, for example, you jot down notes, comments, observations, and queries in the margins. These notations are next to bits of data that strike you as interesting, potentially relevant, or important to your study. Think of yourself as having a conversation with the data, asking questions of it, making comments to it, and so on. "The notes serve to isolate the initially most striking, if not ultimately most important, aspects of the data" (LeCompte, Preissle, and Tesch, 1993, p. 236).

After working through the entire transcript in this manner, you go back over your marginal notes and comments and try to group those comments and notes that seem to go together. This is akin to sorting items in the grocery store example. Keep a running list of these groupings attached to the transcript or on a separate paper or memo to yourself. At the beginning of an inquiry, this list is likely to be fairly long because you do not yet know what will surface across the rest of the data. You also will not yet know which groupings might be subsumed under others.

Moving to your next set of data (transcript, field notes, or document), you scan it in exactly the same way as just outlined. You do this, keeping in mind the list of groupings that you extracted from the first transcript, checking to see if they are also present in this second set. You also make a separate list of comments, terms, and notes from this set and then *compare* this list with the one derived from the first transcript. These two lists should then be merged into one master list of concepts derived from both sets of data. This master list constitutes a primitive outline or classification system reflecting the recurring regularities or patterns in your study. These patterns and regularities become the categories or themes into which subsequent items are sorted.

It should be clear that categories are abstractions derived from the data, not the data themselves. To paraphrase Glaser and Strauss (1967), these categories have a life of their own apart from the data

from which they came. Categories are conceptual elements that "cover" or span many individual examples of the category. This is illustrated in Figure 9.1.

The shaded background boxes represent incidents of the category from which the category was derived. The position of the category vis-à-vis the data is represented by the lighter ellipse in the figure.

Naming the Categories

The names of your categories can come from at least three sources: the researcher, the participants, or sources outside the study such as the literature. The most common situation is when the investigator comes up with terms, concepts, and categories that reflect what he or she sees in the data. In the second approach, the data

Figure 9.1. Deriving Categories from Data.
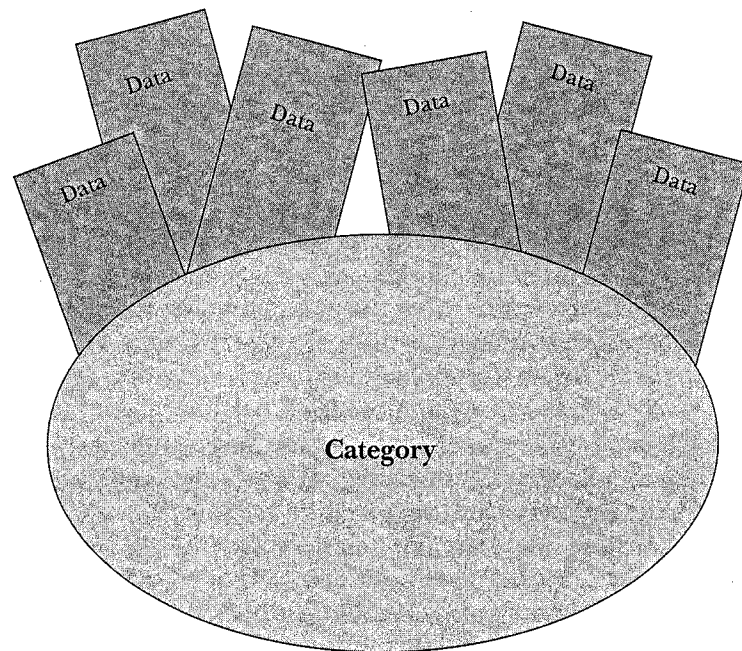
can be organized into a scheme suggested by the participants themselves. "This approach requires an analysis of the verbal categories used by participants and/or staff in a program to break up the complexity of reality into parts. It is a fundamental purpose of language to tell us what is important by giving it a name and therefore separating it from other things with other names" (Patton, 1990, p. 393). Patton gives the example of teachers' classification of dropouts into "chronics" and "borderlines" (p. 394). Bogdan and Biklen (1992) found that parents were classified by professional staff as "good parents," "not-so-good parents," or "troublemakers" (p. 169).

In addition to the participants' own categories, classification schemes can be borrowed from sources outside the study at hand. Applying someone else's scheme requires that the categories be compatible with the purpose and theoretical framework of the study. The database is scanned to determine the fit of a priori categories, and then the data are sorted into the borrowed categories.

There is some danger in using borrowed classification schemes, however. As Glaser and Strauss (1967) point out, "Merely selecting data for a category that has been established by another theory tends to hinder the generation of new categories, because the major effort is not generation, but data selection. Also, emergent categories usually prove to be the most relevant and the best fitted to the data. . . . Working with borrowed categories is more difficult since they are harder to find, fewer in number, and not as rich; since in the long run they may not be relevant, and are not exactly designed for the purpose, they must be respecified" (p. 37).

Several important guidelines can be used to determine the efficacy of categories derived from the constant comparative method of data analysis:

- Categories should *reflect the purpose of the research*. In effect, categories are the *answers* to your research question(s). One of Rowden's (1994) findings (or categories) regarding the role of human resource development (HRD) in successful small companies was that HRD activities serve to preserve the market niche of these companies. This category, "preserve the market niche," was one "answer" to the study's question of how human resource development functions in the success of small businesses.

- Categories should be *exhaustive*; that is, you should be able to place all data that you decided were important or relevant to the study in a category or subcategory.
- Categories should be *mutually exclusive*. A particular unit of data should fit into only one category. If the exact same unit of data can be placed into more than one category, more conceptual work needs to be done to refine your categories.
- Categories should be *sensitizing*. The naming of the category should be as sensitive as possible to what is in the data. An outsider should be able to read the categories and gain some sense of their nature. The more exacting in capturing the meaning of the phenomenon, the better. For example, the category "leadership" does not reveal as much as the category "charismatic leadership." In another example, "defiant behavior" is not as sensitizing as "defiance of adult authority figures."
- Categories should be *conceptually congruent*. This means that the same level of abstraction should characterize all categories at the same level. In the grocery store example described earlier, the items should not be sorted according to produce, canned goods, and fruit. While produce and canned goods are at the same conceptual level, fruit is a type of produce and should form a subcategory of produce.

Conceptual congruence is probably the most difficult criterion to apply. Investigators are usually so immersed in their data and their analysis that it is hard for them to see whether or not a set of categories make sense together. One of the best strategies for checking all the criteria against your category scheme is to display your set of categories in the form of a chart or table. This can be as simple as a list of one-word categories. In a study of the structure of simple reminiscence (Merriam, 1989), for example, the categories or findings were displayed in a list consisting of four terms—selection, immersion, withdrawal, and closure. Data displays can also be quite complex (Miles and Huberman, 1994). The point is that by laying out the basic structure of your findings in front of you, you can see how well all of the parts fit together. Also, by writing out the purpose statement at the top of your display, you can immediately see whether the categories are answers to the research question.

How Many Categories?

The number of categories a researcher constructs depends on the data and the focus of the research. In any case, the number should be manageable. In my experience, the fewer the categories, the greater the level of abstraction, and the greater ease with which you can communicate your findings to others. A large number of categories is likely to reflect an analysis too lodged in concrete description. Guba and Lincoln (1981) suggest four guidelines for developing categories that are both comprehensive and illuminating. First, the number of people who mention something or the frequency with which something arises in the data indicates an important dimension. Second, the audience may determine what is important—that is, some categories will appear to various audiences as more or less credible. Third, some categories will stand out because of their uniqueness and should be retained. And fourth, certain categories may reveal "areas of inquiry not otherwise recognized" or "provide a unique leverage on an otherwise common problem" (p. 95).

Several guidelines can help a researcher determine whether a set of categories is complete. First, "there should be a minimum of unassignable data items, as well as relative freedom from ambiguity of classification" (Guba and Lincoln, 1981, p. 96). Moreover, the set of categories should seem plausible given the data from which they emerge, causing independent investigators to agree that the categories make sense in light of the data. This strategy helps to ensure reliability and is discussed further in Chapter Ten.

Systems for Placing Data into Categories

Once you are satisfied with the set of categories derived from the data, the categories can be fleshed out and made more robust by searching through the data for more and better units of relevant information. The mechanical handling of the data at this stage of the analysis warrants some attention. Four basic strategies for organizing all the data in preparation for further analysis or for writing the results of the study include using index cards, file folders, information retrieval cards, and computer programs.

Each unit of information can be put onto a separate *index card* and coded according to any number of categories ranging from situational factors (who, what, when, where) to categories representing

emerging themes or concepts. You could also first code the units in the margins of the interview transcripts, field notes, or documents. Photocopies can be made of the pages where you have made comments or notations or have identified tentative categories. Each photocopied unit of information is then placed on an index card; these can be sorted into piles by constantly comparing the information on one card with the information on the next. The piles are labeled, and the cards within that pile are coded accordingly. Once all the cards have been coded, cards relevant to a certain category can be retrieved by the code on the card.

A variation of the index card method is the use of *information retrieval cards*. At least two commercially available systems are available that can be used to sort data (Werner and Schoepfle, 1987). McBee or Indecks cards are large, index-type cards with numbered holes around the margins. First, the researcher pastes a photocopied unit of data on a card or types data onto a card. Second, the data are assigned a number representing a category. Third, the researcher punches out the corresponding numbered hole found on the margin of the card. A large rod, somewhat like a knitting needle, can then be passed through the same numbered hole on all of the cards; lifting and shaking the needle full of cards allows the coded and punched cards to fall out. When this has been done for all categories, all the data pertaining to specific, coded categories are grouped. This technique allows the cards to be left in random order, since all relevant information is coded by using the holes at the edge of the card.

If *file folders* are used, a photocopy of the entire database or case record is made. Working page by page as outlined earlier in this chapter, the researcher writes notations in the margins, including tentative categories or themes emerging from the raw data. The photocopied pages are then cut up, and coded sections are placed into file folders labeled by category or theme. Each unit of data needs to be coded not only by category but by its original page number and possibly by other identifying codes such as respondent's name and so on. If need be, each cut-up piece of information can be located later in the master copy.

Numerous *computer programs* have been developed to store, sort, and retrieve qualitative data. Some researchers have also devised systems using powerful word processing packages or database pro-

grams. Interview transcripts, observation notes, and so on are entered verbatim into the computer. The program then numbers each line of the database. The researcher uses a hard copy of the numbered database to analyze the data, making notes in the margins and developing themes or categories. Going from the hard copy back to the computer file, categories and their corresponding line numbers are entered. The researcher can then retrieve and print, by category, any set of data desired. Multiple levels of coding are possible for the same unit of information. (See Chapter Eight for more discussion of computers in qualitative research).

No doubt every researcher devises his or her own scheme for handling qualitative data. The four strategies for sorting data just presented allow for the easy retrieval of data and for cross-analysis of coded categories. Cross-analysis is especially important in a level of analysis that goes beyond a categorical or taxonomic integration of the data toward the development of theory. The development of theory is the most sophisticated level of data analysis and is discussed in the next section.

Developing Theory

Several levels of data analysis are possible in a qualitative case study. At the most basic level, data are organized chronologically or sometimes topically and presented in a narrative that is largely, if not wholly, descriptive. Moving from concrete description of observable data to a somewhat more abstract level involves using concepts to describe phenomena. Rather than just describing a classroom interaction, for example, a researcher might cite it as an instance of "learning" or "confrontation" or "peer support," depending on the research problem. This is the process of systematically classifying data into some sort of schema consisting of categories, themes, or types. The categories describe the data, but to some extent they also interpret the data. A third level of analysis involves making inferences, developing models, or generating theory. It is a process, Miles and Huberman (1994) write, of moving up "from the empirical trenches to a more conceptual overview of the landscape. We're no longer just dealing with observables, but also with unobservables, and are connecting the two with successive layers of inferential glue" (p. 261).

Thinking about data—*theorizing*—is a step toward developing a theory that explains some aspect of educational practice and allows a researcher to draw inferences about future activity. Theorizing is defined as "the cognitive process of discovering or manipulating abstract categories and the relationships among those categories" (LeCompte, Preissle, and Tesch, 1993, p. 239). It is fraught with ambiguity. "The riskiness of going beyond the data into a never-never land of inference" is a difficult task for most qualitative researchers because they are too close to the data, unable to articulate how the study is significant, and unable to shift into a speculative mode of thinking (p. 269). Theorizing about data can also be hindered by thinking that is linear rather than contextual. Patton (1990) notes the temptation to "fall back on the linear assumptions of quantitative analysis," which involves specifying "isolated variables that are mechanically linked together out of context" (p. 423). Such noncontextual statements "may be more distorting than illuminating. It is the ongoing challenge, paradox, and dilemma of qualitative analysis that we must be constantly moving back and forth between the phenomenon of the program and our abstractions of that program, between the descriptions of what has occurred and our analysis of those descriptions, between the complexity of reality and our simplifications of those complexities, between the circularities and interdependencies of human activity and our need for linear, ordered statements of cause-effect" (pp. 423–424).

Nevertheless, data often seem to beg for continued analysis past the formation of categories. A key here is when the researcher knows that the category scheme does not tell the whole story—that there is more to be understood about the phenomenon. This often leads to trying to link the conceptual elements—the categories—together in some meaningful way. One of the best ways to try this out is to visualize how the categories work together. Even a simple diagram or model using the categories and subcategories of the data analysis can effectively capture the interaction or relatedness of the findings. In Johnson-Bailey and Cervero's (1996) study of black reentry women, for example (see Chapter Three), her categories of silence, negotiation, and resistance were the strategies the black women in their study used to survive or succeed in formal education. However, these strategies were used both inside and outside the classroom, and they intersected with systems of race, gender, class, and color. Figure 9.2 shows a model displaying these interrelationships.

Johnson-Bailey and Cervero (1996) explain it as follows:

The issues of race, gender, class, and color are depicted in the background surrounding the circle to indicate powerful forces which are ever-present in society. The center circle overlaps the smaller circles which represent the different segments of society: school, workplace, community, and family. The obstacles they encountered in school were no different than those experienced in the other three areas. To cope with old dilemmas, the women relied on familiar strategies (silence, negotiation, and resistance)

Figure 9.2. Linking Categories and Concepts in a Model of Reentry Black Women's Experience.

Rentry of Black Women
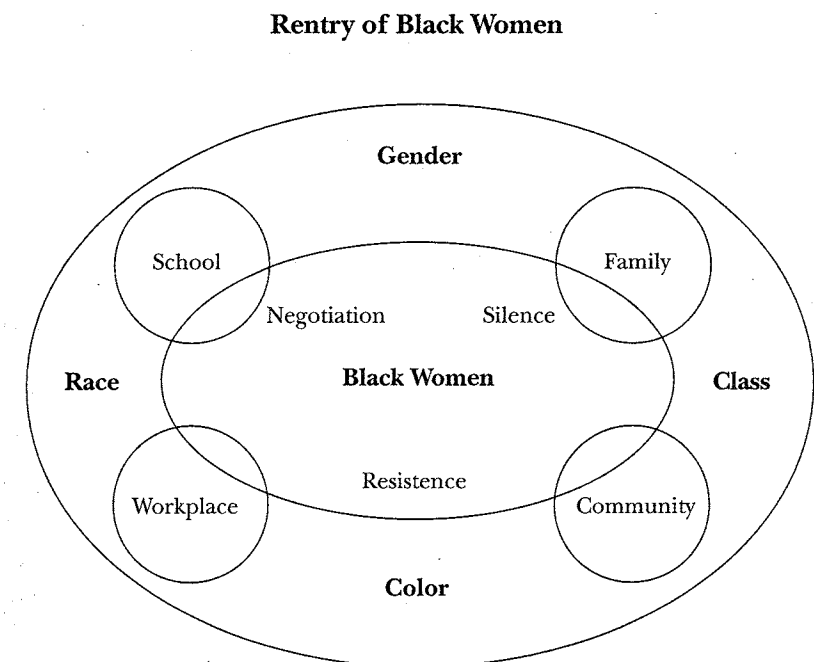
Source: Johnson-Bailey and Cervero (1996, p. 153). Reprinted with permission.

they had used throughout their lives, to respond to the direct impact of racism, sexism, classism, and colorism in these four social sites [p. 154].

Speculation is a key to developing theory in a qualitative study. Speculation involves "playing with ideas probabilistically. It permits the investigator to go beyond the data and make guesses about what will happen in the future, based on what has been learned in the past about constructs and linkages among them and on comparisons between that knowledge and what presently is known about the same phenomena. These guesses are projections about how confidently the relationships found or explanations developed can be expected to obtain in the future" (Goetz and LeCompte, 1984, p. 173).

The qualitative investigator who wishes to derive theory from data can turn to Glaser and Strauss (1967) and Strauss (1987) for assistance. They have devised a strategy for developing substantive theory—theory that applies to a specific aspect of educational practice. Since the theory is grounded in the data and emerges from them, the methodology is called *grounded theory* (see Chapter One). A grounded theory begins with categories. In addition to categories, a theory consists of two other elements—properties and hypotheses. Properties are also concepts but ones that describe a category; properties are not examples of a category but dimensions of it. The category "career malaise," for example, is defined by the properties of "boredom," "inertia," and "trapped" (Merriam, 1980). In a grounded theory study of the epistemological development of Malaysian women, Reybold (1996) found that the process was one of negotiating their culture with their personal sense of self. The category, "defining a personal model of self," had three properties representing how this was done: outright *adoption* of cultural definitions; *adaptation* of the culture; or *construction* of a personal model that disregarded cultural expectations.

Hypotheses are the suggested links between categories and properties. In Reybold's (1996) study of epistemological development cited earlier, she hypothesizes that family support of education for girls and women, diverse learning experiences, and extended international opportunities, are factors that foster epistemological development. In another grounded theory study of a

faculty's participation in in-service workshops, for example, the researcher cited "workshop credibility" as one of several categories explaining faculty participation (Rosenfeldt, 1981). A property that helped to define workshop credibility was called "identification with sponsoring agent." The author hypothesized that "workshop participation will depend on the extent to which faculty members identify with the workshop sponsors. Namely, the greater the identification of the potential participants with the sponsoring agent, the greater the likelihood that professors will participate in a given workshop" (Rosenfeldt, 1981, p. 189). Such hypotheses emerge simultaneously with the collection and analysis of data. The researcher tries to support tentative hypotheses while at the same time remaining open to the emergence of new hypotheses. "Generating hypotheses requires evidence enough only to establish a suggestion—not an excessive piling up of evidence to establish a proof" (Glaser and Strauss, 1967, pp. 39–40).

The development of categories, properties, and tentative hypotheses through the constant comparative method (Glaser and Strauss, 1967) is a process whereby the data gradually evolve into a core of emerging theory. This core is a theoretical framework that guides the further collection of data. Deriving a theory from the data involves both the integration and the refinement of categories, properties, and hypotheses. As the theory solidifies, "major modifications become fewer and fewer as the analyst compares the next incidents of a category to its properties. Later modifications are mainly on the order of clarifying the logic, taking out non-relevant properties, integrating elaborating details of properties into the major outline of interrelated categories" (p. 110). In short, more data can be processed with fewer adjustments because the theory emerges "with a smaller set of higher level concepts" (p. 110).

At this point, "with reduction of terminology and consequent generalizing . . . the analyst starts to achieve two major requirements of theory: (1) *parsimony* of variables and formulation, and (2) *scope* in the applicability of the theory to a wide range of situations" (pp. 110–111, emphasis in original). Besides parsimony and scope, the emergent theory can be evaluated in terms of its overall explanatory power, by how well the generalizations are supported, by how well integrated the elements are, and by whether there is a logical consistency to every dimension of the theory.

Those who build theory in an applied field such as education need also be concerned with how well the theory fits the substantive area to which it will be applied, whether laypersons will be able to understand and use the theory, and whether the person who uses the theory will "have enough control in everyday situations to make its application worthwhile" (p. 245).

While building theory in the manner described by Glaser and Strauss is largely an inductive process, there are times throughout the investigation when a deductive strategy is used. Tentative categories, properties, and hypotheses continually emerge and must be tested against the data—that is, the researcher asks if there are sufficient data to support a certain category or hypothesis. If so, the element is retained; if not, it is discarded. Thus the researcher is continually shifting back and forth between deductive and inductive modes of thinking. For Glaser and Strauss (1967) the difference is one of emphasis. "Verifying as much as possible with as accurate evidence as feasible is requisite while one discovers and generates his theory—but *not* to the point where verification becomes so paramount as to curb generation" (p. 28).

In summary, data analysis is a process of making sense out of data. It can be limited to determining how best to arrange the material into a narrative account of the findings. More commonly, researchers extend analysis to developing categories, themes, or other taxonomic classes that interpret the meaning of the data. The categories become the findings of the study. When categories and their properties are reduced and refined and then linked together by tentative hypotheses, the analysis is moving toward the development of a theory to explain the data's meaning. This third level of analysis transcends the formation of categories, for a theory seeks to explain a large number of phenomena and tell how they are related. In a nice summary of the process, Dey (1993) compares qualitative data analysis to climbing a mountain to see the view.

> First of all, we must insist that our mountain rises above the plain world of common sense to afford a more "scientific" perspective. . . . We can allow our mountain to be of any size and shape; the small hill of a short undergraduate project, or the precipitous peak of a large-scale research project. . . . for the most part

much the same tasks are required of both. The mountain is climbed bit by bit, and while we are climbing, we focus on one step at a time. But the view we obtain is more than the sum of the sequence of steps we take along the way. Every so often, we can turn and look to the horizon, and in doing so we see the surrounding country from a fresh vantage point. . . . This climb, with its circuitous paths, its tangents and apparent reversals, and its fresh vistas, reflects the creative and non-sequential character of the analytic process. Progress may be slow and laborious, but it can be rewarded with some breath-taking revelations [pp. 53–54].

Data Analysis in Case Studies

While the basic strategy for analyzing data outlined earlier in this chapter and in Chapter Eight applies to all types of qualitative research, some features of case studies affect data analysis. First, a case study is an intensive, holistic description and analysis of a single, bounded unit. Conveying an understanding of the case is the paramount consideration in analyzing the data. Stake (1995) explains:

> Keeping in mind that it is *the case* we are trying to understand, we analyze episodes or text materials with a sense of correspondence [by which Stake means "consistency within certain conditions"]. We are trying to understand behavior, issues, and contexts with regard to our particular case. . . . We try to find the pattern or the significance through direct interpretation, just asking ourselves "What did that mean?" For more important episodes or passages of text, we must take more time, looking them over again and again, reflecting, triangulating, being skeptical about first impressions and simple meanings [p. 78, emphasis in original].

In case studies, communicating understanding—the goal of the data analysis—is linked to the fact that data have usually been derived from interviews, field observations, and documents. In addition to a tremendous amount of data, this range of data sources may present disparate, incompatible, even apparently contradictory information. The case study researcher can be seriously challenged in trying to make sense out of the data. Attention to data management is particularly important under these circumstances.

To begin the more intensive phase of data analysis in a case study, all the information about the case should be brought together—interview logs or transcripts, field notes, reports, records, the investigator's own documents, physical traces, and reflective memos. All this material needs to be organized in some fashion so that data are easily retrievable. Yin (1994) calls this organized material the *case study data base*, which he differentiates from the case study report. In a similar fashion, Patton (1990) differentiates the *case record* from the final *case study*. "The case record pulls together and organizes the voluminous case data into a comprehensive primary resource package. The case record includes all the major information that will be used in doing the case analysis and case study. Information is edited, redundancies are sorted out, parts are fitted together, and the case record is organized for ready access either chronologically and/or topically. The case record must be complete but manageable" (p. 386–387). The case study database (or record) then, is the data of the study organized so the researcher can locate specific data during intensive analysis.

The various procedures for deriving meaning from qualitative data described in this and the preceding chapter apply to the single case study. While the final write-up or case report may have a greater proportion of description than other forms of qualitative research in order to convey a holistic understanding of the case, the level of interpretation may also extend to the presentation of categories, themes, models, or theory. (For discussion of the final report, see Chapter Eleven).

Multiple or comparative case studies involve collecting and analyzing data from several cases. Instead of studying one good high school, for example, Lightfoot (1983) studied six. Her findings are presented first as six individual case studies (or "portraits" as she calls them); she then offers a cross-case analysis leading to generalizations about what constitutes a good high school.

In a multiple case study, there are two stages of analysis—the within-case analysis and the cross-case analysis. For the *within-case analysis*, each case is first treated as a comprehensive case in and of itself. Data are gathered so the researcher can learn as much about the contextual variables as possible that might have a bearing on the case. The data of the single qualitative case are analyzed as described here and in Chapter Eight. Once the analysis of each case

is completed, *cross-case analysis* begins. A qualitative, inductive, multicase study seeks to build abstractions across cases. The researcher attempts "to build a general explanation that fits each of the individual cases, even though the cases will vary in their details" (Yin, 1994, p. 112). The researcher attempts to see "processes and outcomes that occur across many cases, to understand how they are qualified by local conditions, and thus develop more sophisticated descriptions and more powerful explanations" (Miles and Huberman, 1994, p. 172). Miles and Huberman warn that "cross-case analysis is tricky. Simply summarizing superficially across some themes or main variables by itself tells us little. We have to look carefully at the complex configuration of processes within each case, understand the local dynamics, before we can begin to see patterning of variables that transcends particular cases" (pp. 205–206).

As with the single case study, one of the challenges in a multicase study is in the *management* of the data; the researcher probably has considerably more raw information and must find ways to handle it without becoming overwhelmed. Cross-case studies are also likely to involve a team of investigators, each studying an assigned site. Clearly then, coordination of both personnel and data is called for. Indeed, ongoing collaboration in data analysis is essential in large-scale studies. Also essential are coordinated systems for recording data. Miles and Huberman (1994) have developed numerous methods for analyzing data from several cases or sites. The methods range from simple to complex, from descriptive to explanatory, and all involve devising matrices and other visual aids for displaying data across sites.

Ultimately, cross-case analysis differs little from analysis of data in a single qualitative case study. The level of analysis can result in little more than a unified description across cases; it can lead to categories, themes, or typologies that conceptualize the data from all the cases; or it can result in building substantive theory offering an integrated framework covering multiple cases.

A variation of cross-case or cross-site studies is the *case survey*. This is a form of secondary analysis in that the case studies have already been conducted and are available to the researcher. They function as databases to answer new questions or confirm new interpretations. Stenhouse (1978) proposed establishing a repository of case records in education so that these databases can be

easily accessed by researchers. The Human Relations Area File (HRAF) I mentioned in Chapter Six is an anthropological database containing numerous ethnographic case studies. And in some ways, the Educational Resources Information Center (ERIC) functions as a repository, since it contains thousands of documents, reports, and studies in education. It does not separate case studies from other materials, however, so it would take some effort to identify case studies in particular.

The purpose of a case survey is to aggregate "diverse case studies together under a common conceptual framework so that findings will be cumulative . . . to identify what it is we already 'know,' what it is we do not know, and what it is we suspect" (Lucas, quoted in Guba and Lincoln, 1981, p. 247). The basic strategy for conducting a case survey differs somewhat from data analysis in cross-case studies. First, the researcher must determine the criteria by which cases are to be selected for analysis. This step requires "a tight definition of the phenomenon under investigation" (Guba and Lincoln, 1981, p. 250). Second, case surveys tend to be quantitative in nature, although they need not be. West and Oldfather (1995) suggest an innovative way to do a cross-case comparison of qualitative case studies. Pooled case comparison allows the comparison of "separate but similar studies *ex post facto*; like overlaying of one transparency on another, this method highlights both the uniqueness and the commonality of participants' experiences and allows us to understand each study more fully" (p. 454). Raw data from separate studies of similar phenomena are "pooled" into a data set for a fresh analysis.

In summary, data analysis in case studies must account for some of the identifying features of this particular type of qualitative research, including the focus on understanding and the typically broad range of data available for analysis. In a multiple case study, a within-case analysis is followed by a cross-case analysis.

Summary

The analysis of qualitative data can range from organizing a narrative description of the phenomenon, to constructing categories or themes that cut across the data, to building theory. Each of these levels of analysis calls upon the investigator's intuitive as well

as analytical powers. The process can certainly be enhanced by employing techniques that have helped others, such as using data displays, as well as devising a systematic approach to the task. One particular approach, that of the constant comparative method of data analysis, was used in this chapter to demonstrate how to construct meaning from qualitative data.

I covered in some detail the ins and outs of category construction. The step-by-step process includes naming the categories, determining the number of categories, and figuring out systems for placing data into categories. Using categories as the basic conceptual element, I discussed how analysis can be extended to theory building. Finally, I discussed data analysis in case study research, with particular attention to within-case and cross-case analysis.

Mining Data from Documents

Interviewing and observing are two data collection strategies designed to gather data that specifically address the research question. Documents, however, are usually produced for reasons other than the research at hand and therefore are not subject to the same limitations. The presence of documents does not intrude upon or alter the setting in ways that the presence of the investigator often does. Nor are documents dependent upon the whims of human beings whose cooperation is essential for collecting good data through interviews and observations. Documents are, in fact, a ready-made source of data easily accessible to the imaginative and resourceful investigator. This chapter examines the nature of documents, various types of documents, their use in qualitative research, and their limitations and strengths. The last section of the chapter presents a look at a relatively new type of documents and data—those obtained on-line.

Nature of Documents

A number of terms are used to refer to sources of data in a study other than interviews or observations. I have chosen the term *document* as the umbrella term to refer to a wide range of written, visual, and physical material relevant to the study at hand. This term includes materials "in the broad sense of any communication"—for example novels, newspapers, love songs, diaries, psychiatric interviews, and the like (Holsti, 1969, p. 1). Documents, as the term is used in this chapter, also include what LeCompte

and Preissle (1993) define as artifacts—"symbolic materials such as writing and signs and nonsymbolic materials such as tools and furnishings" (p. 216). Others use the term "available" materials or data. This means just about anything in existence prior to the research at hand. "The researcher may make use of letters or television transcripts, historical documents or journalistic accounts, tribal artifacts or works of art. He may analyze the records of corporations, police courts, or the U.S. Bureau of the Census. He may reexamine . . . the already completed studies of other scholars. As all these and diverse other materials accumulate, it may well be that increasing numbers of researchers will find that the data they need have already been gathered" (Selltiz, Jahoda, Deutsch, and Cook, 1959, pp. 240–241). Photographs, film, and video can also be used as data sources, as can physical evidence or traces (Webb and others, 1981; Harper, 1994). Although this chapter concentrates on written documents, the general discussion applies to all forms of data not gathered through interviews or observations.

Types of Documents

Public records, personal documents, and physical material are three major types of documents available to the researcher for analysis. Moreover, a researcher can create documents for the purpose of the investigation.

Public Records

Public records are "the ongoing, continuing records of a society" (Webb and others, 1981, p. 78). As Guba and Lincoln (1981) note, "The first and most important injunction to anyone looking for official records is to presume that if an event happened, some record of it exists" (p. 253). Public documents include actuarial records of births, deaths, and marriages, the U.S. census, police records, court transcripts, agency records, association manuals, program documents, mass media, government documents, and so on. Locating public records is limited only by the researcher's imagination and industriousness. Auster (1985), for example, demonstrates how to conduct a study of changing social expectations for family, career, gender roles, and sexual behavior through

the sole data source of Girl Scout handbooks. Youth organization handbooks, she points out, "represent the intersection of biography and history" (p. 359), providing an excellent data source for studying changing social mores.

For those interested in educational questions, there are numerous sources of public documents—discussions of educational issues and bills in the *Congressional Record;* federal, state, and private agency reports; individual program records; and the statistical database of the Center for Educational Statistics. Since many case studies are at the program level, it is particularly important to seek out the paper trail for what it can reveal about the program—"things that cannot be observed," things "that have taken place before the evaluation began. They may include private interchanges to which the educator would not otherwise be privy. They can reveal goals or decisions that might be unknown to the evaluator" (Patton, 1990, p. 233). Ideally this paper trail includes "all routine records on clients, all correspondence from and to program staff, financial and budget records, organizational rules, regulations, memoranda, charts, and any other official or unofficial documents generated by or for the program" (p. 233). Such documents not only provide valuable information about the program itself, but they can also stimulate thinking "about important questions to pursue through more direct observations and interviewing" (p. 233).

If you were interested in studying the role of parent involvement in a neighborhood school, for example, you could look for public record documents in the form of the following: notices sent home to parents; memos between and among teachers, staff, and the parents' association; formal policy statements regarding parent involvement; school bulletin boards or other displays featuring aspects of parent involvement; newspaper and other media coverage of activities featuring parent involvement; and any official records of parent attendance or presence in the school.

Other sources of public information that are easily accessible but often overlooked include previous studies and data "banks" of information. However, in using these resources the researcher has to rely on someone else's description and interpretation of data rather than use the raw data as a basis for analysis. These meta-analyses, as they are called, are more common in quantitative research, although there has been some recent thinking as to how this strategy might

apply to qualitative studies (LeCompte and Preissle, 1992; West and Oldfather, 1995). For large-scale or cross-cultural research, relying on previous studies may be the only realistic way to conduct the investigation.

An example of a data bank that is potentially useful in qualitative research, especially ethnographic studies (see Chapter One), is the Human Relations Area File (Murdock, 1983; Murdock and others, 1982). This file is a compilation of ethnographic studies of more than 350 societies; data are classified and coded by cultural group and also by more than 700 topics. Education is one broad topic under which such subtopics as elementary education, educational theory and methods, students, and vocational education can be found. The index is organized so that a researcher can retrieve documents related to the educational practices of one particular cultural group, or documents can be retrieved about a specific practice such as "student uprisings" across many cultures. Types of documents found in this file include ethnographer field notes, diary entries, reports to various agencies, books, newspaper articles, works of fiction about the culture, and photographs.

"Every literate society," writes Kidder (1981b), "produces a variety of material intended to inform, entertain, or persuade the populace" (p. 286). Popular media forms such as television, films, radio, newspapers, literary works, photography, cartoons, and the like are another source of "public" data. Mass communication materials are especially good sources for dealing with questions about some aspect of society at a given time, for comparing groups on a certain dimension, or for tracking cultural change and trends. They "concentrate on what is of current interest, and that concentration makes it possible to track many phenomena and index the growth and decline of public interest in them" (Webb and others, 1981, p. 120). Studies have been conducted, for example, on the roles of blacks in television, the presence of ageism in cartoons, and teenage culture in movies.

Personal Documents

In contrast to public sources of data, personal documents "refer to any first-person narrative that describes an individual's actions, experiences, and beliefs" (Bogdan and Biklen, 1992, p. 132). Such

documents include diaries, letters, home videos, sermons, children's growth records, scrapbooks and photo albums, calendars, autobiographies, and travel logs. Selltiz, Jahoda, Deutsch, and Cook (1959) note that "the rationale for the use of personal documents is similar to that for the use of observational techniques. What the latter may achieve for overt behavior, the former can do for inner experiences: to reveal to the social scientist life as it is lived without the interference of research" (p. 325). Such documents can tell the researcher about the inner meaning of everyday events, or they may yield descriptions of "rare and extraordinary events in human life" (p. 327) such as can be found in Admiral Byrd's report of his experiences alone at the South Pole or Helen Keller's account of overcoming extraordinary physical handicaps.

Personal documents are a reliable source of data concerning a person's attitudes, beliefs, and view of the world. But because they are personal documents, the material is highly subjective in that the writer is the only one to select what he or she considers important to record. Obviously these documents are not representative or necessarily reliable accounts of what actually may have occurred. They do, however, reflect the participant's perspective, which is what most qualitative research is seeking. Burgess (1982) summarizes the nature of personal documents:

> The field researcher needs to consider: Is the material trustworthy? Is the material atypical? Has the material been edited and refined? Does the autobiographical material only contain highlights of life that are considered interesting? Furthermore, it could be argued that the material is automatically biased as only certain people produce autobiographies and keep diaries; there is self-selectivity involved in the sample of material available; they do not provide a complete historical record. Nevertheless, such material does provide a subjective account of the situation it records; it is a reconstruction of part of life. Furthermore, it provides an account that is based on the author's experience [p. 132].

An entire study can be based on personal documents. Abramson's (1992) case study of Russian Jewish emigration is based solely on his grandfather's diaries written over a twelve–year period. A well-known earlier study of Polish immigrant life relied heavily

upon personal letters written between immigrants and relatives in Europe (Thomas and Znaniecki, 1927). Many of these letters were obtained by placing ads in local newspapers asking for them.

Physical Material

Physical material as a form of document, broadly defined, consists of physical objects found within the study setting. Anthropologists typically refer to these objects as *artifacts,* which include the tools, implements, utensils, and instruments of everyday living. Hodder (1994) includes artifacts and written texts that have physically endured over time as "mute material evidence" (p. 398) in the study of culture. In a study of students with physical disabilities, for example, specially designed or modified tools for learning (computers, sports equipment, and so on) could be included as part of the database.

Physical trace material is yet another potential source of information. Physical traces are defined as "any changes in the physical environment due to human actions" (Rathje, 1979, pp. 75–76). Examples of physical evidence being used in research studies are provided by Webb and others (1981, p. 2):

- One investigator wanted to know the relationship between reported and actual beer consumption. He obtained a "front door" measure by asking residents of houses how much beer they consumed each week and a "back door" measure by counting the beer cans in their garbage cans. The back door measure resulted in a considerably higher estimate of beer consumption.
- The degree of fear induced by a ghost-story-telling session can be measured by noting the shrinking diameter of a circle of seated children. . . .
- Library withdrawals were used to demonstrate the effect of the introduction of television into a community. Fiction titles dropped, nonfiction titles were unaffected. . . .
- A child's interest in Christmas was demonstrated by distortions in the size of Santa Claus drawings.
- Racial attitudes in two colleges were compared by noting the degree of clustering of blacks and whites in lecture halls.

Two basic means of studying physical traces are to note their erosion, which is the degree of wear, and to note their accretion, which is the degree of accumulation. The wear and tear on floor tiles in front of a museum exhibit as a sign of public interest is a well-known example of erosion (Webb and others, 1966); the accumulation of beer cans in the preceding list is a good example of accretion.

Because physical traces can usually be measured, they are most often suited for obtaining information on the incidence and frequency of behavior. They are also a good check on information obtained from interviews or surveys. In case study research, most physical trace measures are used to supplement data gathered through interviews and observations. A researcher might, for example, compare the wear and tear on computer terminals in a school program that purports to include computer literacy in its basic curriculum. Other advantages of using trace measures are noted by Rathje (1979, pp. 78–79):

- Trace measures record the results of actual behavior, not reported or experimental approximations.
- Trace measures are usually *nonreactive* and *unobtrusive*. Since they are applied after behavior has occurred they do not modify the behavior they seek to study.
- Material traces are ubiquitous and readily available for study.
- Because material traces are applied to inanimate objects, they usually require minimal cooperation and inconvenience from human subjects.
- Because the number of measures of traces depends upon the recorder's interest rather than informant patience, a variety of interrelated behaviors can often be studied at once.
- Because of the minimal inconvenience and expense to informants, trace measures can be used over long time periods as longitudinal monitoring devices.

Researcher-Generated Documents

Most commonly, when documents are included in a study, what is being referred to are public records, personal documents, and physical material *already present* in the research setting. Because they

have not been produced for the research purpose, they often contain much that is irrelevant to the study; by the same token, they can contain clues, even startling insights, into the phenomenon under study. Most researchers find them well worth the effort to locate and examine.

Researcher-generated documents are documents prepared by the researcher or for the researcher by participants after the study has begun. The specific purpose for generating documents is to learn more about the situation, person, or event being investigated. The researcher might request that someone keep a diary or log of activities during the course of the investigation. Or a life history of an individual or historical account of a program might be solicited to illuminate the present situation.

A researcher's photographs are another example of this type of document. Such photographs, often taken in conjunction with participant observation, provide a "means of remembering and studying detail that might be overlooked if a photographic image were not available for reflection" (Bogdan and Biklen, 1992, p. 143). Preskill (1995) reports using photographs to document aspects of the organizational culture of a magnet high school; the photos were also used as resources in interviewing teachers and students—what Preskill calls "reading photographs" and "photographic interviewing" (p. 189). Photographs can also be taken by the participants. In a study of differing perceptions of white and African American Greek members of their university environment, researchers provided disposable cameras to participants to take photos exemplifying what their university experience meant (Perka, Matherly, Fishman, and Ridge, 1992). These photos and interviews asking participants to interpret the photos provided the data for analysis. (See Becker, 1986a, for more on photography in qualitative research).

Quantitative data produced by the investigator also fall into this category of documents. Projective tests, attitudinal measures, content examinations, statistical data from surveys on any number of topics—all can be treated as documents in support of a qualitative investigation. In a case study of a county health workers' training program, for example, data were collected from written questionnaires as well as through observation and interviews. Results of the survey became supporting documentary material

for the observation and interview-based findings of the study (Dominick and Cervero, 1987).

In summary, then, documents include a broad range of materials available to the researcher who is creative in seeking them out. Literally millions of public and private documents, as well as physical traces of human behavior, can be used as primary or secondary sources of data. Further, documents can be generated by the researcher once the study has begun.

Using Documents in Qualitative Research

Using documentary material as data is not much different from using interviews or observations. Glaser and Strauss (1967) compare fieldwork with library research. "When someone stands in the library stacks, he is, metaphorically, surrounded by voices begging to be heard. Every book, every magazine article, represents at least one person who is equivalent to the anthropologist's informant or the sociologist's interviewee. In those publications, people converse, announce positions, argue with a range of eloquence, and describe events or scenes in ways entirely comparable to what is seen and heard during fieldwork" (p. 163).

Whether in fieldwork or library work, the data collection is guided by questions, educated hunches, and emerging findings. Although the search is systematic, both settings also allow for the accidental uncovering of valuable data. Tracking down leads, being open to new insights, and being sensitive to the data are the same whether the researcher is interviewing, observing, or analyzing documents. Since the investigator is the primary instrument for gathering data, he or she relies on skills and intuition to find and interpret data from documents.

Finding relevant materials is the first step in the process. As I mentioned, this is generally a systematic procedure that evolves from the topic of inquiry itself. A case study of a back-to-industry program for postsecondary faculty logically led the researcher to memos, background papers, advertising material, application forms, and final reports on the project (Kline, 1981). A qualitative study of classroom instruction would lead to documents in the form of instructors' lesson plans, student assignments, objects in the classroom, official grade reports and school records, teacher

evaluations, and so on. Besides the setting itself, the logical places to look are libraries, historical societies, archives, and institutional files. Others have located personal documents like letters and diaries by placing advertisements in newspapers and newsletters (Taylor and Bogdan, 1984).

Selltiz, Jahoda, Deutsch, and Cook (1959) observe that finding pertinent documents hinges to some extent on the investigator's ability to think creatively about the problem under study. "The use of such data demands *a capacity to ask many different questions related to the research problem.* By definition, the purpose for which available records have been collected is different from the purpose for which the social scientist wishes to use them" (p. 318, emphasis in original). Thus the researcher must keep an open mind when it comes to discovering useful documents. Being open to any possibility can lead to serendipitous discoveries. Recent tobacco company exposés have been buttressed by the discovery of buried memos in which the addictive quality of nicotine is discussed; the famous Watergate tapes were literally stumbled upon during routine questioning of White House staff.

Once documents have been located, their authenticity must be assessed. Since they were not produced for the researcher, the investigator must try to "reconstruct the process by which the data were originally assembled by somebody else" (Riley, 1963). It is important to determine "the conditions under which these data were produced, what specific methodological and technical decisions may have been made, . . . and the consequent impact on the nature of the data now to be taken over" (p. 252). A news release to the general public serves a quite different purpose than an internal memo on the same issue. In evaluating artifacts—that is, objects used or produced by a particular cultural group—LeCompte and Preissle (1993) suggest that the researcher ask such questions as, What is the history of its production and use? How is its use allocated? Is its selection biased? How might it be distorted or falsified?

Determining the authenticity and accuracy of documents is part of the research process. Burgess (1982) writes that documents should not be used in isolation. It is the investigator's responsibility to determine as much as possible about the document, its origins and reasons for being written, its author, and the context in which it was written. Guba and Lincoln (1981), citing Clark (1967),

list the questions a researcher might ask about the authenticity of documents:

- What is the history of the document?
- How did it come into my hands?
- What guarantee is there that it is what it pretends to be?
- Is the document complete, as originally constructed?
- Has it been tampered with or edited?
- If the document is genuine, under what circumstances and for what purposes was it produced?
- Who was/is the author?
- What was he trying to accomplish? For whom was the document intended?
- What were the maker's sources of information? Does the document represent an eyewitness account, a secondhand account, a reconstruction of an event long prior to the writing, an interpretation?
- What was or is the maker's bias?
- To what extent was the writer likely to want to tell the truth?
- Do other documents exist that might shed additional light on the same story, event, project, program, context? If so, are they available, accessible? Who holds them? [pp. 238–239]

An important distinction for historians that qualitative researchers might also attend to is whether documents are primary or secondary sources. Primary sources are those in which the originator of the document is recounting firsthand experience with the phenomenon of interest. The best primary sources are those recorded closest in time and place to the phenomenon by a qualified person. Given this definition, most personal documents and eyewitness accounts of social phenomena could be considered primary resources. Secondary sources are reports of a phenomenon by those who have not directly experienced the phenomenon of interest; these are often compiled at a later date. Interestingly, the same document could be classified as primary or secondary depending upon the purpose of the study. The diary of a loved one caring for someone with terminal cancer, for example, would be a primary source of data for a study on caretaking; it would be con-

sidered a secondary source of data for understanding how patients themselves cope with a terminal disease.

After assessing the authenticity and nature of documents or artifacts, the researcher must adopt some system for coding and cataloging them. If at all possible, written documents should be copied and artifacts photographed or videotaped. By establishing basic descriptive categories early on for coding, the researcher will have easy access to information in the analysis and interpretation stage. In a case study of a career enhancement award program, for example, applications for the award were part of the database (Zeph, 1991). The applications were coded according to the applicant's type of employment, dollar amount of request, sex, geographic location, and nature of the project proposed.

In qualitative case studies, a form of content analysis is used to analyze documents. Essentially, content analysis is a systematic procedure for describing the content of communications. Historians and literary critics have long used content analysis to analyze historical documents and literary works. Modern content analysis has most often been applied to communications media (newspapers, periodicals, television, film) and has had a strong quantitative focus. A major concern has been measuring the frequency and variety of messages and confirming hypotheses. Most research designs using content analysis are sequential in nature—"moving from category construction to sampling, data collection, data analysis and interpretation" (Altheide, 1987, p. 68). Data collection and coding are often carried out by novices using protocols and trained to count units of analysis.

Quantification need not be a component of content analysis, however. The nature of the data can also be assessed. Altheide (1987) describes how qualitative content analysis differs from conventional content analysis. "Ethnographic content analysis is used to document and understand the communication of meaning, as well as to verify theoretical relationships. Its distinctive characteristic is the reflexive and highly interactive nature of the investigator, concepts, data collection and analysis. . . . The investigator is continually central, although protocols may be used in later phases of the research. . . . The aim is to be systematic and analytic, but not rigid" (p. 68).

Limitations and Strengths of Documents

In judging the value of a data source, a researcher can ask whether it contains information or insights relevant to the research question and whether it can be acquired in a reasonably practical yet systematic manner. If these two questions can be answered in the affirmative, there is no reason not to use a particular source of data. Documents or artifacts have been underused in qualitative research, however. Glaser and Strauss (1967) attribute this underuse to the fact that researchers prefer to produce their own data, that the use of documents is too much like historical research, that researchers want "to see the concrete situation and informants in person" (p. 163), and that they distrust their own competency in using documentary materials.

Preferences for other sources of data may reflect a researcher's uncertainty about the potential of documents for yielding knowledge and insight. But the researcher's caution may also reflect some of the limitations inherent in this data source. Several limitations stem from the basic difference between this source and data gleaned from interviews or observations—that most documentary data have not been developed for research purposes. The materials may therefore be incomplete from a research perspective. In contrast to field notes, available materials may not "afford a continuity of unfolding events in the kind of detail that the theorist requires" (Glaser and Strauss, 1967, p. 182). Whether personal accounts or official documents are involved, the source may provide unrepresentative samples. "Often no one on the project keeps very good notes on processes, few memoranda are generated, and, even more often, the only writing that is done is in response to funders' requests for technical reports or other periodic statements about the progress of the program or project. If no documents exist, however, or if the documents are sparse and seem uninformative, this ought to tell the inquirer something about the context" (Guba and Lincoln, 1981, pp. 234–235).

Because documents are not produced for research purposes, the information they offer may not be in a form that is useful (or understandable) to the investigator. Furthermore, such data "may not fit present definitions of the concepts under scrutiny; they may lack correspondence with the conceptual model" (Riley, 1963,

p. 254). This is, of course, more of a problem when documents are used as secondary data sources to verify findings based on other data. If documents are used as part of the process of inductively building categories and theoretical constructs as in qualitative case studies, then their "fit" with preestablished concepts or models is less of a concern.

A third major problem with documentary materials is determining their authenticity and accuracy. Even public records that purport to be objective and accurate contain built-in biases that a researcher may not be aware of. For example, the incidence and frequency of crimes reported in police records may be a function of how certain crimes are defined and a particular department's procedures for reporting them. Personal documents are subject to purposeful or nonpurposeful deception. There is likely to be, for example, an underestimation of income in a personal income tax report versus an overestimation of expenses in a grant proposal. Distortion in personal documents may be unintentional in that the writer is unaware of his or her biases or simply does not remember accurately. Selltiz, Jahoda, Deutsch, and Cook (1959, p. 325) quote Augustine, who noted this problem of authenticity in his famous personal document, *Confessions*. "And when they hear me confessing of myself, how do they know whether I speak the truth?" Concern about authenticity applies to historical documents as well as to anonymous project reports and sources who wish to remain anonymous, such as "Deep Throat" of the 1974 Watergate case (Webb and others, 1981).

Despite these limitations, documents are a good source of data for numerous reasons. To begin with, they often meet Dexter's (1970) criteria for selecting a particular data collection strategy, that is, documents should be used when it appears they will yield "*better* data or *more* data . . . than other tactics" (p. 11). Many documents are easily accessible, free, and contain information that would take an investigator enormous time and effort to gather otherwise.

Furthermore, documents may be the *only* means of studying certain problems. Riley (1963) notes four situations in which documents are crucial to an investigation: (1) historical studies in which events can no longer be observed and informants may not recall or be available for recall; (2) cross-cultural studies in which settings are remote or inaccessible; (3) studies that rely on technical expertise

such as a doctor's report; and (4) studies of intimate personal rela-
tionships that cannot be observed and that people are often reluc-
tant to discuss.

The data found in documents can be used in the same man-
ner as data from interviews or observations. The data can furnish
descriptive information, verify emerging hypotheses, advance new
categories and hypotheses, offer historical understanding, track
change and development, and so on. Glaser and Strauss (1967)
point to the usefulness of documents for theory building—a
process that "begs for comparative analysis. The library offers a *fan-
tastic range* of comparison groups, if only the researcher has the
ingenuity to discover them" (p. 179, emphasis in original).

One of the greatest advantages in using documentary material
is its stability. Unlike interviewing and observation, the presence
of the investigator does not alter what is being studied. Documen-
tary data are "objective" sources of data compared to other forms.
Such data have also been called "unobtrusive." Webb and others'
(1966) classic book on unobtrusive measures in its revised form is
titled *Nonreactive Measures in the Social Sciences* (1981) because, they
write, "we came to realize over the years that the original title was
not the best one since it was the nonreactivity of the measures
rather than their unobtrusiveness that was of major concern" (p.
ix). Nonreactive measures include physical traces, official records,
private documents, and simple and contrived observations.

Finally, documentary data are particularly good sources for
qualitative case studies because they can ground an investigation in
the context of the problem being investigated. Analysis of this data
source "lends contextual richness and helps to ground an inquiry
in the milieu of the writer. This grounding in real-world issues and
day-to-day concerns is ultimately what the naturalistic inquiry is
working toward" (Guba and Lincoln, 1981, p. 234).

Thus, like any other source of data, documents have their limi-
tations and their advantages. Because they are produced for rea-
sons other than research, they may be fragmentary, they may not fit
the conceptual framework of the research, and their authenticity
may be difficult to determine. However, because they exist inde-
pendent of a research agenda, they are nonreactive, that is, unaf-
fected by the research process. They are a product of the context
in which they were produced and therefore grounded in the real

world. Finally, many documents or artifacts cost little or nothing
and are often easy to obtain.

On-Line Data Sources

Anyone who reads a newspaper has seen the term *information super-
highway* applied to the Internet and heard about the explosive
growth it has undergone in the last few years. From its humble
beginnings as a communication tool exclusively for university pro-
fessors and scientists (initially designed to withstand the results of
a war), the Internet has become a standard resource for college
students, businesses, and anyone else who has access to a computer
with a modem.

In addition to providing a number of reference sources—albeit
of uneven quality—the Internet supports interactions among peo-
ple through various forms of computer-mediated communication.
E-mail, listservs, usenet groups, chat rooms, and other interactive
environments allow people who have never met to encounter one
another and even establish relationships conducted primarily
through on-line contacts. These interactions, still ill-defined within
our society, are of obvious interest to qualitative researchers. In
addition to being a focus for study in and of themselves, they pro-
vide multiple sources of data relating to other studies. What fac-
tors must be considered when accessing and analyzing these data
sources?

In this section I will explore some of the issues associated with
the use of on-line data sources. How are these sources similar to
more familiar sources, such as documents, interviews, and obser-
vations? How are they different? What issues and concerns are
raised by the effects of the media on the data-gathering process?
What ethical considerations arise in this new research context?

These are not questions easily answered, nor are they the
exclusive province of qualitative researchers. Articles in computer-
related magazines and the popular press regularly discuss various
effects of the Internet on society at large, ranging from explo-
rations of the "multiple selves" possible on-line, to mentions of
"on-line affairs" between people who have never seen one another
in person, to news about religious groups using the Internet to
build community, evangelize, and make sacred texts available to

the public (Geller, 1996; McCorduck, 1996; Namuth and others, 1996). Even standard news magazines highlight issues related to cyberspace—the ambiguous destination to which the information superhighway leads. Since the changing electronic landscape outpaces the publication of specific maps or guides, this survey will merely outline a general range of concerns for discussion. For any particular area of study, the specific application of these considerations will vary.

On-Line Versus Off-Line Data

In qualitative research, the three basic ways to collect data have traditionally been through interviews, observations, and examinations of documents and artifacts (Merriam and Simpson, 1995). Many of the references and data sources available on-line reflect characteristics of these familiar data sources. Web pages, papers available through file transfer protocol, and various forms of "electronic paper" can be considered documents that are simply accessed on-line. Illustrations and programs—even games—available in static form to be downloaded by the user can be treated as artifacts. E-mail can be used to question individuals as in an interview, and researchers can observe the on-line interactions among individuals in a variety of formats.

To some extent then, on-line data collection offers an electronic extension of familiar research techniques, widening the scope of data available to the researcher. Certainly, many of the decisions faced in off-line situations emerge in parallel form in on-line research: whether to join an on-line community as a complete observer, a complete participant, or something in between; how to select a sample group; how to approach potential participants when initiating a study; how to gain trust; and so on.

However, on-line data collection has some important differences due to the nature of the medium through which it is conducted. These differences have a profound influence on the study that must not be ignored or trivialized. For example, individuals who do not have access to computers will be automatically excluded from the study. Is this appropriate for the study, or will demographic differences that correlate with computer access distort the findings?

Though the amount of information increases to an overwhelming degree, not all critical interactions are necessarily available for study. Members of a usenet group may also communicate through private e-mail messages that the researcher never sees (Schrum, 1996). Quantity of information is no guarantee of comprehensiveness.

In addition, each form of computer-mediated communication has a unique effect on the information it transmits. For example, an e-mail interview may have the same verbal content as one conducted in person, but it lacks inflection, body language, and the many other nuances that often communicate more vividly than words. Frequent users of e-mail recognize its limitations; new users are regularly warned that jokes and sarcasm do not travel well online, and they are taught "emoticons" that attempt to replicate the emotional richness common in speech. At the same time that some communication characteristics are curtailed or modified, others are artificially enhanced. The asynchronous nature of e-mail can add reflection time to an on-line interview that would be unavailable in a face-to-face session. Immediate reactions, strong emotional responses, and unguarded expressions are all lost to the researcher unless, after second thought, the participant chooses to make these transient first thoughts available—and is capable of articulating them in writing. These reactions could completely change the interpretation of a response. Conversely, a casual response may have an unexpected and unsettling permanency; e-mail exchanges long forgotten can resurface, sometimes in totally different and even misleading contexts.

Even as they become familiar with the evolving conventions of on-line expression, researchers need to remain alert to the variables of electronic communication. Participants in listservs and usenet groups have an entire terminology to describe certain types of exchanges, such as "flaming" (responding in a hostile manner), "trolling" (deliberately misleading), and "flame-baiting" (provoking flames from others) that occur as implicit meanings find new outlets.

In terms of group interactions, writing skills and computer literacy strongly influence how individuals are perceived on-line. Often someone will seem to have an entirely different character: a funny, charming person can seem caustic and sarcastic when the

smile accompanying his words disappears. Another individual whose writing is mature and thoughtful may prove to have limited social skills when deprived of reflection time and forced to react spontaneously.

This discrepancy between real and on-line personalities occurs even when people are trying to be themselves—or at least an idealized version of themselves (Phillips and Barnes, 1995). It is compounded when individuals purposefully create different on-line personas, which is fairly frequent in some electronic environments. On-line interaction can vary widely, from scholarly communities in which individuals list their real names with their university affiliations and degrees, to fantasy games in which participants make up names and descriptions that reflect little of their off-line characteristics. Where does role-playing shade into deception? As Phillips and Barnes observe, "there is a great deal of opportunity to create fraud through role playing. . . . There is no way to corroborate the image you get. . . . Through text-only exchanges, you have no way to really know who you are talking to" (1995, pp. 39–40). Under these conditions, the assumption that the world is composed of multiple, changing realities—part of the qualitative paradigm— becomes at once a trivially self-evident observation and a magnified complication. Judging individuals by the way they choose to present themselves on-line is a risky business, and verification or triangulation may be far less reliable than in the "real world."

Even on-line documents and artifacts take on new qualities. The Web page cited today may be gone tomorrow or the content changed so radically as to be unrecognizable. Managing data assumes a new dimension when its stability can no longer be taken for granted. Version control, once only of concern to programmers and editors, emerges as a critical issue for anyone using the Internet as a reference or a resource.

This is a new territory, with unfamiliar rules that change as quickly as they are identified. My best advice for researchers is to recognize that the results of their research are strongly influenced by the characteristics of the data revealed, concealed, or altered because of the nature of the medium through which they are presented. Analyzing, describing, and discussing the potential effects of these characteristics will be an important aspect of research conducted from on-line data.

Effects of the Medium on Data Gathering

In addition to the differences between on-line and off-line data, differences caused by the manner in which data are gathered must be considered. In qualitative research, the researcher is the primary instrument for data collection and analysis. This factor is usually perceived as an advantage, because humans are both responsive and adaptive. At the same time, it carries the responsibility of assessing and reporting researcher biases that might have an impact on the study.

When collecting data from the Internet, the researcher is no longer the primary instrument for data collection; a variety of software tools must be used to locate, select, and process information. Like the researcher, these tools have inherent biases that may affect the study, but their biases may be very subtle—and often much more difficult for a researcher to detect and describe. As Norman (1993) observes, "different technologies afford different operations. That is, they make some things easy to do, others difficult or impossible. It should come as no surprise that those things that the affordances make easy are apt to get done, those things that the affordances make difficult are not apt to get done" (p. 106).

Software tools not only shape what is easy or difficult to do, but also shape the user's perception of what is possible, according to Carroll and Kellogg (1989), as cited in Kellogg and Richards (1995). "Software tools will also shape a user's 'policy' for handling information; after all, they determine which actions are possible to take towards a particular piece of information. A user's mental model of goals . . . possible to have vis-à-vis Internet information will arise, at least in part, from the functionality of the software" (pp. 13–14).

These passages raise critical concerns for qualitative researchers accessing data from the Internet: How are their tools shaping the task? Again this is a rapidly evolving area; the researcher's responsibility must be to describe tools and methods, as well as their potential effects on the work.

Ethical Issues

In any qualitative study, ethical issues relating to protection of the participants are of concern. In an on-line environment, these issues overlap the public debate about ownership of intellectual property,

copyright, and free speech. The ability to read, save, copy, archive, and easily edit huge volumes of material written by faceless masses can lead a researcher to forget that these are the words of individuals. Even when the names are changed, some people are easily identified by the details of their messages. The highly public nature of some of the electronic environments in which people exchange ideas can lull researchers into forgetting the right to privacy that these individuals have, just as the seeming anonymity of electronic communication can lull individuals into revealing highly intimate details of their lives to anyone who happens to be reading their messages.

Schrum and Harris (1996), who are among the leaders in exploring the ethical implications of research conducted in the Internet's "virtual communities," frame a number of suggestions to researchers operating in this arena. They suggest that it is the researcher's responsibility to be informed about the basic tenets of ethical qualitative research, to inform participants about the research in a variety of accessible forms, and to respect the participants' ownership of materials they generate (Schrum and Harris, 1996, p. 19).

The term *participants* is commonly used by qualitative researchers to describe the individuals being studied. It is a carefully chosen identifier, with connotations of inclusion and willing cooperation. This single word captures a number of attitudes about research from the qualitative paradigm. It also serves as a litmus test concerning ethics. If this term cannot be accurately used—if *subjects* more appropriately describes the inclusion of unwilling or uninformed individuals under the researcher's scrutiny—then the researcher should honestly reevaluate the methods and procedures of the study.

The growing importance of on-line interaction makes it a natural arena for qualitative research. Three critical areas that the qualitative researcher must consider are the effects of the context on the data, the effects of software functionalities on the data-gathering process, and the effects the medium tends to have on ethical practice. Explicitly considering and describing the impact of these factors is a new responsibility of the qualitative researcher.

Summary

Documents, broadly defined to include public records, personal papers, physical traces, and artifacts, are a third major source of data in qualitative research. Although some documents might be prepared at the investigator's request (such as a respondent keeping a diary or writing a life history), most are produced independently of the research study. They are thus nonreactive and grounded in the context under study. Because they are produced for reasons other than the study at hand, some ingenuity is needed in locating documents that bear on the problem and then in analyzing their content. Congruence between documents and the research problem depends on the researcher's flexibility in construing the problem and the related questions. Such a stance is particularly fitting in qualitative studies, which, by their very nature, are emergent in design and inductive in analysis. Documents of all types can help the researcher uncover meaning, develop understanding, and discover insights relevant to the research problem.

Data gathering on-line is an emerging area of keen interest for qualitative researchers. However, a number of issues must be considered when using data from an on-line interaction; I reviewed some of these issues in this chapter.