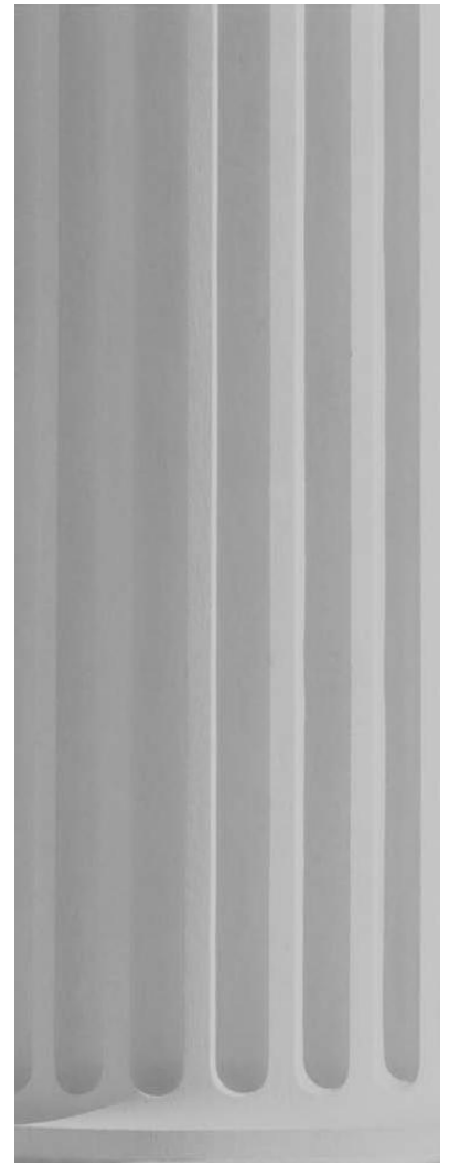# Qualitative and Unobtrusive Measures

KEY TERMS

case study
coding
content analysis
data audit
direct observation
exception dictionary
external validity
field research
hypothesis
indirect measure
mixed methods research
participant observation

qualitative measures
qualitative data
quantitative
reliability
sample
sampling
secondary analysis
true score theory
unitizing
unobtrusive measures
unstructured interviewing
validity

field research
A research method in which the researcher goes into the field to observe the phenomenon in its natural state.

quantitative
The numerical representation of some object. A quantitative variable is any variable that is measured using numbers.

qualitative data
Data in which the variables are not in a numerical form, but are in the form of text, photographs, sound bytes, and so on.

unobtrusive measures
Methods used to collect data without interfering in the lives of the respondents.

content analysis
The analysis of text documents. The analysis can be quantitative, qualitative, or both. Typically, the major purpose of content analysis is to identify patterns in text.

secondary analysis
Analysis that makes use of already existing data sources.

qualitative measures
Data not recorded in numerical form.

This chapter presents two broad areas of measurement—qualitative measurement and unobtrusive measurement. Each of them is distinct from the traditional survey methods described in Chapter 4 and from the scaling and index measures described in Chapter 5.

Qualitative measurement comes from a long tradition of **field research**, originally in anthropology and then subsequently in psychology, sociology, and the other social sciences. This tradition is extremely complex and diverse, and there is probably as much variation and dispute within the tradition as there is in more **quantitative** traditions. Even the simple notion that qualitative means nonquantitative has begun to break down as we recognize the intimate interconnectedness between the two. This chapter introduces the qualitative tradition, the idea of **qualitative data** and the different approaches to collecting it, the different types of qualitative methods, and the standards for judging the validity of qualitative measurement.

Unobtrusive measures are ones that are collected without interfering in the lives of the respondents (Webb et al., 1981). They also represent a broad tradition of measurement in social research. They range from traditional **content analysis** of existing text documents and **secondary analysis** of existing data to some of the cleverest and most indirect methods of measurement you'll see.

6-1 Qualitative Measures

Qualitative research is a vast and complex area of methodology that can easily take up whole textbooks on its own. The purpose of this section is to introduce you to the idea of qualitative research (and how it is related to quantitative research) and show you the major types of qualitative research data, approaches, and methods.

So, what is qualitative research, and what are **qualitative measures**? Qualitative measures are any measures where the data is not recorded in numerical form. (I know, it's a pain to define something by telling you what it is not, but this really is the most accurate way to look at the breadth of qualitative measures.) Qualitative measures include brief written responses on surveys, interviews, anthropological field research, video and audio data recording, and many other approaches, all of which are characterized by a nonnumerical format. Qualitative research is any research that relies primarily or exclusively on qualitative measures.

6-1a When to Use Qualitative Research

Qualitative research is typically the approach of choice in circumstances that have one or more of the following four characteristics:

- For generating new theories or hypotheses
- For achieving a deep understanding of the issues
- For developing detailed stories to describe a phenomenon
- For mixed methods research

These are addressed in the sections that follow.

Generating New Theories or Hypotheses One of the major reasons for doing qualitative research is to understand a phenomenon well enough to be able to form some initial theories, hypotheses, or hunches about how it works. Too often in applied social research (especially in economics and psychology), graduate students jump from doing a quick literature review of a topic of interest to writing a research proposal complete with theories and *hypotheses* based on their own thinking. What they miss is the direct experience of the phenomenon. Before mounting a study, all students should probably be required to spend some time living with the phenomenon they are studying. If they do, they are likely to approach the existing literature on the topic with a fresh perspective born of their direct experience, as well as formulate their own ideas about what causes what to happen. This is where the more interesting and valuable new theories and hypotheses originate, and good qualitative research can play a major role in such theory development.

Achieving Deeper Understanding of the Phenomenon Qualitative research enables us to get at the rich complexity of the phenomenon, to deepen our understanding of how things work. Although quantitative research can describe a phenomenon generally, across a group of respondents, it is very difficult to learn from a quantitative study how the phenomenon is understood and experienced by the respondents, how it interacts with other issues and factors that affect their lives. In addition, in social research, there are many complex and sensitive issues that almost defy simple quantitative summarization. For example, if you are interested in how people view topics like religion, human sexuality, the death penalty, gun control, and so on, my guess is that you would be hard pressed to develop a quantitative methodology that would do anything more than summarize a few key positions on these issues. Although this does have its place (and it's done all the time), if you really want to try to achieve a deep understanding of how people think about these topics, some type of in-depth interviewing or observation is almost certainly required.

Developing Detailed Stories to Describe a Phenomenon Qualitative research excels at generating detailed information to tell stories. We can see how that is important when we look at how social research is used in policy development and decision making. There's an informal saying among social research that goes something like "one good personal story trumps pages of quantitative results." In legislative hearings and organizational boardrooms, the well-researched anecdote is often what compels decision makers. I'm not suggesting that all we need to do in social research is produce stories. There is a persuasiveness to a pattern of evidence based on quantitative assessment. But if that is all we present, the numbers in our case may not translate well for decision makers because impersonal numbers may not connect to their experience. Illustrating the implications of quantitative data through well-researched qualitative anecdotes and stories is essential to effective use of social research.

Qualitative research, and the stories it can generate, enables you to describe the phenomenon of interest with great richness, often in the original language of the research participants. Because of its complexity, some of the best qualitative research is published in book form, often in a style that almost approaches a narrative story. One of my favorite writers (and, I daresay, one of the finest qualitative researchers) is Studs Terkel. He has written intriguing accounts of the Great Depression (*Hard Times*), World War II (*The Good War*), and socioeconomic divisions in America (*The Great Divide*), among others. In each book, he follows a similar qualitative methodology, identifying informants who directly experienced the phenomenon in question, interviewing them at length, and then editing the interviews so that the collection tells a rich and multilayered story that addresses the question of interest in a way that no one story alone would convey.

mixed methods research
Any research that uses multiple research methods to take advantage of the unique advantages that each method offers. For instance, a study that combines case study interviews with an experimental design can be considered mixed methods.

Mixed Methods Research One of the most important areas in applied social research these days is called **mixed methods research**. In mixed methods research, we simultaneously conduct both qualitative and quantitative research to achieve the advantages of each and mitigate their weaknesses. There are several different ways to accomplish the mixing of methods. These tend to differ in how and at what stage of the research you bring the quantitative and qualitative traditions together. For instance, you can conduct qualitative and quantitative substudies as though they are independent of each other on separate parallel tracks where you bring together the results of each at the end in a synthesis or summary. Or you can mix quantitative and qualitative data collection methods throughout, analyzing the results together and examining the similarities and contrasts. Or you can integrate the qualitative and quantitative approaches into a new synthetic method, such as when we combine qualitative brainstorming and quantitative rating approaches into a single method. Or you can integrate the paradigmatic perspectives of qualitative and quantitative traditions at all stages of a research project, repeatedly and dynamically using each to question and improve the results of the other.

Quantitative research excels at summarizing large amounts of data and reaching generalizations based on statistical estimations. Qualitative research excels at telling the story from the participant's viewpoint, providing the rich, descriptive detail that sets quantitative results into their human context. We are only beginning to learn about how we can best integrate these great traditions of qualitative and quantitative research, and many of today's social research students will spend much of their careers exploring this idea.

6-1b Qualitative and Quantitative Data

It may seem odd that I would argue that there is little difference between qualitative and quantitative data. After all, qualitative data typically consists of words, whereas quantitative data consists of numbers. Aren't these fundamentally different? I don't think so, for the following reasons:

- All qualitative data can be coded quantitatively.
- All quantitative data is based on qualitative judgment.

I'll consider each of these reasons in turn.

All Qualitative Data Can Be Coded Quantitatively What I mean here is simple. Anything that is qualitative can be assigned meaningful numerical values. These values can then be manipulated numerically or quantitatively to help you achieve greater insight into the meaning of the data so you can examine specific hypotheses. Consider an example. Many surveys have one or more short, open-ended questions that ask the respondent to supply text responses. The most familiar instance is probably the sentence that is often tacked onto a short survey: "Please add any additional comments." The immediate responses are text-based and qualitative, but you can always (and usually will) perform some type of simple classification of the text responses. You might sort the responses into simple categories, for example. Often, you'll give each category a short label that represents the theme in the response. What you don't often recognize is that even the simple act of categorizing can be viewed as a quantitative one. For instance, let's say that you develop five themes that the respondents express in their open-ended responses. Assume that you have ten respondents. You could easily set up a simple **coding** table like the one in Table 6–1 to represent the coding of the ten responses into the five themes.

coding
The process of categorizing qualitative data.

This is a simple qualitative thematic coding analysis. But you can represent exactly the same information quantitatively as in Table 6–2.

Notice that this is exactly the same data. The first table (Table 6–1) would probably be called a qualitative coding, while the second (Table 6–2) is clearly quantitative. The quantitative coding gives you additional useful information and makes it

TABLE 6–1	Coding of Qualitative Data into Five Themes for Ten Respondents				
Person	Theme 1	Theme 2	Theme 3	Theme 4	Theme 5
1	✓	✓		✓	
2	✓		✓		
3	✓	✓		✓	
4		✓		✓	
5		✓		✓	✓
6	✓	✓			✓
7			✓	✓	✓
8		✓		✓	
9			✓		✓
10				✓	✓

TABLE 6–2	Quantitative Coding of the Data in Table 6–1					
Person	Theme 1	Theme 2	Theme 3	Theme 4	Theme 5	Totals
1	1	1	0	1	0	3
2	1	0	1	0	0	2
3	1	1	0	1	0	3
4	0	1	0	1	0	2
5	0	1	0	1	1	3
6	1	1	0	0	1	3
7	0	0	1	1	1	3
8	0	1	0	1	0	2
9	0	0	1	0	1	2
10	0	0	0	1	1	2

possible to do analyses that you couldn't do with the qualitative coding. For instance, simply by adding down the columns in Table 6–2, you can say that Theme 4 was the most frequently mentioned, and by adding across the rows, you can say that all respondents touched on two or three of the five themes.

The point is that the line between qualitative and quantitative is less distinct than we sometimes imagine. All qualitative data can be quantitatively coded in an almost infinite variety of ways. This doesn't detract from the qualitative information. You can still do any judgmental syntheses or analyses you want, but recognizing the similarities between qualitative and quantitative information opens up new possibilities for interpretation that might otherwise go unutilized. Now to the other side of the coin. . . .
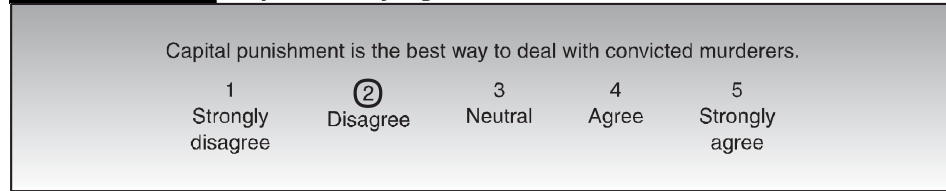
All Quantitative Data Is Based on Qualitative Judgment Numbers in and of themselves can't be interpreted without understanding the assumptions that underlie them. Take, for example, a simple 1 to 5 rating variable, shown in Figure 6–1.

Here, the respondent answered 2 = Disagree. What does this mean? How do you interpret the value 2 here? You can't really understand this quantitative value unless you dig into some of the judgments and assumptions that underlie it:

- Did the respondent understand the term *capital punishment?*
- Did the respondent understand that 2 means that he or she is disagreeing with the statement?

FIGURE 6–1	A rating illustrates that quantitative data is based on qualitative judgments

Capital punishment is the best way to deal with convicted murderers.

1	②	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree

- Does the respondent have any idea about alternatives to capital punishment (otherwise, how can he or she judge what's best)?
- Did the respondent read carefully enough to determine that the statement was limited only to convicted murderers (for instance, rapists were not included)?
- Does the respondent care, or was he or she just circling anything arbitrarily?
- How was this question presented in the context of the survey (for example, did the questions immediately before this one bias the response in any way)?
- Was the respondent mentally alert (especially if this is late in a long survey or the respondent had other things going on earlier in the day)?
- What was the setting for the survey (lighting, noise, and other distractions)?
- Was the survey anonymous? Was it confidential?
- In the respondent's mind, is the difference between a 1 and a 2 the same as between a 2 and a 3 (meaning, is this an interval scale)?

I could go on and on, but my point should be clear. All numerical information involves numerous judgments about what the number means. Quantitative and qualitative data are, at some level, virtually inseparable. Neither exists in a vacuum; neither can be considered totally apart from the other. To ask which is better or more valid or has greater verisimilitude or whatever ignores the intimate connection between them. To do the highest quality research, you need to incorporate both the qualitative and quantitative approaches.

6-1c Qualitative Data

Qualitative data is extremely varied in nature. It includes virtually any information that can be captured that is not numerical in nature (Miles & Huberman, 1994). Here are some of the major categories or types of qualitative data:

- *In-depth interviews.* These include both individual interviews (one-on-one) as well as group interviews (including focus groups). The data can be recorded in numerous ways, including stenography, audio recording, video recording, and written notes. In-depth interviews differ from **direct observation** primarily in the nature of the interaction. In interviews, it is assumed that there is a questioner and one or more interviewees. The purpose of the interview is to probe the ideas of the interviewees about the phenomenon of interest.

- *Direct observation.* I use the term *direct observation* broadly here. It differs from interviewing in that the observer does not actively query the respondent. It can include everything from field research, where one lives in another context or culture for a period of time, to photographs that illustrate some aspect of the phenomenon. The data can be recorded in many of the same ways as interviews (stenography, audio, and video) and through pictures (photos or drawings). (For example, courtroom drawings of witnesses are a form of direct observation.)

- *Written documents.* Usually, this refers to existing documents (as opposed to transcripts of interviews conducted for the research). It can include newspapers, magazines, books, websites, memos, transcripts of conversations, annual reports, and so on. Usually, written documents are analyzed with some form of content analysis (see Section 6-2b, Content Analysis).

direct observation
The process of observing a phenomenon to gather information about it. This process is distinguished from participant observation in that a direct observer does not typically try to become a participant in the context and does strive to be as unobtrusive as possible so as not to bias the observations.

6-1d Qualitative Measures and Observations

A variety of methods are common in qualitative measurement. In fact, the methods are limited primarily by the imagination of the researcher. Here, I discuss a few of the more common methods.

Participant Observation One of the most common methods for qualitative data collection—**participant observation**—is also one of the most demanding. It requires that the researcher become a participant in the culture or context being observed. The literature on participant observation discusses how to enter the context, the role of the researcher as a participant, the collection and storage of field notes, and the analysis of field data. Participant observation often requires months or years of intensive work because the researcher needs to become accepted as a natural part of the culture to ensure that the observations are of the natural phenomenon.

participant observation
A method of qualitative observation in which the researcher becomes a participant in the culture or context being observed.

Direct Observation Direct observation is distinguished from participant observation in a number of ways. First, a direct observer doesn't typically try to become a participant in the context. However, the direct observer does strive to be as unobtrusive as possible so as not to bias the observations. Second, direct observation suggests a more detached perspective. The researcher is watching, rather than both watching and taking part. Consequently, technology can be a useful part of direct observation. For instance, you can videotape the phenomenon or observe from behind one-way mirrors. Third, direct observation tends to be more structured than participant observation. The researcher is observing certain sampled situations or people, rather than trying to become immersed in the entire context. Finally, direct observation tends not to take as long as participant observation. For instance, one might observe mother-child interactions under specific circumstances in a laboratory setting, looking especially for the nonverbal cues being used.

Unstructured Interviewing **Unstructured interviewing** involves direct interaction between the researcher and a respondent or group. It differs from traditional structured interviewing in several important ways. First, although the researcher may have some initial guiding questions or core concepts to ask about, there is no formal structured instrument or protocol. Second, the interviewer is free to move the conversation in any direction of interest that may come up. Consequently, unstructured interviewing is particularly useful for exploring a topic broadly. However, there is a price for this lack of structure. Because each interview tends to be unique with no predetermined set of questions asked of all respondents, it is usually more difficult to analyze unstructured interview data, especially when synthesizing across respondents.

unstructured interviewing
An interviewing method that uses no predetermined interview protocol or survey and in which the interview questions emerge and evolve as the interview proceeds.

Unstructured interviewing may very well be the most common form of data collection of all. You could say it is the method being used whenever anyone asks someone else a question! It is especially useful when conducting site visits or casual focus groups designed to explore a context or situation.

Case Studies A **case study** is an intensive study of a specific individual or specific context. For instance, Freud developed case studies of several individuals as the basis for the theory of psychoanalysis, and Piaget did case studies of children to study developmental phases. Case studies are extensively used in business, law, and policy analysis, with the level of analysis varying from a particular individual to the history of an organization or an event. There is no single way to conduct a case study, and a combination of methods (such as unstructured interviewing and direct observation) is often used. We include case studies in our discussion of qualitative research strategies, but quantitative approaches to studying cases are quite possible and

case study
An intensive study of a specific individual or specific context.

becoming more common with new technology. For example, sometimes researchers provide participants with electronic data collection devices (sometimes called *ambulatory data loggers*) to capture a stream of live events in the natural context. This kind of data can be examined using many kinds of graphic and time series analyses.

Sometimes qualitative case studies can become a form of intervention as well as evaluation. An interesting recent example of this is the Most Significant Change (MSC) technique (Dart & Davies, 2003). The MSC approach generates stories directly from program participants by asking them to describe the most significant change they have experienced or observed in a given period as a result of the program. This form of case study is well suited to understanding change processes as they unfold, but as Dart and Davies pointed out, it can also be used to summarize change at the conclusion of a program and may include both quantitative and qualitative indicators.

Focus Groups Focus groups have become extremely popular in marketing and other kinds of social research because they enable researchers to obtain detailed information about attitudes, opinions, and preferences of selected groups of participants. These methods can be used to generate as many ideas on a topic as possible and to achieve consensus in a group. Sometimes a focus group can be effectively used as a first stage in development of a survey through the identification of potential items relevant to a topic or population. Careful planning of a focus group includes the following considerations:

- What will the specific focus be? It is wise to keep the number of focus questions limited to about five to seven.
- Who will participate? Generally speaking, seven to twelve participants per group will be optimal, but the number of groups you conduct will depend on how much diversity you want to include your sample.
- How will you record the observations? (Audiotaping and videotaping, transcripts, and detailed note taking can be used solely or in combination.)
- How will you analyze the data? There are several approaches to focus group analysis, but perhaps the main consideration is to have a written plan prior to conducting your groups.

It is also very important to carefully think about the ethics of inviting people to discuss topics in a focus group format, especially if the topic is a sensitive one and if your participants are in some way considered vulnerable or have ongoing relationships with one another.

6-1e The Quality of Qualitative Research

Some qualitative researchers reject the framework of validity that is commonly accepted in more quantitative research in the social sciences. They reject the idea that there is a single reality that exists separate from our perceptions. In their view, each of us sees a different reality because we see it from a different perspective and through different experiences. They don't think research can be judged using the criteria of validity. Research is less about getting at the truth than it is about reaching meaningful conclusions, deeper understanding, and useful results. These qualitative researchers argue for different standards of judging the quality of qualitative research.

For instance, Guba and Lincoln (1981) proposed four criteria for judging the soundness of qualitative research and explicitly offered these as an alternative to the four criteria often used in the quantitative tradition (Cook & Campbell, 1979). They thought that their four criteria better reflected the underlying assumptions involved in much qualitative research. Their proposed criteria and the analogous quantitative criteria are listed in Table 6–3.

TABLE 6–3	Criteria for Judging Research Quality from a More Qualitative Perspective

Traditional Criteria for Judging Quantitative Research	Alternative Criteria for Judging Qualitative Research
Internal validity	Credibility
External validity	Transferability
Reliability	Dependability
Objectivity	Confirmability

Credibility The **credibility** criteria involve establishing that the results of qualitative research are credible or believable from the perspective of the participant in the research. Since from this perspective the purpose of qualitative research is to describe or understand the phenomena of interest from the participants' eyes, the participants are the only ones who can legitimately judge the credibility of the results.

credibility
Establishing that the results of qualitative research are believable from the perspective of the participant in the research.

Transferability **Transferability** refers to the degree to which the results of qualitative research can be generalized or transferred to other contexts or settings. From a qualitative perspective, transferability is primarily the responsibility of the one doing the generalizing. The qualitative researcher can enhance transferability by doing a thorough job of describing the research context and the assumptions that were central to the research. The person who wishes to transfer the results to a different context is then responsible for making the judgment of how sensible the transfer is.

transferability
The degree to which the results of qualitative research can be generalized or transferred to other contexts or settings.

Dependability The traditional quantitative view of *reliability* is based on the assumption of replicability or repeatability (see Section 3-2, Reliability). Essentially, it is concerned with whether you would obtain the same results if you could observe the same thing twice. However, you can't actually measure the same thing twice; by definition, if you are measuring twice, you are measuring two different things. This thinking goes back at least to the ancient Greek Democritus, who argued that we can never step into the same river twice because the river is constantly changing. To estimate reliability, quantitative researchers construct various hypothetical notions (for example, true score theory as described in (see Section 3-2a) to try to get around this fact.

The idea of **dependability**, on the other hand, emphasizes the need for the researcher to account for the ever-changing context within which research occurs. The researcher is responsible for describing the changes that occur in the setting and how these changes might affect the conclusions that are reached. Reliability emphasizes the researcher's responsibility to develop measures that, in the absence of any real change, would yield consistent results. Dependability emphasizes the researcher's responsibility to describe the ever-changing research context.

dependability
In qualitative research, the degree to which the research adequately describes the continuously changing context and its effects on conclusions.

Confirmability Qualitative research tends to assume that each researcher brings a unique perspective to the study. **Confirmability** refers to the degree to which others can confirm or corroborate the results. There are a number of strategies for enhancing confirmability. The researcher can actively search for and describe negative instances that contradict prior observations. After the study, a researcher can conduct a **data audit** that examines the data collection and analysis procedures and makes judgments about the potential for bias or distortion.

confirmability
The degree to which others can confirm or corroborate the results in qualitative research.

data audit
A systematic assessment of data and data collection procedures conducted to establish and document the credibility of data collection processes and potential inaccuracies in the data.

6-2 Unobtrusive Measures

Unobtrusive measures are measures that allow the researcher to gather data without becoming involved in respondents' interaction with the measure used (Webb et al., 1981). In most of the methods I've presented to this point (index scores are a notable exception, and can often be considered unobtrusive measures), researchers have some interaction with respondents in the course of conducting studies. For example, direct observation and participant observation require the researcher to be physically present. This can lead the respondents to alter their behavior to look good in the eyes of the researcher or to conform to what they think the researcher would like to see. A questionnaire is an interruption in the natural stream of behavior. Respondents may tire of filling out a survey or become resentful of the questions asked.

Unobtrusive measurement presumably reduces the biases that result from the intrusion of the researcher or measurement instrument. However, unobtrusive measures depend on the context and, in many situations, are simply not available or feasible. For some constructs, there may not be any sensible way to develop unobtrusive measures.

Three approaches to unobtrusive measurement are discussed here: indirect measures, content analysis, and secondary analysis of data.

6-2a Indirect Measures

indirect measure
An unobtrusive measure that occurs naturally in a research context.

An **indirect measure** is an unobtrusive measure that occurs naturally in a research context. The researcher is able to collect data without the respondent being aware of it.

The types of indirect measures that may be available are limited only by the researcher's imagination and inventiveness. For instance, let's say you would like to measure the popularity of various exhibits in a museum. It may be possible to set up some type of mechanical measurement system that is invisible to the museum patrons. In one study, the system was simple. The museum installed new floor tiles in front of each exhibit it wanted a measurement on, and after a period of time, researchers measured the wear-and-tear on the tiles as an indirect measure of patron traffic and interest. You might be able to improve on this approach considerably by using more contemporary electronic instruments. For instance, you might construct an electrical device that senses movement in front of an exhibit or place hidden cameras and code patron interest based on videotaped evidence.

One of my favorite indirect measures occurred in a study of radio station listening preferences. Rather than conducting an obtrusive, costly, and time-consuming survey or interviewing people about favorite radio stations, the researchers went to local auto dealers and garages and checked all cars that were being serviced to see what station the radios were tuned to when the cars were brought in for servicing. In a similar manner, if you want to know magazine preferences, you might observe magazine sales rates, rather than trying to survey readers to ask which magazines they buy. Of course, we need to be careful about how we interpret indirect measures. Just checking radio stations of cars brought in for servicing can be deceptive. We can't automatically conclude that the driver of the car was the one who actually preferred that station (wait till you have kids!), or when it was being listened to, or how often it was tuned in.

These examples illustrate one of the most important points about indirect measures: You have to be careful about ethics when using this type of measurement. In an indirect measure, you are, by definition, collecting information without the respondents' knowledge. In doing so, you may be violating their right to privacy, and you are certainly not using informed consent. Of course, some types of information may be public and therefore do not involve an invasion of privacy, but you should be especially careful to review the ethical implications of the use of indirect measures.

6-2b Content Analysis

Content analysis is the systematic analysis of text (Krippendorff, 2004). The analysis can be quantitative, qualitative, or both. Typically, the major purpose of content analysis is to identify patterns in text. Content analysis is an extremely broad area of research. It includes the following types of analysis:

- *Thematic analysis of text.* The identification of themes or major ideas in a document or set of documents. The documents can be any kind of text, including field notes, newspaper articles, technical papers, or organizational memos.
- *Indexing.* A variety of automated methods for rapidly indexing text documents exists. For instance, Key Words in Context (KWIC) analysis is a computer analysis of text data. A computer program scans the text and indexes all key words. A key word is any term in the text that is not included in an **exception diction-ary**. Typically, an exception dictionary would exclude all nonessential words like *is*, *and*, and *of*. All remaining key words in the text are alphabetized and listed with the text that precedes and follows it, so the researcher can see the word in the context in which it occurred in the text. In an analysis of interview text, for instance, you could easily identify all uses of the term *abuse* and the context in which it was used.
- *Quantitative descriptive analysis.* Here, the purpose is to describe features of the text quantitatively. For instance, you might want to find out which words or phrases were used most frequently in the text. Again, this type of analysis is most often done directly with computer programs.

Content analysis typically includes several important steps or phases. First, when there are many texts to analyze (for example, newspaper stories, organizational reports), the researcher often has to begin by sampling from the population of potential texts to select the ones that will be used. Second, the researcher usually needs to identify and apply the rules that are used to divide each text into segments or "chunks" that will be treated as separate units of analysis in the study, a process referred to as **unitizing**. For instance, you might extract each identifiable assertion from a longer interview transcript. Third, the content analyst constructs and applies one or more codes to each unitized text segment, a process called coding. The development of a coding scheme is based on the themes that you are searching for or uncover as you classify the text. Finally, you analyze the coded data, very often both quantitatively and qualitatively, to determine which themes occur most frequently, in what contexts, and how they might be correlated.

Content analysis has several potential limitations that you should keep in mind. First, you are limited to the types of information available in text form. If you were studying the way a news story is being handled by the news media, you probably would have a ready population of news stories from which you could *sample*. However, if you are interested in studying people's views on capital punishment, you are less likely to find an archive of text documents that would be appropriate. Second, you have to be especially careful with sampling to avoid bias. For instance, a study of current research on methods of treatment for cancer might use the published research literature as the population. This would leave out both the writing on cancer that was not published for one reason or another (publication bias), as well as the most recent work that has not yet been published. Finally, you have to be careful about interpreting results of automated context analyses. A computer program cannot always determine what someone meant by a term or phrase. It is relatively easy in a large analysis to misinterpret a result because you did not take into account the subtleties or context of meaning.

However, content analysis has the advantage of being unobtrusive and, depending on whether automated methods exist, can be a relatively rapid method for analyzing large amounts of text.

exception dictionary
A dictionary that includes all nonessential words like *is*, *and*, and *of*, in a content analysis study.

unitizing
In content analysis, the process of breaking continuous text into separate units that can subsequently be coded.

6-2c Secondary Analysis of Data

Secondary analysis, like content analysis, makes use of already existing data sources. However, *secondary analysis* typically refers to the reanalysis of quantitative data, rather than text.

In our modern world, an unbelievable mass of data is routinely collected by governments, businesses, schools, and other organizations. Much of this information is stored in electronic databases that can be accessed and analyzed. In addition, many research projects store raw data in electronic form in computer archives so that others can also analyze the data. Examples of data available for secondary analysis include:

- Census Bureau data
- Crime records
- Standardized testing data
- Economic data
- Consumer data

Secondary analysis often involves combining information from multiple databases to examine research questions. For example, you might join crime data with census information to assess patterns in criminal behavior by geographic location and group.

Secondary analysis has several advantages. First, it is efficient. It makes use of data that was already collected by someone else. It is the research equivalent of recycling. Second, it often allows you to extend the scope of your study considerably. In many small research projects, it is impossible to consider taking a national sample because of the costs involved. Many archived databases are already national in scope, and by using them, you can leverage a relatively small budget into a much broader study than if you collected the data yourself.

However, secondary analysis is not without difficulties. Frequently, it is no trivial matter to access and link data from large complex databases. Often, you have to make assumptions about which data to combine and which variables are appropriately aggregated into indexes (see Chapter 5). Perhaps more important, when you use data collected by others, you often don't know what problems occurred in the original data collection. Large, well-financed national studies are usually documented thoroughly, but even detailed documentation of procedures is often no substitute for direct experience collecting data.

One of the most important and least utilized purposes of secondary analysis is to replicate prior research findings. In any original data analysis, there is the potential for errors. In addition, data analysts tend to approach the analysis from their own perspective, using the analytic tools with which they are familiar. In most research, the data is analyzed only once by the original research team. It seems an awful waste. Data that might have taken months or years to collect is examined only once in a relatively brief way and from one analyst's perspective. In social research, we generally do a terrible job of documenting and archiving the data from individual studies and making it available in electronic form for others to reanalyze, and we tend to give little professional credit to studies that are reanalyzed. Nevertheless, in the hard sciences, the tradition of replicability of results is a critical one, and we in the applied social sciences could benefit by directing more of our efforts to secondary analysis of existing data.

Summary

This chapter began by comparing qualitative and quantitative data. I made the point that each type of data has its strengths and weaknesses, and that they are often best when used together. Qualitative data can always be quantified, and quantitative data is always based on qualitative assumptions.

Qualitative data can be collected through a variety of methods, including in-depth interviews, direct observation, and written documents. Standards for judging the quality of qualitative data include credibility, transferability, dependability, and confirmability.

Unobtrusive measures are ways of collecting data that don't require researcher interaction with the population of interest. Indirect measures require the researcher to set up conditions so that those being studied are unaware that they are being studied. Content analysis involves the systematic assessment of existing texts and, because it does not require original data collection, is typically considered unobtrusive. Similarly, by definition, the secondary analysis of existing data makes use of information that was previously collected and, as such, does not intrude on respondents.

Login to the Online Edition of your text at www.atomicdog.com to find additional resources located in the Study Guide at the end of each chapter.