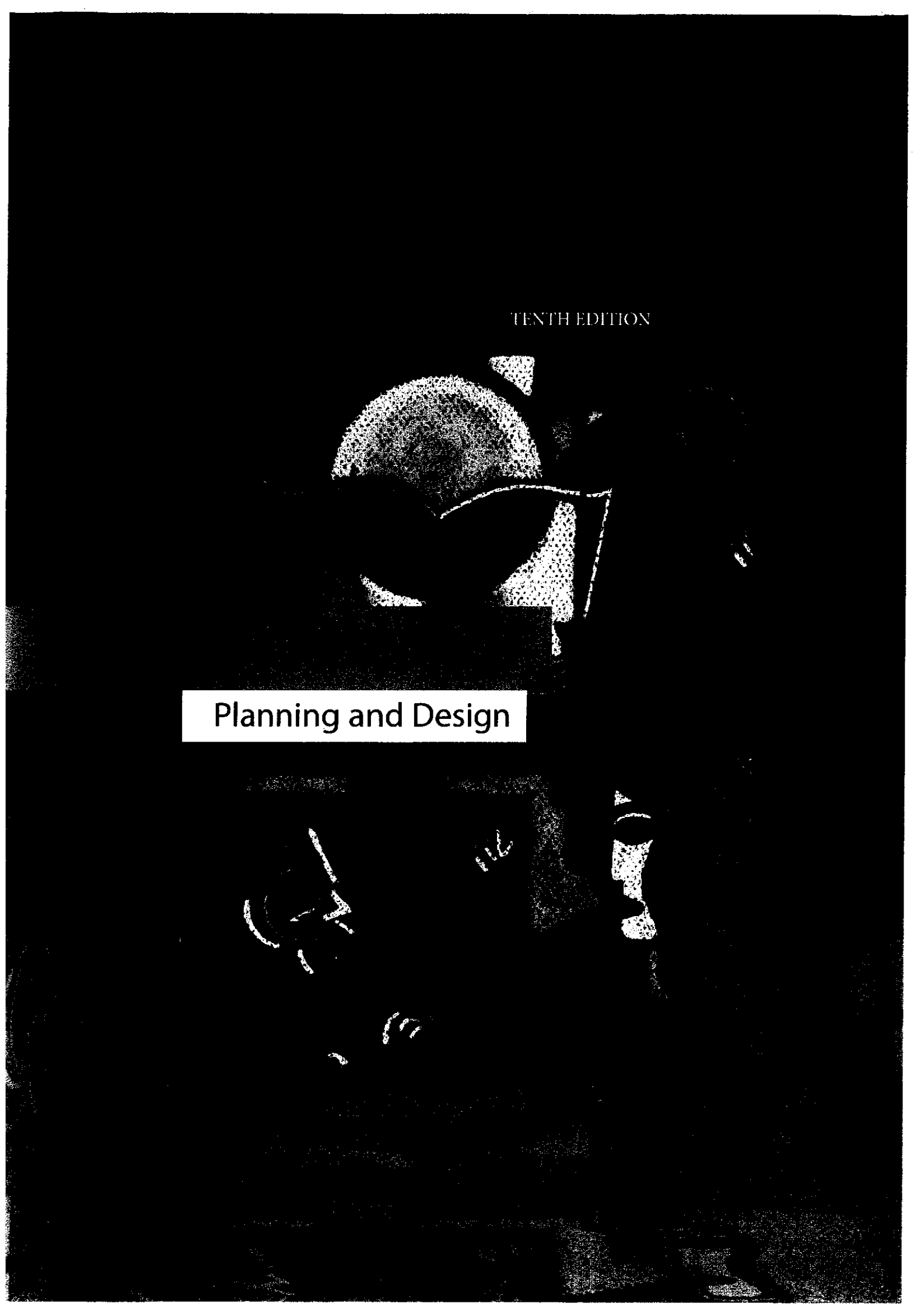
TENTH EDITION



Planning and Design

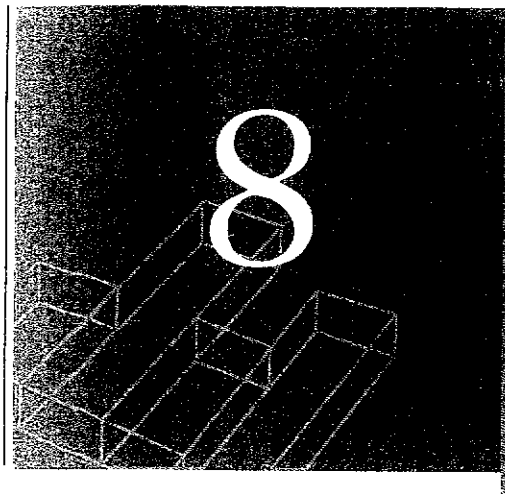
ISBN-13: 978-0-13-269324-0
ISBN-10: 0-13-269324-0

EAN 9 780132 693240 90000 >

www.pearsonhighered.com

PEARSON

ALWAYS LEARNING



8 Descriptive Research

To behold is to look beyond the fact; to observe, to go beyond the observation. Look at the world of people and objects around you, and you will be overwhelmed by what you see. But select from that world a well-chosen few and observe them with insight, and they will tell you a great deal.

In this chapter we discuss types of quantitative study that fall under the broad heading *descriptive quantitative research*. This general category of research designs involves either identifying the characteristics of an observed phenomenon or exploring possible associations among two or more phenomena. In every case, descriptive research examines a situation *as it is*. It does not involve changing or modifying the situation under investigation, nor is it intended to determine cause-and-effect relationships.

As you proceed through the chapter, you will find several strategies—sampling, making observations, interviewing—that you encountered previously in the discussion of qualitative research in Chapter 6. This is old news, you might think. On the contrary, such strategies take on a very different form when we want them to yield quantitative data.

Descriptive Research Designs

In the next few pages, we describe observation studies, correlational research, developmental designs, and survey research, all of which yield quantitative information that can be summarized through statistical analyses. We devote a significant portion of the chapter to survey research, because this approach is used quite frequently in such diverse disciplines as business, government, public health, sociology, and education.

Observation Studies

In the qualitative studies described in Chapter 6, observations are usually recorded in great detail, perhaps with field notes or videotapes that capture the wide variety of ways in which people or other animal species act and interact. From these data, the researcher constructs a complex yet integrated picture of how certain humans or nonhumans spend their time.

In *quantitative* research, however, an **observation study** is quite different. Perhaps most importantly, an observation study does not necessarily involve members of the animal kingdom. Certainly it might involve humans or other animals. But it might, instead, be aimed at studying plant species, nonliving objects (e.g., rock formations, soil samples), or dynamic physical phenomena (e.g., weather patterns, black holes).

In addition, a quantitative observation study tends to have a particular, prespecified focus. When human beings are the topic of study, the focus is typically on a certain aspect of behavior. Furthermore, the behavior is quantified in some way. In some situations, each occurrence of the behavior is *counted* to determine its overall frequency. In other situations, the behavior is *rated* for accuracy, intensity, maturity, or some other dimension. But regardless of approach, the researcher strives to be *as objective as possible* in assessing the behavior

being studied. To maintain such objectivity, he or she is likely to use strategies such as the following:

- Define the behavior being studied in such a precise, concrete manner that the behavior is easily recognized when it occurs.
- Divide the observation period into small segments and then record whether the behavior does or does not occur during each segment. (Each segment might be 30 seconds, 5 minutes, 15 minutes, or whatever other time span is suitable for the behavior being observed.)
- Use a rating scale to evaluate the behavior in terms of specific dimensions (more about rating scales a bit later in the chapter).
- Have two or three people rate the same behavior independently, without knowledge of one another's ratings.
- Train the rater(s) to use specific criteria when counting or evaluating the behavior, and continue training until consistent ratings are obtained for any single occurrence of the behavior.

A study by Kontos (1999) can give you a flavor for what a researcher might do in an observation study. Kontos's research question was this: What roles do preschool teachers adopt during children's free-play periods? (She asked the question within the context of theoretical issues that are irrelevant to our purposes here.) The study took place during free-play sessions in Head Start classrooms, where 40 preschool teachers wore cordless microphones that transmitted what they said (and what people near them said as well) to a remote audiotape recorder. Each teacher was audiotaped for 15 minutes on each of two different days. Following data collection, the tapes were transcribed and broken into 1-minute segments. Each segment was coded in terms of the primary role the teacher assumed during that time, with five possible roles being identified: *interviewer* (talking with children about issues unrelated to a play activity), *stage manager* (helping children get ready to engage in a play activity), *play enhancer/playmate* (joining a play activity in some way), *safety/behavior monitor* (managing children's behavior), or *uninvolved* (not attending to the children's activities in any manner). Two research assistants were trained in using this coding scheme until they were consistent in their judgments at least 90% of the time, indicating a reasonably high *interrater reliability*. They then independently coded each of the 1-minute segments and discussed any segments on which they disagreed, eventually reaching consensus on all segments. (The researcher found, among other things, that teachers' behaviors were to some degree a function of the activities in which the children were engaging. Her conclusions, like her consideration of theoretical issues, go beyond the scope of this book.)

As should be clear from the preceding example, an observation study involves considerable advance planning, meticulous attention to detail, a great deal of time, and, often, the help of one or more research assistants. Furthermore, a pilot study is essential for ironing out any wrinkles in identifying and classifying the behavior(s) or other characteristic(s) under investigation. Embarking on a full-fledged study without first pilot-testing the methodology can result in many hours of wasted time.

Ultimately, an observation study can yield data that portray much of the richness and complexity of human behavior. In some situations, then, it provides a quantitative alternative to such approaches as ethnographies and grounded theory studies.

Correlational Research

A **correlational study** examines the extent to which differences in one characteristic or variable are related to differences in one or more *other* characteristics or variables. A **correlation** exists if, when one variable increases, another variable either increases or decreases in a somewhat predictable fashion.

In correlational studies, researchers gather data about two or more characteristics for a particular group of people or other appropriate units of study. These data are numbers that reflect specific measurements of the characteristics in question. When human beings are the focus of investigation, the data might be test scores, ratings assigned by an expert observer, or frequencies of certain behaviors. Data in animal studies, too, might be frequencies of

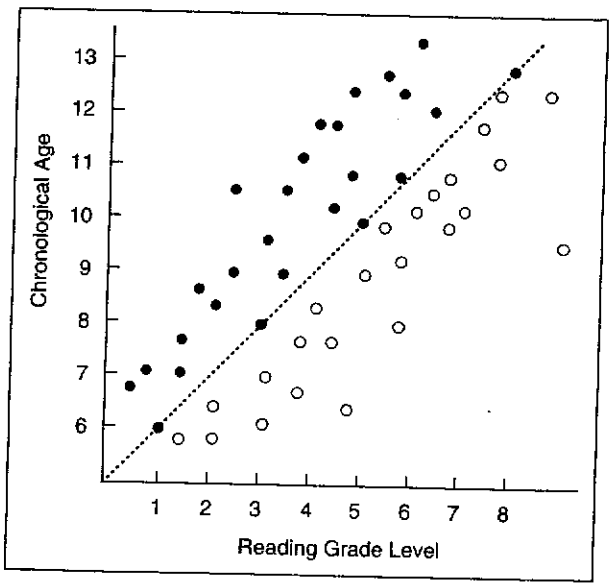
particular behaviors, but alternatively they might be fertility rates, metabolic processes, or measures of health and longevity. Data in studies of plants, inanimate objects, or dynamic physical phenomena might be measures of growth, chemical reactions, density, temperature, or virtually any other characteristic that human measurement instruments can assess with some objectivity. Whatever the nature of the data, at least two different characteristics are measured in order to determine whether and in what way these characteristics are interrelated.

An example of a correlational study may be helpful. As you well know, as children grow older, they become better readers. In other words, there is a *correlation* between age and reading ability. Imagine that a researcher has a sample of 50 children and knows two things about these children: their age and their scores on a reading achievement test (the test scores indicate the approximate "grade level" at which the children are reading). The researcher might plot the data on a **scatter plot** (also known as a *scattergram)* to allow a visual inspection of the relationship between the two variables. Figure 8.1 presents this hypothetical scatter plot. Chronological age is on the vertical axis (the *ordinate)* of the graph, and reading level is on the horizontal axis (the *abscissa)*. Each dot represents a particular child; its placement on the scatter plot indicates both the child's age and his or her reading level.

If age and reading ability were two completely unrelated characteristics, the dots would be scattered all over the graph in a seemingly random manner. When the dots instead form a rough elliptical shape (as the dots in Figure 8.1 do) or perhaps a skinnier sausage shape, then we know that the two characteristics are correlated to some degree. The diagonal line running through the middle of the dots in Figure 8.1—sometimes called the *line of regression*—reflects a hypothetical perfect correlation between age and reading level; if all the dots fell on this line, a child's age would tell us *exactly* what the child's reading level is. In actuality, only four dots—the solid black ones—fall on the line. Some dots lie below the line, showing children whose reading level is, relatively speaking, advanced for their age (these children are designated by hollow black dots). Other dots lie above the line, indicating children who are lagging a bit in reading (these children are designated by solid blue dots).

As we examine the scatter plot, we can say several things about it. First, we can *describe* the homogeneity or heterogeneity of the two variables—the extent to which the children are similar to or different from one another with respect to age and reading level. For instance, if the data were to include only children of ages 6 and 7, we would have greater homogeneity with respect to age than would be the case for a sample of children ages 6 through 13. Second, we can *describe* the degree to which the two variables are intercorrelated, perhaps by computing a statistic known as a *correlation coefficient* (Chapter 11 provides details). But third—and most

FIGURE 8.1
Example of a scatter plot: Correlation between age and reading level



importantly—we can *interpret* these data and give them meaning. For instance, the upward trend of the dots from left to right tells us that as children grow older, their reading level improves. This correlation enables us to use a child's age to estimate—that is, to *predict*—the child's reading level to some degree.

A Caution about Interpreting Correlational Results

In all correlational studies, be alert for faulty logic. When two variables are correlated, researchers sometimes conclude that one of the variables must in some way influence the other. In some instances, such an influence may indeed exist; for example, chronological age—or at least the amount of experience that one's age reflects—almost certainly has a direct bearing on children's mental development, including their reading ability. But ultimately we can never infer a cause-and-effect relationship on the basis of correlation alone. Simply put, *correlation does not, in and of itself, indicate causation.*

Take a logically absurd yet statistically demonstrable example. We could conceivably find a statistical correlation between the number of elephants in Thailand in any given year and the size of the Florida orange crop the same year. The facts may be very clear: As the size of the elephant population increases over time, the Florida orange crop also increases. Yet it is ludicrous to think that because we can show a positive correlation there must therefore be a causal bond at the root of the relationship. There is no connection whatsoever between the elephant population in Thailand and the production of oranges in Florida—at least none that we authors can think of! The correlation is simply a fluke and has no meaning.

In the extreme situation of the elephants–oranges correlation, the faulty logic is readily apparent. Yet we often see similarly faulty reasoning proposed in correlational research reports. For instance, imagine that a researcher finds a correlation between family income level and children's school performance: On average, the lower a child's family income is, the lower the child's academic grade-point-average (GPA) is likely to be. It would be all too easy to draw the conclusion that socioeconomic status directly *affects* (i.e., has a causal influence on) academic achievement. We might think that *because* family paychecks, family living conditions, and so forth are below average, the achievement of the boys and girls of such families is also below par. If a family's economic status could be improved, we conclude, the school achievement of the family's children would also improve.

No, no, no! We *cannot* make an inference about causation on the basis of correlated data alone. It's possible that paycheck size does have an impact on children's grades, but it's equally possible that it does not. Perhaps, instead, an undetermined third variable—maybe parents' education levels or maybe the degree to which family members are victims to society-wide discrimination—influences both the size of the family paycheck and the children's school performance.

If we were to infer that socioeconomic status directly affects academic achievement, not only would we be going far beyond the data we have but we would also have trouble accounting for all of the world's intellectual giants, some of whom have grown up in severe poverty. For example, Robert Burns, a great Scottish poet, had very little formal schooling as a result of his family's poverty. George Washington Carver, a famous American botanist, was born of slaves. The father of Franz Schubert, the Austrian composer, was a peasant, and his mother was a cook; Schubert himself lived in poverty most of his life. The case of Abraham Lincoln, born in a log cabin, walking miles to borrow a book, also discredits the impoverished environment-deprived child notion.

The data may not lie, but the causal conclusions we draw from the data may, at times, be extremely suspect. Nevertheless, a good researcher must not be content to stop at the point of finding a correlational relationship, because *beneath the correlation* lie some potentially quite interesting dynamics. One way to explore these dynamics is through *structural equation modeling (SEM)*, a statistical procedure we describe briefly in Table 11.5 in Chapter 11. Another approach—one that can yield more solid conclusions about cause-and-effect relationships—is to follow up a correlational study with one or more experimental studies (described in Chapter 9) to test various hypotheses about what causes what.

Developmental Designs

Earlier we presented a hypothetical example of how children's ages might correlate with their reading achievement levels. Oftentimes when researchers want to study how a particular characteristic changes as people grow older, they use one of two developmental designs, either a cross-sectional study or a longitudinal study.

In a **cross-sectional study**, people from several different age-groups are sampled and compared. For instance, a developmental psychologist might study the nature of friendships for children at ages 4, 8, 12, and 16. A gerontologist might consider how retired people in their 70s, 80s, and 90s tend to spend their leisure time.

In a **longitudinal study**, a single group of people is followed over the course of several months or years, and data related to the characteristic(s) under investigation are collected at various times.[1] For example, a psycholinguist might examine how children's spoken language changes between 6 months and 5 years of age. Or an educational psychologist might get measures of academic achievement and social adjustment for a group of fourth graders and then, 10 years later, find out which students had completed high school (and what their high school GPAs were) and which ones had not. The educational psychologist might also compute correlations between the measures taken in the fourth grade and the students' high school GPAs; thus, the project would be a correlational study—in this case enabling predictions from Time 1 to Time 2—as well as a longitudinal one.

When longitudinal studies are also correlational studies, they enable researchers to identify potential mediating and moderating variables in correlational relationships. As previously explained in Chapter 2, *mediating variables*—also known as *intervening variables*—may help explain why a characteristic observed at Time 1 is correlated with a characteristic observed at Time 2. Mediating variables are typically measured at some point between Time 1 and Time 2—we might call it Time 1½. In contrast, *moderating variables* influence the nature and strength of a correlational relationship; these might be measured at either Time 1 or Time 1½. A statistical technique mentioned earlier—structural equation modeling (SEM)—can be especially helpful for identifying mediating and moderating variables in a longitudinal study (again we refer you to Table 11.5 in Chapter 11). Yet keep in mind that even with a complex statistical analysis such as SEM, *correlational studies cannot conclusively demonstrate cause-and-effect relationships.*

Obviously, cross-sectional studies are easier and more expedient to conduct than longitudinal studies, because the researcher can collect all the needed data at a single time. In contrast, a researcher who conducts a longitudinal study must collect data over a lengthy period and will almost invariably lose some participants along the way, perhaps because they move to unknown locations or perhaps because they no longer want to participate. An additional disadvantage of a longitudinal design is that when people respond repeatedly to the same measurement instrument, they are likely to improve simply because of their *practice* with the instrument, even if the characteristic being measured hasn't changed at all.

But cross-sectional designs have their disadvantages as well. For one thing, the different age-groups sampled may have been raised under different environmental conditions. For example, imagine that we want to find out whether logical thinking ability improves or declines between the ages of 20 and 70. If we take a cross-sectional approach, we might get samples of 20-year-olds and 70-year-olds and then measure their ability to think logically about various scenarios, perhaps using a standardized multiple-choice test. Now imagine that, in this study, the 20-year-olds obtain higher scores on our logical thinking test than the 70-year-olds. Does this mean that logical thinking ability declines with age? Not necessarily. At least two other possible explanations readily come to mind. The quality of education has changed in many ways over the past few decades, and the younger people have probably had a superior education to that of the older people. And the younger folks have almost certainly had more experience taking multiple-choice tests than the older folks have. Such problems pose threats to the *interval validity*

[1] Some longitudinal studies are conducted over a much shorter time period—perhaps a few minutes or a couple of hours. Such studies, often called *microgenetic studies*, can be useful in studying how children's thinking processes change as a result of short-term, targeted interventions (e.g., Kuhn, 1995).

of this cross-sectional study: We cannot eliminate other possible explanations for the results observed (recall the discussion of internal validity in Chapter 4).

A second disadvantage of a cross-sectional design is that we cannot compute correlations between characteristics at different age levels. Consider, again, the educational psychologist who wants to use students' academic achievement and social adjustment in fourth grade to predict their tendency to complete their high school education. If the educational psychologist were to use a cross-sectional study, there would be different students in each age-group—and only one set of measures for each student—making predictions across time impossible.

To address some of the weaknesses of longitudinal and cross-sectional designs, researchers occasionally combine both approaches in what is known as a **cohort-sequential study**. In particular, a researcher begins with two or more age-groups (this is the cross-sectional piece) and follows each age-group over a period of time (this is the longitudinal piece). As an example, let's return to the issue of how people's logical thinking ability changes over time. Imagine that instead of doing a simple cross-sectional study involving 20-year-olds and 70-year-olds, we begin with a group of 20-year-olds and a group of 65-year-olds. We give them a multiple-choice test designed to assess logical reasoning both at the beginning of the study and then again five years later when they are 25 and 70 years old, respectively. If both groups improve over the five-year time span, we might wonder if practice in taking multiple-choice tests or practice in taking this *particular* test may partly account for the improvement. Alternatively, if the test scores increase for the younger group but decrease for the older group, we might reasonably conclude that logical thinking ability *does* decrease somewhat in the later decades of life.

A cohort-sequential study offers a second advantage, in that, like a longitudinal study, it enables us to calculate correlations between measures taken at two different time periods and therefore to make predictions across time. For instance, we might determine whether people who score highest on the logical thinking test at Time 1 (when they are either 20 or 65 years old) are also those who score highest on the test at Time 2 (when they are either 25 or 70 years old). If we find such a correlation, we can reasonably conclude that logical thinking ability is a relatively stable characteristic—that certain people currently think and will continue to think in a more logical manner than others. We could also add other variables to the study—for instance, determining the extent of participants' postsecondary education or the frequency with which participants engage in activities that require logical reasoning—and determine whether such variables mediate or moderate the long-term stability of logical reasoning ability.

Cross-sectional, longitudinal, and cohort-sequential designs are used in a variety of disciplines, but, as you might guess, they are most commonly seen in developmental research (e.g., studies in child development or gerontology). Should you wish to conduct a developmental study, we urge you to browse in such journals as *Child Development* and *Developmental Psychology* for ideas about specific research strategies.

Survey Research

Some scholars use the term *survey research* to refer to almost *any* form of descriptive, quantitative research. We use a more restricted meaning here: **Survey research** involves acquiring information about one or more groups of people—perhaps about their characteristics, opinions, attitudes, or previous experiences—by asking them questions and tabulating their answers. The ultimate goal is to learn about a large population by surveying a sample of that population; thus, we might call this approach a *descriptive survey* or *normative survey*.

Reduced to its basic elements, a **survey** is quite simple in design: The researcher poses a series of questions to willing participants; summarizes their responses with percentages, frequency counts, or more sophisticated statistical indexes; and then draws inferences about a particular population from the responses of the sample. It is used with more or less sophistication in many areas of human activity—for instance, in a neighborhood petition in support of or against a proposed town ordinance or in a national telephone survey seeking to ascertain people's attitudes about various candidates for political office. This is not to suggest, however, that because of their frequent use, surveys are any less demanding in their design requirements or any easier for the researcher to conduct than other types of research. Quite the contrary, the survey design

makes critical demands on the researcher that, if not carefully respected, can place the entire research effort in jeopardy.

Survey research captures a fleeting moment in time, much as a camera takes a single-frame photograph of an ongoing activity. By drawing conclusions from one transitory collection of data, we may generalize about the state of affairs for a longer time period. But we must never assume that the results can be entered into the Book of Eternal Certainties as ever-abiding Truth. Remember the wisdom of the Greek philosopher Heraclitus: There is nothing permanent but change.

An additional consideration in survey research is that we are relying on *self-report* data: People are telling us what they believe to be true or, perhaps, what they think we want to hear. As noted in the discussion of interviews in Chapter 6, people's memories for events are often distortions of reality—what they think happened isn't always what *did* happen—and people aren't always insightful about their true thoughts and feelings. Furthermore, people's descriptions of their attitudes and opinions are often constructed on the spot—sometimes they haven't really thought about a certain issue until a researcher poses a question about it—and so may be colored by recent events or the current context (Schwarz, 1999). An additional problem is that some participants may intentionally misrepresent the facts—at least, the "facts" as they know them—in order to give the researcher a favorable impression. For example, if we were to ask parents the question "Have you ever abused your children?" the percentage of parents who told us *yes* would be close to zero, and so we would almost certainly underestimate the prevalence of child abuse in our society.

Survey research typically employs a face-to-face interview, a telephone interview, or a written questionnaire. We discuss these techniques briefly here and then offer practical suggestions for conducting them in "Practical Application" sections later on. We describe a fourth approach—using the Internet—in a subsequent "Practical Application" that addresses technology-based methods of data collection.

Face-to-Face and Telephone Interviews

In qualitative research studies, interviews are often quite open-ended, perhaps addressing one or a few central issues but otherwise going in different directions for different participants. In survey research, however, interviews are fairly structured. In a **structured interview**, the researcher asks a standard set of questions and nothing more. In a **semistructured interview**, the research may follow the standard questions with one or more individually tailored questions to get clarification or probe a person's reasoning.

Another difference between qualitative and quantitative studies is the general "feel" of the interview: It tends to be informal and friendly in a qualitative study but more formal and emotionally neutral in a quantitative one. Participants in a qualitative interview may feel as if they're simply engaging in a friendly chat with the researcher, who is typically someone they have come to know and trust. In contrast, participants in survey research are continually aware that, yes, this is an interview and that the temporary relationship they've formed with the researcher will end once the interview is complete. This is not to say, however, that a survey researcher shouldn't strive to establish rapport with participants. Quite the contrary, the researcher is more likely to gain participants' cooperation and encourage them to respond honestly if he or she is likable and friendly and shows a genuine interest in what people have to say.

Face-to-face interviews have the distinct advantage of enabling the researcher to establish rapport with potential participants and therefore gain their cooperation. Thus, such interviews yield the highest **response rates**—the percentages of people agreeing to participate—in survey research. However, the time and expense involved may be prohibitive if the needed interviewees reside in a variety of states, provinces, or countries.

Telephone interviews are less time-consuming and less expensive (they involve only the cost of any long-distance calls), and the researcher has potential access to virtually anyone on the planet who has a landline telephone or cell phone. Although the response rate is not as high as for a face-to-face interview—many people are apt to be busy, annoyed at being bothered, concerned about using costly cell phone minutes, or otherwise not interested in participating—it is considerably higher than for a mailed questionnaire. Unfortunately, the researcher conducting telephone interviews cannot establish the same kind of rapport that is possible in a face-to-face

situation, and the sample will be biased to the extent that people without phones are part of the population about whom the researcher wants to draw inferences.[2]

Whether they are conducted face-to-face or over the telephone, personal interviews allow the researcher to clarify ambiguous answers and, when appropriate, seek follow-up information. Because such interviews take time, however, they may not be practical when large sample sizes are important.

Questionnaires

Paper-and-pencil questionnaires can be sent to a large number of people, including those who live thousands of miles away. Thus, they may save the researcher travel expenses, and postage is typically cheaper than a lengthy long-distance telephone call. The social scientist who collects data with a questionnaire and the physicist who determines the presence of radioactivity with a Geiger counter are at just about the same degree of remoteness from their respective sources of data: Neither sees the source from which the data originate. From the perspective of survey participants, this distance can be an additional advantage: Participants can respond to questions with some assurance that their responses won't come back to haunt them. Thus, they may be more truthful than they would be in a personal interview, especially when addressing sensitive or controversial issues.

Yet questionnaires have their drawbacks as well. Typically, the majority of people who receive questionnaires don't return them—in other words, there may be a low **return rate**—and the people who do return them are not necessarily representative of the originally selected sample. Even when people are willing participants in a questionnaire study, their responses will reflect their reading and writing skills and, perhaps, their misinterpretation of one or more questions. Furthermore, by specifying in advance all of the questions that will be asked—and thereby eliminating other questions that *could* be asked about the issue or phenomenon in question—the researcher is apt to gain only limited, and possibly distorted, information (Dowson & McInerney, 2001).

If questionnaires are to yield useful data, they must be carefully planned, constructed, and distributed. In fact, *any* descriptive study requires careful planning, with close attention to each methodological detail. We now turn to the topic of planning.

Planning for Data Collection in a Descriptive Study

In quantitative research, a descriptive study invariably involves measuring one or more variables in some way. With this point in mind, let's return to a distinction we first made in Chapter 4—the distinction between substantial and insubstantial phenomena. When studying the nature of *substantial phenomena*—phenomena that have physical substance, an obvious basis in the physical world—a researcher can often use measurement instruments that are clearly valid for their purpose. Tape measures, balance scales, oscilloscopes, MRI machines—these instruments are indisputably valid for measuring length, weight, electrical waves, and internal body structures, respectively. Some widely accepted measurement techniques also exist for studying *insubstantial phenomena*—concepts, abilities, and other intangible entities that cannot be pinned down in terms of precise physical qualities. For example, an economist might use Gross Domestic Product statistics as measures of a nation's economic growth, and a psychologist might use the *Stanford-Binet Intelligence Scale* to measure children's general cognitive ability.

Yet many descriptive studies address complex variables—perhaps people's or animals' day-to-day behaviors, or perhaps people's opinions and attitudes about a particular topic—for which no ready-made measurement instruments exist. In their observations, interviews, questionnaires, and so on, descriptive researchers use a variety of strategies to measure complex variables, as you will see in the upcoming Practical Application sections.

[2] Midway between a face-to-face interview and a telephone interview is an interview conducted on the Internet using Skype (www.skype.com) or other video conferencing software. As this tenth edition of the book goes to press, however, people's access to such software is quite limited—and, in our experience, not always dependable—thereby introducing a significant bias into the sample a researcher might get.

PRACTICAL APPLICATION Using Checklists and Rating Scales

Two techniques that facilitate evaluation and quantification of complex phenomena are the checklist and the rating scale. A **checklist** is a list of behaviors or characteristics for which a researcher is looking. The researcher—or in some studies, each participant—simply indicates whether each item on the list is observed, present, or true or, in contrast, is *not* observed, present, or true.

A **rating scale** is more useful when a behavior, attitude, or other phenomenon of interest needs to be evaluated on a continuum of, say, "inadequate" to "excellent," "never" to "always," or "strongly disapprove" to "strongly approve." Rating scales were developed by Rensis Likert in the 1930s to assess people's attitudes; accordingly, they are sometimes called **Likert scales**.[3]

Checklists and rating scales can presumably be used in research related to a wide variety of phenomena, including those involving human beings, nonhuman animals, plants, or inanimate objects (e.g., works of art and literature, geomorphological formations). We illustrate the use of both techniques with a simple example involving human participants. In the late 1970s, park rangers at Rocky Mountain National Park in Colorado were concerned about the heavy summertime traffic traveling up a narrow mountain road to Bear Lake, a popular destination for park visitors. So in the summer of 1978, they provided buses that would shuttle visitors to Bear Lake and back again. This being a radical innovation at the time, the rangers wondered about people's reactions to the buses; if there were strong objections, other solutions to the traffic problem would have to be identified for the following summer.

Park officials asked a sociologist friend of ours to address their research question: How do park visitors feel about the new bus system? The sociologist decided that the best way to approach the problem was to conduct a survey. He and his research assistants waited at the parking lot to which buses returned after their trip to Bear Lake; they randomly selected people who exited the bus and administered the survey. With such a captive audience, the response rate was extremely high: 1,246 of the 1,268 people who were approached agreed to participate in the study, yielding a response rate of 98%.

We present three of the interview questions in Figure 8.2. Based on people's responses, the sociologist concluded that people were solidly in favor of the bus system

FIGURE 8.2
Excerpts from a survey at Rocky Mountain National Park. Item 4 is a *checklist*. Items 5 and 6 are *rating scales*.

From Trahan (1978, Appendix A).

4. Why did you decide to use the bus system?
   - ____ Forced to; Bear Lake was closed to cars
   - ____ Thought it was required
   - ____ Environmental and aesthetic reasons
   - ____ To save time and/or gas
   - ____ To avoid or lessen traffic
   - ____ Easier to park
   - ____ To receive some park interpretation
   - ____ Other (specify): _____

5. In general, what is your opinion of public bus use in national parks as an effort to reduce traffic congestion and park problems and help maintain the environmental quality of the park?

Strongly approve	Approve	Neutral	Disapprove	Strongly disapprove

If "disapprove" or "strongly disapprove," why? _____

6. What is your overall reaction to the present Bear Lake bus system?

Very satisfied	Satisfied	Neutral	Dissatisfied	Very dissatisfied

[3]Although we have often heard *Likert* pronounced as "lie-kert," Likert himself pronounced it "lick-ert."

(Trahan, 1978). As a result, it continues to be in operation today, many years after the survey was conducted.

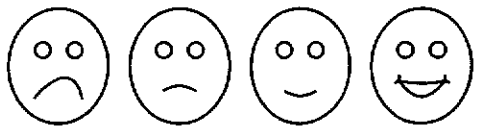
One of us authors was once a member of a dissertation committee for a doctoral student who developed a creative way of presenting a Likert scale to children (Shaklee, 1998). The student was investigating the effects of a particular approach to teaching elementary school science and wanted to determine whether students' beliefs about the nature of school learning—especially learning science—would change as a result of the approach. Both before and after the instructional intervention, she read a series of statements and asked students either to agree or to disagree with them by pointing to one of four faces. The statements and the rating scale that students used to respond to them are presented in Figure 8.3.

Notice that in the rating scale items in the Rocky Mountain National Park survey, park visitors were given the option of responding "Neutral" to each question. In the elementary school study, however, the children always had to answer "yes" or "no." Experts have mixed views about letting respondents remain neutral in interviews and questionnaires. If you use rating scales in your own research, you should consider the implications of letting your respondents straddle the fence by including a "no opinion" or other neutral response, and design your scales accordingly.

Whenever you use checklists or rating scales, you simplify and more easily quantify people's behaviors or attitudes. In the process, however, you may lose valuable information that clarifies or qualifies people's responses. Ultimately you will have to determine whether the trade-off is worth it for the particular research problem you are investigating.

RE 8.3
entary school
out science
g
tary Children's
cal Beliefs and
gs of Science in
of Computer-
deo
with Scientists
by J. M.
unpublished
ertation,
Northern
eeley. Reprinted
ion.

Students responded to each statement by pointing to one of the faces below.



1. No 2. Sort of No 3. Sort of Yes 4. Yes

Students who were unfamiliar with Likert scales practiced the procedure using items A and B; others began with item 1.

A. Are cats green?
B. Is it a nice day?
1. The best thing about science is that most problems have one right answer.
2. If I can't understand something quickly, I keep trying.
3. When I don't understand a new idea, it is best to figure it out on my own.
4. I get confused when books have different information from what I already know.
5. An expert is someone who is born really smart.
6. If scientists try hard enough, they can find the truth to almost everything.
7. Students who do well learn quickly.
8. Getting ahead takes a lot of work.
9. The most important part about being a good student is memorizing the facts.
10. I can believe what I read.
11. Truth never changes.
12. Learning takes a long time.
13. Really smart students don't have to work hard to do well in school.
14. Kids who disagree with teachers are show-offs.
15. Scientists can get to the truth.
16. I try to use information from books and many other places.
17. It is annoying to listen to people who can't make up their minds.
18. Everyone needs to learn how to learn.
19. If I try too hard to understand a problem, I just get confused.
20. Sometimes I just have to accept answers from a teacher even if they don't make sense to me.

PRACTICAL APPLICATION Computerizing Observations

In quantitative research studies, most observations are, of course, quantified in some way, sometimes with simple frequency counts—perhaps using a checklist for guidance—and at other times with rating scales. Either way, you might record your observations on a computer laptop or tablet as you are making them. For example, when using a checklist, you might create a spreadsheet with a small number of columns—one for each item on the checklist—and a row for every entity you will observe. Then, as you conduct your observations, you can enter an "X" or other symbol into the appropriate cell whenever you see an item in the checklist.

For more complex observations, you might create a general *template* document in either a spreadsheet or word processing software and then electronically "save" a separate version of the document for each person, situation, or other entity you are observing. You can either print out these entity-specific documents for handwritten coding during your observations, or, if time and your keyboarding skills allow, you can fill in each document while on-site in the research setting.

You might also look for *peripheral devices*—devices you can hook to your computer to enhance its capabilities—to aid your data collection. For example, Acheson and Gall (2003) have described an approach in which observers use a bar code reader and a special list of bar codes to record different types of behaviors. Each code is associated with a specific type of behavior. When a particular behavior is observed, the bar code reader is swept over the appropriate bar code, and immediately the type and time of the behavior are recorded. In this way, all observations are electronically entered and categorized in a single motion.

PRACTICAL APPLICATION Planning and Conducting Interviews in a Quantitative Study

Interviewing involves much more than just asking questions. The questions for an interview should be carefully planned and precisely worded to yield the kinds of data the researcher needs to answer his or her research question. In the following two sections, we offer guidelines for conducting interviews in a quantitative study and describe how one student successfully planned and conducted the interviews he needed for a research project.

GUIDELINES Conducting Interviews in a Quantitative Study

In Chapter 6 we presented guidelines for conducting interviews in qualitative research. Most of those guidelines are equally applicable in quantitative research, and so we list them again here:

1. Identify questions in advance.
2. Consider how participants' cultural backgrounds might influence their responses.
3. Make sure your interviewees are representative of the group.
4. Find a suitable location.
5. Get written permission.
6. Establish and maintain rapport.
7. Focus on the actual rather than on the abstract or hypothetical.
8. Don't put words in people's mouths.
9. Record responses verbatim.
10. Keep your reactions to yourself.
11. Remember that you are not necessarily getting the facts.

But interviews are typically more structured in quantitative studies than they are in qualitative studies. Therefore, we add several additional guidelines for conducting interviews in quantitative research.

12. *As you write the questions, consider how you can quantify the responses, and modify the questions accordingly.* Remember, you are conducting a *quantitative* study. Thus you will, to some extent,

be coding people's responses as numbers and, quite possibly, conducting statistical analyses on those numbers. You will be able to assign numerical codes to responses more easily if you identify an appropriate coding scheme ahead of time.

13. *Consider asking questions that will elicit qualitative information as well.* You don't necessarily have to quantify *everything*. People's responses to a few open-ended questions may support or provide additional insights into the numerical data you obtain from more structured questions (for example, see the open-ended question in Item 5 of Figure 8.2). By combining quantitative and qualitative data in this manner, you are essentially employing a *mixed-methods design*.

14. *Pilot-test the questions.* When you plan your interview, you will, of course, be trying to develop questions that elicit the kinds of information you are seeking. But despite your best intentions, you may create questions that are ambiguous or misleading or that yield uninterpretable or otherwise useless responses. You can save yourself a great deal of time over the long run if you fine-tune your questions before you begin data collection. You can easily find the weak spots in your questions by asking a few volunteers to answer them in a pilot study.

15. *Restrict each question to a single idea.* Don't try to get too much information in any single question; in doing so, you may get multiple kinds of data—"mixed messages," so to speak—that are difficult to interpret (Gall, Gall, & Borg, 2007).

16. *Save controversial questions for the latter part of the interview.* If you will be touching on sensitive topics (e.g., opinions about gun control, attitudes toward people with diverse sexual orientations), put them near the end of the interview, after you have established rapport and gained the person's trust (Gall et al., 2007).

17. *Seek clarifying information when necessary.* Be alert for responses that are vague or otherwise difficult to interpret. Simple, nonleading questions—for instance, "Can you tell me more about that?"—may yield the additional information you need (Gall et al., 2007, p. 254).

18. *Consider how you might use a computer to streamline the process.* As mentioned in Chapter 6, some computer software programs allow you to record interviews directly onto a laptop computer and then transform these conversations into written text. Alternatively, if interviewees' responses are likely to be short, you might type them directly into a spreadsheet or word processing program.

An Example in International Relations

A student wanted to interview certain United Nations personnel to get their opinions concerning issues related to his study. He planned to go to New York City for a series of interviews and, to conserve both time and expense, wanted to schedule them as tightly as possible. His procedure was organized and logical.

Approximately 6 weeks before going for the interviews, the student wrote the United Nations representatives with whom he wished to confer; he told them when he would be in New York and requested an interview that would last 30 minutes at most. He asked each prospective interviewee for an indication of several time slots when the interview might be scheduled. In his letter, he clearly explained what information he was seeking and why he was seeking it. His reasons were mature and meaningful and were phrased to pique the interest for those he wanted to interview. (*Not* among his reasons was the fact that he was writing a thesis! If you must reveal that you are collecting data for a thesis, use the word *study* instead of *thesis*. Aside from the student and the graduate advisor, theses hold very little glamour in the everyday world. "Studies" are much more acceptable.)

With the letter, the student enclosed a separate sheet containing the questions he intended to ask during the interview, arranged in the order he would ask them. He also suggested that if the interviewee had no objections, he would tape the interview in order to conserve time and lessen the distraction of handwritten notes. He provided a check box on a return postcard for the interviewee to indicate whether he or she had any objection to recording the interview.

After receiving potential interviewees' replies, he set up a master chart of appointments and, by letter, immediately confirmed each interviewee's appointment time and thanked the

interviewee for his or her cooperation. When a time conflict arose, he sought to resolve it by suggesting alternative times that were still open.

Ten days before the scheduled interviews, the student mailed reminders, together with another copy of the interview questions, just in case any interviewees had misplaced the previously sent copies. He also enclosed his full interview schedule so that the interviewees might appreciate the time constraints under which he was working.

The student arrived promptly for each scheduled interview, introduced himself, asked whether the interviewee wanted a copy of the questions he had previously sent, and began with the first question. He tried to guide the interview, always keeping to his agenda of questions and seeking to preserve a relaxed, friendly, yet also professional atmosphere. He wrapped up each interview by thanking the interviewee for the courtesy of giving his or her time. In 3½ days, he interviewed 35 United Nations representatives and had more than four-fifths of his data on tape.

The student transcribed the substance of the interviews and, within 10 days of his visit, sent each interviewee a typed, double-spaced transcript accompanied by a thank-you letter for granting the interview. He asked each individual to read the transcript carefully and, if it was correct, to sign a statement that it was a correct record of the interview. If the person found it inexact or incorrect in any place, he or she could correct the script as desired. In the same mailing, the researcher included a request for permission to use any quotations from the interview in his final report, with the understanding that he would again send the interview content for the interviewee's approval. In his final thesis, the researcher acknowledged his interviewees and noted that they had inspected and approved all of their quoted statements. With the use of such strategies, the researcher and the readers of his report could all be confident that the participants' thoughts and opinions were accurately represented.

In summary, the researcher's use of the following steps led to a highly productive research effort:

1. Set up the interview well in advance.
2. Send the agenda of questions to ask the interviewee.
3. Ask for permission to tape the conference.
4. Confirm the date immediately in writing.
5. Send a reminder, together with another copy of the questions, 10 days before the interview.
6. Be prompt; follow the agenda; offer a copy of the questions in case the original copy has been mislaid.
7. After the interview, submit a transcript of the interview, and get from the interviewee either a written acknowledgment of its accuracy or a corrected copy.
8. After incorporating the material into a semifinal draft of the research report, send that section of the report to the interviewee for final approval and written permission to use the data in the report.

PRACTICAL APPLICATION Constructing and Administering a Questionnaire

Questionnaires seem so simple, yet in our experience they can be tricky to construct and administer. One false step can lead to uninterpretable data or an abysmally low return rate. We have numerous suggestions that can help you make your use of a questionnaire both fruitful and efficient. We have divided our suggestions into three categories: constructing your questionnaire, using technology to facilitate questionnaire administration and data analysis, and maximizing your return rate. In a subsequent Practical Application, we discuss how you might solicit survey participants and collect your data online.

GUIDELINES Constructing Your Questionnaire

Following are 12 guidelines for developing a questionnaire that encourages people to be cooperative and yields responses you can use and interpret. We apologize for the length of the list, but, as we just said, questionnaire construction is a tricky business.

1. *Keep it short.* Your questionnaire should be as brief as possible and solicit only information that is essential to the research effort. You should test every item by two criteria: (a) What do I intend to do with the information I am requesting? and (b) Is it absolutely essential to have this information to solve part of the research problem?

2. *Keep the respondent's task simple and concrete.* Make the instrument as simple to read and respond to as possible. Remember, you are asking for people's *time*, a precious commodity for many people these days. People are more likely to respond to a questionnaire—and to do so quickly—if they perceive it to be quick and easy to complete (McCrea et al., 2008).

Discussion items—those that present open-ended questions and ask people to respond with lengthy answers—are time-consuming and mentally exhausting for both the participants and the researcher. Don't forget that you will have to wrestle with the participants' words to try to determine exactly what their answers mean. The usefulness of responses to discussion items rests entirely on participants' skill to express their thoughts in writing. Those who write in the "Yes/no, and I'll tell you exactly why" style are few and far between. Some respondents may ramble, engaging in discussions that aren't focused or don't answer the questions. Furthermore, after answering 15 to 20 discussion questions, your respondents will think you are demanding a book! Such a major compositional exercise is unfair to those from whom you are requesting a favor.

3. *Provide straightforward, specific instructions.* Communicate exactly how you want people to respond. For instance, don't assume that they are familiar with Likert scales. Some of them may never have seen such scales before.

4. *Use simple, clear, unambiguous language.* Write questions that communicate exactly what you want to know. Avoid terms that your respondents may not understand, such as obscure words or technical jargon. Also avoid words that have imprecise meanings, such as *several* and *usually*.

5. *Give a rationale for any items whose purpose may be unclear.* We cannot say this enough: You are asking people to do you a favor by responding to your questionnaire. Give them a reason to *want* to do the favor. Each question should have a purpose, and in one way or another, you should make that purpose clear.

6. *Check for unwarranted assumptions implicit in your questions.* Consider a very simple question: "How many cigarettes do you smoke each day?" It seems to be a clear and unambiguous question, especially if it is accompanied with certain choices so that all the respondent has to do is to check one of them:

> How many cigarettes do you smoke each day? Check one of the following:
> ___ More than 25 ____ 25–16 ____ 15–11 ____ 10–6 ____ 5–1 ____ none

One underlying assumption here is that a person is likely to be a smoker rather than nonsmoker, which isn't necessarily the case. A second assumption is that a person smokes the same number of cigarettes each day, but for many smokers this assumption isn't viable. At work, when the pressure is on, people may be chain smokers. But on weekends and holidays, they may relax and smoke only one or two cigarettes a day or go without smoking at all. How are the people in this group supposed to answer the question? What box does this type of smoker check?

Had the author of the question considered the assumptions on which the question was predicated, he or she might first have asked questions such as these:

> Do you smoke cigarettes?
> ____ Yes
> ____ No (If you mark "no," skip the next two questions.)
>
> Are your daily smoking habits reasonably consistent; that is, do you smoke about the same number of cigarettes each day?
> ____ Yes
> ____ No (If you mark "no," skip the next question.)

7. *Word your questions in ways that don't give clues about preferred or more desirable responses.* Take another question: "What strategies have you used to try to quit smoking?" By implying that the

respondent has, in fact, tried to quit, it may lead him or her to describe strategies that have never been seriously tried at all.

8. *Determine in advance how you will code the responses.* As you write your questions—perhaps even *before* you write them—develop a plan for recoding participants' responses into numerical data you can statistically analyze. Data processing procedures may also dictate the form a questionnaire should take. If, for example, people's response sheets will be fed into a computer scanner, the questionnaire must be structured differently than if the responses will be tabulated using paper and pencil (we'll say more about computer scanning in the subsequent set of guidelines).

9. *Check for consistency.* When a questionnaire asks questions about a potentially controversial topic, some respondents might give answers that are socially acceptable rather than accurate in order to present a favorable impression. To allow for this possibility, you may wish to ask the same question two or more times—using different words each time—at various points in your questionnaire. For example, take the following two items, appearing in a questionnaire as Items 2 and 30. (Their distance from each other increases the likelihood that a person will answer the second without recalling how he or she answered the first.) Notice how one individual has answered them:

2. Check one of the following choices:
   _X_ In my thinking, I am a liberal.
   ___ In my thinking, I am a conservative.

30. Check one of the following choices:
   ___ I find new ideas stimulating and attractive, and I would find it challenging to be among the first to try them.
   _X_ I subscribe to the position of Alexander Pope:
   "Be not the first by whom the new is tried,
   nor yet the last to lay the old aside."

The two responses are inconsistent. In the first, the respondent claims to be a liberal thinker but later, when given liberal and conservative positions in other forms, indicates a position generally thought to be more conservative than liberal. Such an inconsistency might lead you to question whether the respondent is truly the liberal thinker he or she claims to be.

When developing a questionnaire, researchers sometimes include several items designed to assess essentially the same characteristic. This approach is especially common in studies that involve personality characteristics, motivation, attitudes, and other complex psychological traits. For example, one of us authors once worked with two colleagues to explore factors that might influence the teaching effectiveness of college education majors who were completing their teaching internship year (Middleton, Ormrod, & Abrams, 2007). The research team speculated that one factor potentially affecting teaching effectiveness was willingness to try new teaching techniques and in other ways take reasonable risks in the classroom. The team developed eight items to assess risk taking. Following are four examples, which were interspersed among items designed to assess other characteristics:

	Not at All True		*Somewhat True*		*Very True*
11. I would prefer to teach in a way that is familiar to me rather than trying a teaching strategy that I would have to learn how to do.	1	2	3	4	5
16. I like trying new approaches to teaching, even if I occasionally find they don't work very well.	1	2	3	4	5
39. I would choose to teach something I knew I could do, rather than a topic I haven't taught before.	1	2	3	4	5
51. I sometimes change my plan in the middle of a lesson if I see an opportunity to practice teaching skills I haven't yet mastered.	1	2	3	4	5

Notice how a response of "Very true" to Items 16 and 51 would be indicative of a *high* risk taker, whereas a response of "Very true" to Items 11 and 39 would be indicative of a *low* risk taker. Such counterbalancing of items—some reflecting a high level of a characteristic and others reflecting a low level of the characteristic—can help address some people's general tendency to agree or disagree with a great many statements, including contradictory ones (Nicholls, Orr, Okubo, & Loftus, 2006). A researcher who uses this counterbalancing approach cannot, of course, simply add up a participant's numerical responses for a particular characteristic. For example, for the four risk-taking items just presented, a researcher who wanted high risk takers to have higher scores than low risk takers might give 5 points each for "Very true" responses to the high-risk-taking items (16 and 51) and 5 points each for "Not at all true" responses to the low-risk-taking items (11 and 39). In general, the values of the low-risk-taking items would, during scoring, be opposite to what they are on the questionnaire, with 1s being worth 5 points each, 2s being worth 4 points, 3s being worth 3, 4s being worth 2, and 5s being worth 1. In Appendix A, we describe how to recode participants' responses in precisely this way.

Especially when multiple items are created to assess a single characteristic, a good researcher mathematically determines the degree to which, overall, participants' responses to those items yield consistent results—for instance, the extent to which each person's responses to all "risk-taking" items yield similar results. Essentially, the researcher is determining the *internal consistency reliability* of the set of items. Most statistical software packages can easily compute internal consistency reliability coefficients for you.[4]

Ideally, preliminary data on internal consistency reliability is collected in advance of full-fledged data collection. This point leads us to our next suggestion: Conduct at least one pilot test.

10. *Conduct one or more pilot tests to determine the validity of your questionnaire.* Even experienced researchers conduct test runs of newly designed questionnaires to make sure that the questions are clear and will effectively solicit the desired information. At a minimum, you should give your questionnaire to several friends or colleagues to see whether they have difficulty understanding any items. Have them actually fill out the questionnaire. Better still, ask your pilot-test participants what thoughts run through their minds as they read a question:

> Please read this question out loud. . . . What is this question trying to find out from you? . . . Which answer would you choose as the right answer for you? . . . Can you explain to me why you chose that answer? (Karabenick et al., 2007, p. 143)

Through such strategies you can see the kinds of responses you are likely to get and make sure that in your actual study, the responses you obtain will be of sufficient quality to help you answer your research question.

If your research project will include participants of both genders and various cultural backgrounds, be sure to include a diverse sample in your pilot test(s) as well. Gender and culture *do* play a role in people's responses to certain types of questionnaire items. For instance, some researchers have found a tendency for males to play up their strengths and overrate their abilities, whereas females are apt to ruminate on their weaknesses and *under*rate their abilities (Chipman, 2005; Lundeberg & Mohan, 2009). And people from East Asian cultures are more likely to downplay their abilities than people from Western cultures (Heine, 2007). Keep such differences in mind when asking people to rate themselves on their strengths and weaknesses, and experiment with different wordings that might minimize the effects of gender and culture on participants' responses.

Conducting a pilot study for a questionnaire—and especially asking participants what they are thinking as they read and respond to particular items—is one step toward determining whether a questionnaire has *validity* for its purpose—in other words, whether it truly measures what it is intended to measure. Some academic disciplines (e.g., psychology and related fields) insist that a researcher use more formal and objective strategies to determine a questionnaire's validity, especially when the questionnaire is intended to measure complex psychological traits

[4]Two common reliability coefficients, known by the researchers who originated them, are the Kuder-Richardson Formula 20 (for either–or responses such as *yes* vs. *no* or *true* vs. *false*) and Cronbach's alpha coefficient (for multinumber rating scales such as the 5-point scale for the risk-taking items).

TABLE 8.1 Guide for the construction of a questionnaire

This question is below	Why are you asking the question? How does it relate to the research problem?

(personality, motivation, attitudes, etc.). We refer you to the section "Determining the Validity of a Measurement Instrument" in Chapter 4 for a refresher on strategies you might use.

11. *Scrutinize the almost-final product one more time to make sure it addresses your needs.* Item by item, a questionnaire should be quality tested again and again for precision of expression, objectivity, relevance, and probability of favorable reception and return. Have you concentrated on the recipient of the questionnaire, putting yourself in the place of someone who is asked to invest time on your behalf? If you received such a questionnaire from a stranger, what would your honest reaction be? These questions are important and should be answered impartially.

Above all, you should make sure that *every question is essential for you to address the research problem.* Table 8.1 can help you examine your items with this criterion in mind. Using either paper and pencil or appropriate software (e.g., a spreadsheet or the *table* feature in a word processing program), insert each item in the left-hand column and then, in the right-hand column, explain why you need to include it. If you can't explain how an item relates to your research problem, eliminate it!

12. *Make the questionnaire attractive and professional looking.* Your final instrument should have clean lines, crystal-clear printing (and certainly no typos!), and perhaps two or more colors. It should ultimately communicate that its author is a careful, well-organized professional who takes his or her work seriously and has high regard for the research participants.

GUIDELINES Using Technology to Facilitate Questionnaire Administration and Data Analysis

Throughout most of the 20th century, questionnaire-based surveys were almost exclusively paper-and-pencil in nature. But with continuing technological advances and people's increasing computer literacy in recent years, many survey researchers are now turning to technology to share some of the burden of data collection and analysis. Following are several suggestions for using technology to make the use of a questionnaire more efficient and cost-effective:

1. *Use e-mail to request participation and obtain participants' responses.* If the people you wish to survey are a computer-savvy group, have easily obtainable e-mail addresses, and are regularly online, then an e-mail request to participate can be quite appropriate. Furthermore, you can send the survey either within the body of your e-mail message or as an attachment. Participants can respond in a return e-mail message or electronically fill out and return your attachment.

2. *If you use paper mail delivery rather than e-mail, use a word processing program to personalize your correspondence.* Inquiry letters, thank-you letters, and other correspondence can be personalized by using the *merge* function of most word processing programs. This function allows you to combine the information in your database with the documents you wish to send out. For example, when printing the final version of your cover letter, you can include the person's name immediately after the greeting (e.g., "Dear Carlos" or "Dear Mr. Asay")—a simple touch that is likely to yield a higher return rate than letters addressed to "Potential Respondent" or "To whom it may concern." The computer inserts the names for you; you need only tell it where to find the names in your database. The same process is used on every item of junk mail you receive that addresses you personally or that shows your name on a $10,000,000 check (that you may already have won!) and asks you to order magazines. Advertisers have realized the importance of personalization for quite some time. When individuals are addressed by name, they feel special and are more inclined to invest a little of their time.

3. *Use a computer database to keep track of who has responded and who has not.* Computer databases can be readily adapted for use with research involving questionnaires. Information regarding participants' names and addresses, as well as information regarding which materials have been sent and received, can be incorporated into such a database. A search through the database can quickly identify people who have and have not received your request for participation, those who have or have not yet responded, and those who need a first or second reminder letter or e-mail message.

Combining the database with a word processing program can also be helpful. For example, mailing labels can be quickly produced, thus decreasing the amount of time needed to address various mailings to potential participants. Many word processing and spreadsheet programs come with prepared templates for a variety of mailing labels. By identifying the type of label you have for your printer, you can print any or all of the names from your database directly onto the labels.

4. *Use a scanner to facilitate data tabulation.* When you need a large sample to address your research problem adequately, you should consider in advance how you will tabulate the responses after the questionnaires are returned to you. One widely used strategy is to have a computer scan preformatted answer sheets and automatically sort and organize the results. To use this strategy, your questions must each involve a small set of possible answers; for instance, they might be multiple-choice, have yes-or-no answers, or incorporate a 5-point rating scale. You will want the participants to respond using a pencil or dark-colored ink. Enclosing a small number 2 pencil with the questionnaire you send is common courtesy. Furthermore, anything you can do to make participants' task easier—even something as simple as providing the writing implement—will increase your response rate.

5. *When participants are in the same location that you are, administer the questionnaire directly on a computer.* Electronic questionnaires can be highly effective if participants feel comfortable with computers. When participants enter their responses directly onto a computer, you obviously save a great deal of time. Furthermore, when appropriately programmed to do so, a computer can record how *quickly* people respond—information that may in some situations be relevant to your research question.

GUIDELINES Maximizing Your Return Rate for a Questionnaire

As university professors, we authors have sometimes been asked to distribute questionnaires in our classes that relate to some aspect of the university's student services or to students' preferences for the university calendar. The end-of-semester teacher evaluation forms you often fill out are questionnaires as well. Even though participation in such surveys is voluntary, the response rate when one has such a captive audience is typically quite high, often 100%.

Mailing or e-mailing questionnaires to people one doesn't know is quite another matter. Potential respondents have little or nothing to gain by answering and returning the questionnaire, and so many of them don't return it. As a result, the typical return rate for a mailed

questionnaire is 50% or less, and that for an e-mailed questionnaire is even lower (Rogelberg & Luong, 1998; Sheehan, 2001).

We think of one doctoral student who conducted dissertation research in the area of reading. As part of her study, she sent a questionnaire to reading teachers to inquire about their beliefs and attitudes regarding a certain kind of children's literature. Initially, the student sent out 103 questionnaires; 14 teachers completed and returned them (a return rate of 13%). In a second attempt, she sent out 72 questionnaires to a different group of teachers; 12 responded (a return rate of 15%). In one final effort, she sought volunteers on the Internet by using two lists of teachers' e-mail addresses; 57 teachers indicated that they were willing to fill out her questionnaire, and 20 of them actually did so (a return rate of 35%).

Was the student frustrated? Absolutely! Yet she had made a couple of mistakes that undoubtedly thwarted her efforts from the beginning. First, the questionnaire had 36 questions, 18 of which were open-ended ones requiring lengthy written responses. A quick glance would tell any discerning teacher that the questionnaire would take an entire evening to complete. Second, the questionnaires were sent out in the middle of the school year, when teachers were probably already quite busy planning lessons, grading papers, and so on. Even teachers who truly wanted to help this struggling doctoral student (who was a former teacher herself) may simply not have found the time to do it. Fortunately for the student, the questionnaire was only one small part of her study, and she was able to complete her dissertation successfully with the limited (and almost certainly nonrepresentative) sample of responses she received.

Should you decide that a mailed or e-mailed questionnaire is the most suitable approach for answering your research question, the following guidelines can help you increase your return rate:

1. *Consider the timing.* The student just described mailed her questionnaires in the winter and early spring because she wanted to graduate at the end of the summer. The timing of her mailing was convenient for her; however, it was *not* convenient for the people to whom she sent the questionnaire, and her response rate—and her study!—suffered as a result. Consider the characteristics of the sample you are surveying, and try to anticipate when respondents will be most likely to have time to answer a questionnaire. And as a general rule, stay away from peak holiday and vacation times, such as mid-December through early January.

2. *Make a good first impression.* Put yourself in the place of a potential respondent. Imagine a stranger sending you the questionnaire you propose to send. What is your initial impression as you open the envelope or e-mail message? Is the questionnaire inordinately long and time-consuming? Is it cleanly and neatly written? Does it give the impression of relaxation and uncluttered ease? Are the areas for response adequate and clearly indicated? Is the tone courteous, and are the requests reasonable?

3. *Motivate potential respondents.* Give people a reason to *want* to respond. Occasionally, researchers may actually have the resources to pay people for their time or offer other concrete inducements. But more often than not, you will have to rely on the power of persuasion to gain cooperation. Probably the best mechanism for doing so is the cover letter you include with your questionnaire.

One potentially effective strategy is to send a letter soliciting people's cooperation *before* actually sending them the questionnaire. For example, Figure 8.4 shows an example of a letter that a researcher might use to gain people's cooperation in responding to a questionnaire about the quality of a particular academic program. Several aspects of the letter are important to note:

- The letter begins with the name of the sponsoring institution. Ideally, the letter is written on the institution's official letterhead stationery.
- Rather than saying "Dear Sir or Madam," the letter is personalized. Such personalization is easily accomplished with the *mail merge* function of a word processing program.
- The letter describes the potential value of the study, both for the individual and for alumni in general, hence giving the potential responder a reason to *want* to respond.
- The letter assures the individual that his or her cooperation will not place any unreasonable burden—in particular, that the questionnaire will take a maximum of 15 minutes to complete.
- By filling out and sending a simple enclosed postcard (for example, see Figure 8.5)—a quick and easy first step—the researcher gains the individual's commitment to completing

FIGURE 8.4
A letter of inquiry

A B C University
Address

Date

Dear [person's name],

Your alma mater is appealing to you for help. We are not asking for funds, merely for a few minutes of your time.

We know you are proud of your accomplishments at A B C University, and your degree has almost certainly helped you advance your professional aspirations. You can help us maintain—and ideally also improve—your program's reputation by giving us your honest opinion of its strengths and weaknesses while you were here. We have a questionnaire that, with your permission, we would like to send you. It should take at most only fifteen minutes of your time.

Our program is growing, and with your help it can increase not only in size but also in excellence and national prominence. We are confident that you can help us make it the best that it can possibly be.

Enclosed with this letter is a return postcard on which you can indicate your willingness to respond to our questionnaire. Thank you in advance for your kind assistance. And please don't hesitate to contact me at [telephone number] or [e-mail address] if you have any questions or concerns.

Respectfully yours,

Your Signature

Your Name

a lengthier, more complex task in the near future. The postcard should be addressed and stamped for easy return.

- The letter includes two means of communicating with the researcher in case the individual has any reservations about participating in the study.
- The overall tone of the letter is, from beginning to end, courteous and respectful.

Compare the letter in Figure 8.4 with the brief note in Figure 8.6 that was sent to one of us authors and that, unfortunately, is all too typical of students' first attempts at drafting such a letter. A focus only on the researcher's needs in letters of this sort may be another reason for the poor return of questionnaires in some research projects.

The cover letter is extremely important. It should be carefully and thoughtfully composed and should stress the concerns of the recipient rather than any selfish interests of the sender. Some students forget this and, in doing so, unintentionally reveal their own self-centeredness.

FIGURE 8.5
Questionnaire response card

Dear [your name]

☐ Please send the questionnaire; I will be happy to cooperate.

☐ I am sorry, but I do not wish to answer the questionnaire.

Comments:

Date:_____ _____
Name

FIGURE 8.6
A poorly worded request for cooperation

X Y Z UNIVERSITY
Campus Station

Dear Sir:

I am a graduate student at X Y Z University, and the enclosed questionnaire is sent to you in the hope that you will assist me in obtaining information for my master's thesis.

I should appreciate your early reply since I am attempting to get my degree this June.

Yours truly

John Doe
John Doe

4. *If mailing your questionnaire, include a self-addressed envelope with return postage.* To impose on a person's time and spirit of cooperation and then to expect that person also to supply the envelope and pay the postage is unreasonable.

5. *Offer the results of your study.* In return for the investment of time and the courtesy of replying to your questions, offer to send your respondent a summary of your study's results. You might provide a check box, either at the beginning or at the end of your instrument, where a respondent can indicate the desire to have such a summary, together with a place for name and either mailing or e-mailing address. If anonymity is important, a mailed questionnaire might include a separate postcard on which the respondent can request the summary; this postcard should, of course, have a place for the respondent's name and address, along with the suggestion that the card be mailed separately from the questionnaire. For e-mailed questionnaires, a respondent can simply hit the "reply" button twice, once to return the completed questionnaire and a second time (perhaps a few hours later) to request the study's results.

6. *Be gently persistent.* Experts often suggest that when people don't initially respond to a questionnaire, you can increase your response rate by sending two follow-up reminders, perhaps sending each one out a week or two after the previous mailing (e.g., Neuman, 2011; Rogelberg & Luong, 1998). But if the questionnaire is meant to be anonymous, how do you know who has returned it and who has not?

To address this problem, many researchers put a different code number on each copy they send out and keep a list of which number they have sent to each person in their sample. When a questionnaire is returned, they remove the number and person's name from the list. When it is time to send a follow-up letter, they send it only to the people who are still on the list. Researchers should use the list of names and code numbers *only* for this purpose. At no point should they use it to determine who responded in what way to each question—a practice that violates the right to privacy discussed in Chapter 4.

Let's return to the solicitation letter and postcard in Figures 8.4 and 8.5. We have modeled them after a letter and postcard that an American University faculty member successfully used to get alumni feedback about the university's nursing program. After receiving a card that indicated willingness to cooperate, the faculty member immediately mailed the questionnaire. She kept a log of questionnaires mailed, the names and addresses of people to whom they were mailed, and the date of mailing. If a reply was not received within three weeks' time, she sent a reminder letter. The reminder was written in the same tone as the initial letter. An example of such a reminder letter appears in Figure 8.7.

FIGURE 8.7
A follow-up letter

A B C University
Address

Date

Dear [person's name],

We are all very busy these days, and sometimes we have trouble staying on top of our many commitments. Despite our best intentions, we may sometimes overlook something we have said we would do.

Three weeks ago I sent you a questionnaire asking for your input regarding your program at A B C University. To date I have not yet received your completed questionnaire. Perhaps you have simply mislaid it, or perhaps it has been lost in the mail—any one of several reasons might account for its delay in reaching me.

In any event, I am enclosing another copy of the questionnaire, along with another self-addressed, stamped envelope. I am hoping you can find fifteen minutes somewhere in your busy schedule to complete and return the questionnaire. I would really appreciate your personal insights and suggestions regarding your experiences in our program.

Thank you once again for your assistance and generosity in helping us enhance our program. And remember that if you have any questions, you can easily reach me at [telephone number] or [e-mail address].

Respectfully yours,

Your Signature

Your Name

The faculty member's follow-up letter brought results. She was being firm and persuasive, but with considerable skill and tact. Courtesy, understanding, and respect for others pay large dividends in a situation in which a researcher needs others' cooperation. This is especially true in questionnaire studies.

PRACTICAL APPLICATION Using the Internet to Collect Data for a Descriptive Study

In recent years, some researchers have collected descriptive-study data directly on the Internet. For instance, they may put a questionnaire on a website and ask people who visit the site to respond. One site providing links to a wide variety of online research projects is "Psychological Research on the Net."[5] As we write this tenth edition of the book, the site is hosting descriptive research projects on such diverse topics as television commercial effectiveness, gender roles in marriage, cigarette smoking, eating disorders, attitudes and behaviors in the workplace, the psychological impact of automobile commuting, and coping strategies for dealing with stressful life circumstances. The website is maintained by John Krantz, Professor of Psychology at Hanover College, who checks to be sure that each project has been approved by the appropriate internal review board and incorporates informed consent procedures. There is no fee for using the site (J. Krantz, personal communication, May 2003).

Commercial websites for data collection are available as well. Two widely used websites for online surveys are SurveyMonkey (www.surveymonkey.com) and Zoomerang (zoomerang.com).

[5] You can reach the site by going to the website of the Association for Psychological Science (www.psychologicalscience.org); click on "Psychology links," scroll down to "Other Sites of Interest," and then click on "Online Psychology Experiments." Alternatively, you can go directly to the site, which, as this book goes to press, is located at psych.hanover.edu/research/exponnet.html.

Both websites provide templates that make questionnaire design easy and enable a researcher to present a variety of item types (e.g., multiple-choice items, rating scales). They also include features for communicating with a preselected sample of participants (e.g., through e-mail invitations), as well as features through which the researcher can tabulate, statistically analyze, and download the results. For most research projects these websites charge a modest monthly fee.

Conducting a survey online has several advantages (Kraut et al., 2004). When the desired sample size is quite large, an online questionnaire is far more cost-effective than a mailed questionnaire. Often a questionnaire can be adapted based on a participant's previous responses; for instance, if a person responds *no* to the question "Do you smoke cigarettes?" the questionnaire software will subsequently skip questions related to smoking habits. Furthermore, some evidence indicates that online surveys yield data comparable to those obtained through face-to-face contact (Gosling, Vazire, Srivastava, & John, 2004).

If you choose to collect data on the Internet, keep in mind that your ethical standards must be just as rigorous as they would be if you were collecting data through face-to-face contacts or the postal service. Participants must be informed about and agree to the general nature of a study, perhaps by means of a website page that serves as an informed consent letter and a virtual "click to accept" button with which participants can indicate consent (Kraut et al., 2004). Also, participants' responses must remain as confidential as they would in any study. The *protection from harm* ethical standard can be especially troublesome in an online study, as it may be virtually impossible to determine that a participant has found a task or question extremely stressful or upsetting and needs some sort of follow-up intervention. Your research advisor and university's internal review board can help you work through ethical issues and develop appropriate precautions for any study that might potentially cause even minor harm or distress to participants.

Sampling, too, must be a source of concern in an online study. SurveyMonkey and Zoomerang enable a researcher to zero in on a defined set of participants—for example, by uploading a list of e-mail addresses to which the participation request will be sent. Other online research projects, such as those on the "Psychological Research on the Net" website mentioned earlier, are open to anyone who wants to participate. But in virtually any online study, the people who participate will not be representative either of a particular group of people or of the overall population of human beings (Gosling et al., 2004; McGraw, Tew, & Williams, 2000). After all, participants will be limited to people who (a) are comfortable with computers, (b) spend a fair amount of time on the Internet, (c) enjoy partaking in research studies, and (d) have been sufficiently enticed by your research topic to do what you ask of them. In cases where a questionnaire can be completed by anyone who has access to the Internet, many responders are apt to be college students who are earning course credit for their participation. In short, *your sample will be biased to some degree.*

Sampling is a concern for any researcher, but it is especially so for the researcher who wants to draw inferences about a large population. In the following section, we look at strategies for selecting an appropriate sample and consider how bias in sampling procedures can distort the data obtained.

Choosing a Sample in a Descriptive Study

The researcher who conducts a descriptive study wants to determine *the nature of how things are.* Especially when conducting survey research, the researcher may want to describe one or more characteristics of a fairly large population—perhaps the television viewing habits of 10-year-olds, the teaching philosophies of elementary school teachers, or the attitudes that visitors to Rocky Mountain National Park have about a shuttle bus system. Whether the population is 10-year-olds, elementary school teachers, or park visitors, we are talking about *very large* groups of people; for instance, more than 3 million people visit Rocky Mountain National Park every year.

In such situations, the researcher will, of course, usually not study the entire population of interest. Instead, he or she will select a subset, or **sample**, of the population. But the researcher can use the results obtained from the sample to make generalizations about the entire population

only if the sample is truly representative of the population. Here we are talking about a research study's *external validity*, a concept introduced in Chapter 4.

When phrasing their research problems, many novice researchers forget that they will be studying a sample rather than a population. They announce, for example, that their goal is

> to survey the legal philosophies of the attorneys of the United States and to analyze the relationship of these several philosophical positions with respect to the recent decisions of the Supreme Court of the United States.

Someone who words a problem in this way has simply not thought through the meaning of the words themselves: "The attorneys of the United States"! The American Bar Association consists of approximately 400,000 attorneys distributed over more than 3.5 million square miles in the 50 states. As we look at the problem more closely, we begin to discern other, more serious difficulties. What are "legal philosophies"? How does one isolate these philosophies to study them? How can one show a "relationship of philosophical positions" to "recent decisions of the Supreme Court"? How will this relationship be expressed? Will it be expressed statistically? If so, how would one quantify "philosophical positions" and "decisions"? If not, then how will the relationship be shown?

The difficulty in the stated research problem basically arises out of the way the problem is worded. If, on the one hand, the researcher has said what he or she means, he or she proposes to survey "the attorneys"—all of them! If, on the other hand, the researcher intends to survey only a subset of the nation's attorneys, the statement of the problem should have said so with such qualifying words as *selected, representative, typical, certain, a random sample of*, and so on. Careful researchers say precisely what they mean. Notice the difference in the meaning between "The purpose of this research is to survey the representative legal philosophies of a random sample of attorneys . . ." and the original wording, "The purpose of this research is to survey the legal philosophies of the attorneys of the United States. . . ."

The specific sampling procedure used depends on the purpose of the sampling and a careful consideration of the parameters of the population. But in general, *the sample should be so carefully chosen that, through it, the researcher is able to see characteristics of the total population in the same proportions and relationships that they would be seen if the researcher were, in fact, to examine the total population*.

When you look through the wrong end of a set of binoculars, you see the world in miniature. If the lenses aren't precision-made and accurately ground, you get a distorted view of what you are looking at. In the same way, a sample should, ideally, be a population microcosm. If the sampling procedure isn't carefully planned, any conclusions the researcher draws from the data are likely to be distorted. Such distortion is known as *bias*. We discuss various sources of bias a bit later in the chapter. For the moment, however, we concern ourselves with various sampling designs.

Sampling Designs

Different sampling designs may be more or less appropriate in different situations. Here we consider eight approaches to sampling, which fall into two major categories: probability sampling and nonprobability sampling.

Probability Sampling

In **probability sampling**, every part of the population has the potential to be represented in the sample. The sample is chosen from the overall population by **random selection**—that is, it is chosen in such a way that each member of the population has an equal chance of being selected. When such a *random sample* is selected, the researcher can assume that the characteristics of the sample approximate the characteristics of the total population.

An analogy might help. Suppose we have a beaker containing 100 ml of water. Another beaker holds 10 ml of a concentrated acid. We combine the water and acid in proportions

of 10:1. After thoroughly mixing the water and acid, we should be able to extract 1 ml from any part of the solution and find that the sample contains 10 parts water for every 1 part acid. In the same way, if we have a population with considerable variability in race, wealth, education level, social standing, and other factors, and if we have a perfectly selected random sample—a situation usually more theoretical than logistically feasible—we will find in the sample those same characteristics that exist in the larger population, and we will find them in roughly the same proportions.

A sample is no more representative of the total population than the degree to which it has been randomly selected. There are, of course, many methods of random selection. For example, we could assign each person in the population a different number and then use an arbitrary method of picking certain numbers, perhaps by using a roulette wheel (if the entire population consists of 36 or fewer members) or drawing numbers out of a hat. Many computer spreadsheet programs (e.g., Excel) also provide a means of generating random numbers.

A widely used paper-and-pencil method of selecting a random sample is to use a **table of random numbers**, such as that presented in Table 8.2. The researcher typically does not start at the beginning of the table; instead, he or she identifies a starting point randomly. One fundamental principle must be kept in mind: *The purpose of randomness is to let blind chance determine the outcomes of the selection process to as great a degree as possible.* Hence, in determining a starting point for the selection of random numbers, *pure chance* must always initiate the process.

Consider the table of random numbers presented in Table 8.2. It includes 100 blocks of numbers, arranged in 10 rows and 10 columns. The rows and columns are numbered only to assist us in choosing a starting point for using the table. Any block within the table will be at the intersection of two of these "guide" numbers. To enter the table, we need an entry number of two digits; one digit will be used to designate the row, and the other the column, for the block at which we will begin.

But how do we find an entry number? Pull a dollar bill from your wallet. The one we have just pulled as we write this book has the serial number C 45 391827A. We choose the first two digits of the serial number, which makes the entry number 45. But which is the row and which is the column? We flip a coin. If it comes down heads, the first digit will designate the row; otherwise, the digit will designate the column. The coin comes down tails. This means that we will begin in the fourth column and the fifth row. The block where the two intersect is the block where we begin within the table (see Figure 8.8).

We don't have to use a dollar bill to determine the entry point, of course. We could use any source of numbers, such as a telephone directory, a license plate, a friend's social security number, or the stock quotations page in a newspaper. Only one rule governs the final determination of an entry point: *Pure chance dictates the choice.*[6]

Having determined the starting block, we must now consider the size of the proposed sample. If it is to be fewer than 100 individuals, we will need only two-digit numbers. If it is to be more than 99 but fewer than 1,000, we will need three digits to accommodate the sample size.

At this point, let's go back to the total population to consider the group from which the sample is to be drawn. It will be necessary to designate individuals in some manner. A reasonable approach is to arrange the members of the population in a logical order—for instance, alphabetically by surname—and assign each member a serial number for identification purposes.

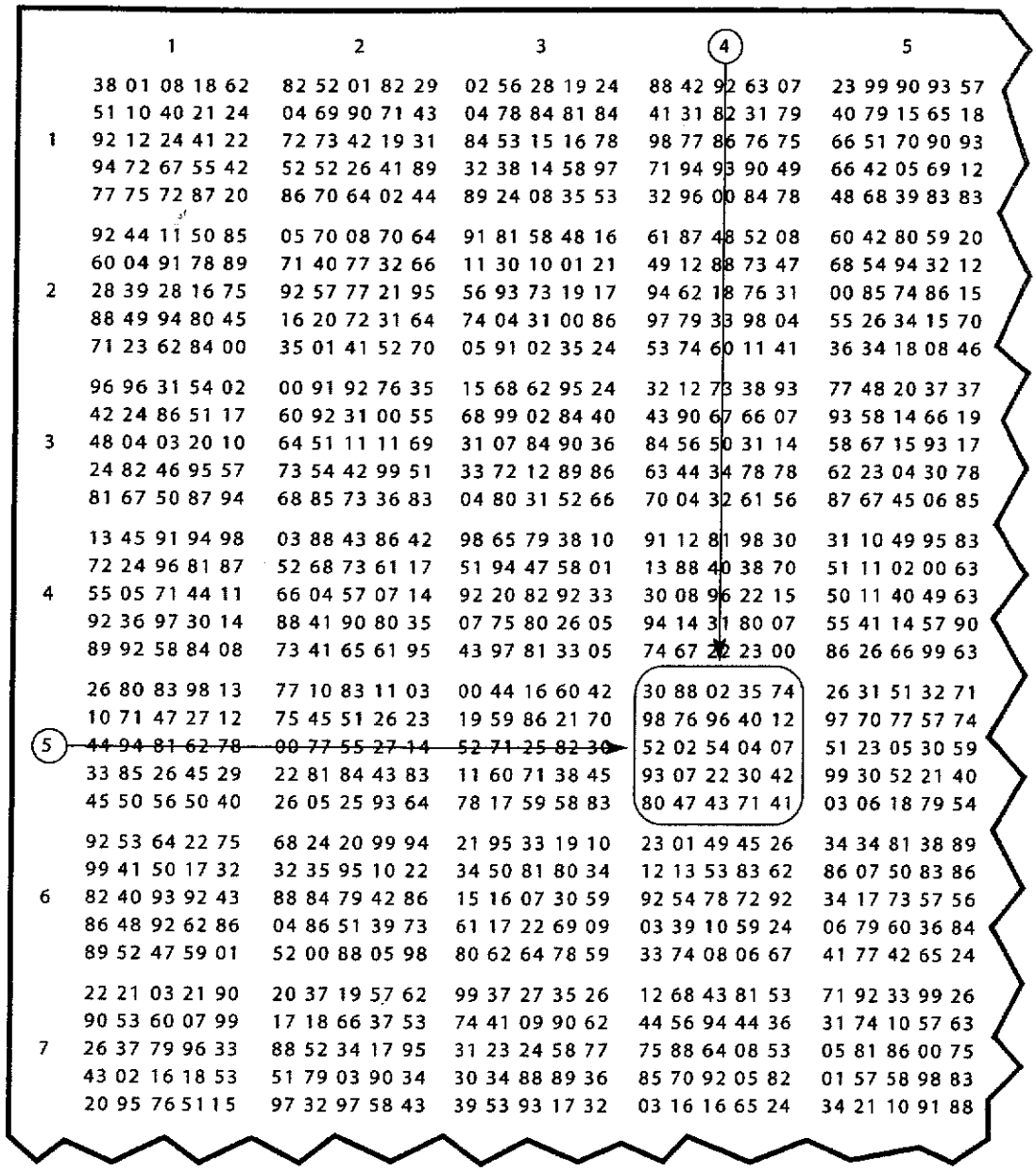
We are now ready for the random selection. We start with the upper left-hand digits in the designated starting block and work downward through the two-digit column in the rest of the table. If we need additional numbers, we proceed to the top of the next column, work our way down, and so on, until we have selected the sample we need. For purposes of illustration, we will assume that the total population consists of 90 individuals from which we will select a sample of 40. We will need random numbers of two digits each. Beginning in the upper left-hand corner of the designated block and remembering that only 90 individuals are in the total population,

[6]One of our readers has correctly pointed out that not all of the sources just suggested reflect strictly random numbers; rather, they may show a predictable pattern, with some numbers appearing more frequently than others. Nevertheless, using such a source ensures that the entry point into the table is chosen *arbitrarily*, eliminating any chance that the researcher might, either intentionally or unintentionally, tilt the sample selection in one direction or another.

TABLE 8.2 Random numbers table

38 01 08 18 62	82 52 01 82 29	02 56 28 19 24	88 42 92 63 07	23 99 90 93 57	78 10 48 55 21	29 84 46 25 60	50 24 21 62 40	56 88 09 95 57	01 78 78 86 35
51 10 40 21 24	04 69 90 71 43	04 78 84 81 84	41 31 82 31 79	40 79 15 65 18	28 86 32 62 70	55 33 27 42 50	35 64 53 48 31	15 24 21 63 23	17 59 94 69 95
92 12 24 41 22	72 73 42 19 31	84 53 15 16 78	98 77 86 76 75	66 51 70 90 93	87 83 76 49 08	32 14 91 93 50	47 49 89 17 52	05 46 40 33 05	77 27 15 16 31
94 72 67 55 42	52 52 26 41 89	32 38 14 58 97	71 94 93 90 49	66 42 05 69 12	28 08 82 64 25	27 79 96 81 66	29 86 00 94 07	84 54 57 62 22	70 80 69 62 87
77 75 72 87 20	86 70 64 02 44	89 24 08 35 53	32 96 00 84 78	48 68 39 83 83	56 47 81 77 42	39 59 64 69 02	10 19 94 50 51	40 34 19 81 31	38 54 97 22 17
92 44 11 50 85	05 70 08 70 64	91 81 58 48 16	61 87 43 52 08	60 42 80 59 20	34 28 49 91 25	19 73 63 65 72	86 97 13 99 06	08 03 07 62 52	33 68 42 59 38
60 04 91 78 89	71 40 77 32 66	11 30 10 01 21	49 12 88 73 47	68 54 94 32 12	20 77 44 49 64	06 44 46 32 33	29 70 98 08 59	64 16 94 48 45	32 22 40 94 41
28 39 28 16 75	92 57 77 21 95	56 93 73 19 17	94 62 18 76 31	00 85 74 86 15	67 31 42 98 65	39 93 31 64 30	75 28 25 82 38	25 99 97 68 63	98 45 70 95 46
88 49 94 80 45	16 20 72 31 64	74 04 31 00 86	97 79 33 98 04	55 26 34 15 70	76 70 92 40 97	79 70 00 64 33	24 29 27 30 25	86 96 27 96 51	67 14 98 05 49
71 23 62 84 00	35 01 41 52 70	05 91 02 35 24	53 74 60 11 41	36 34 18 08 46	80 67 41 60 61	29 06 20 72 42	53 96 06 98 42	81 13 97 04 71	27 27 63 85 31
96 96 31 54 02	00 91 92 76 35	15 68 62 95 24	32 12 73 38 93	77 48 20 37 37	38 54 96 41 02	71 55 92 52 72	83 88 14 67 44	60 20 86 72 56	55 39 83 52 20
42 24 86 51 17	60 92 31 00 55	68 99 02 84 40	43 90 67 66 07	93 58 14 66 19	24 57 83 41 75	81 17 09 22 06	62 51 63 29 07	48 78 30 39 28	07 12 36 20 41
48 04 03 20 10	64 51 11 11 69	31 07 84 90 36	84 56 50 31 14	58 67 15 93 17	36 07 58 97 44	23 49 17 02 59	56 11 41 67 04	46 66 29 36 60	25 24 38 04 18
24 82 46 95 57	73 54 42 99 51	33 72 12 89 86	63 44 34 78 78	62 23 04 30 78	31 36 77 96 03	07 00 17 98 53	40 50 52 29 34	71 77 22 68 97	18 04 81 99 29
81 67 50 87 94	68 85 73 36 83	04 80 31 52 66	70 04 32 61 56	87 67 45 06 85	39 67 65 57 22	79 93 95 80 66	97 60 51 95 17	94 12 85 08 16	20 28 59 36 47
13 45 91 94 98	03 88 43 86 42	98 65 79 38 10	91 12 81 98 30	31 10 49 95 83	14 60 19 90 87	95 99 76 05 42	73 99 91 91 12	62 89 71 43 96	57 23 16 02 17
72 24 96 81 87	52 68 73 61 17	51 94 47 58 01	13 88 40 38 70	51 11 02 00 63	66 16 46 92 60	98 74 78 50 78	17 07 25 83 52	18 65 27 13 89	55 91 08 06 12
55 05 71 44 11	66 04 57 07 14	92 20 82 92 33	30 08 96 22 15	50 11 40 49 53	07 91 17 27 71	46 03 95 14 94	47 94 10 17 21	23 03 23 69 95	40 24 39 29 92
92 36 97 30 14	88 41 90 80 35	07 75 80 26 05	94 14 31 80 07	55 41 14 57 90	52 12 46 94 70	44 39 99 50 47	19 52 92 92 66	26 20 82 10 07	95 75 73 66 54
89 92 58 84 08	73 41 65 61 95	43 97 81 33 05	74 67 22 23 00	86 26 66 99 63	08 89 21 14 83	40 34 50 36 28	86 84 82 22 34	25 82 86 47 85	06 38 07 97 44
26 80 83 98 13	77 10 83 11 03	00 44 16 60 42	30 88 02 35 74	26 31 51 32 71	73 00 22 00 49	22 32 76 37 24	26 97 70 87 34	04 87 06 52 84	90 74 09 74 06
10 71 47 27 12	75 45 51 26 23	19 59 86 21 70	98 76 96 40 12	97 70 77 57 74	44 22 44 97 62	42 28 17 09 55	76 04 82 14 40	76 81 74 96 79	38 81 96 85 99
41 04 81 62 78	06 77 53 27 14	52 71 25 82 93	52 02 54 04 07	51 23 05 30 59	52 40 03 90 79	48 44 83 66 71	57 21 01 45 68	50 25 03 32 31	96 92 82 16 62
33 85 26 45 29	22 81 84 43 83	11 60 71 38 45	93 07 22 30 42	99 30 52 21 40	89 18 59 29 95	23 43 14 58 06	56 71 31 48 25	27 57 85 25 76	72 63 83 95 00
45 50 56 50 40	26 05 25 93 64	78 17 59 58 83	80 47 43 71 41	03 06 18 79 54	16 23 18 55 60	08 65 50 82 07	72 25 01 31 09	53 44 59 31 94	16 87 80 32 14
92 53 64 22 75	68 24 20 99 94	21 95 33 19 10	23 01 49 45 26	34 34 81 38 89	60 81 53 29 72	13 93 48 56 37	67 99 80 12 23	12 06 07 80 23	39 20 96 85 52
99 41 50 17 32	32 35 95 10 22	34 50 81 80 34	12 13 53 83 62	86 07 50 83 86	82 21 93 36 48	56 58 87 36 67	55 45 12 45 69	88 80 59 76 58	28 90 74 84 74
82 40 93 92 43	88 84 79 42 86	15 16 07 30 59	92 54 78 72 92	34 17 73 57 56	31 46 75 31 06	65 20 76 65 17	68 67 24 41 65	83 90 16 72 41	62 28 75 63 34
86 48 92 62 86	04 86 51 39 73	61 17 22 69 09	03 39 10 59 24	06 79 60 36 84	93 97 68 61 23	15 40 89 22 72	49 06 40 32 58	59 56 50 32 05	39 01 75 86 33
89 52 47 59 01	52 00 88 05 98	80 62 64 78 59	33 74 08 06 67	41 77 42 65 24	77 95 26 51 27	38 18 74 92 72	33 78 77 53 41	88 04 45 86 61	44 98 66 68 80
22 21 03 21 90	20 37 19 57 62	99 37 27 35 26	12 68 43 81 53	71 92 33 99 26	78 69 45 98 85	87 39 43 19 80	24 88 10 64 90	78 18 95 35 38	72 70 45 72 58
90 53 60 07 99	17 18 66 37 53	74 41 09 90 62	44 56 94 44 36	31 74 10 57 63	79 62 84 26 54	88 56 10 22 55	03 90 19 50 76	07 50 95 86 69	90 00 55 69 31
26 37 79 96 33	88 52 34 17 95	31 23 24 58 77	75 88 64 08 53	05 81 86 00 75	87 59 13 45 98	63 38 70 31 41	52 50 16 84 22	73 57 74 10 88	98 63 77 24 70
43 02 16 18 53	51 79 03 90 34	30 34 88 89 36	85 70 92 05 82	01 57 58 98 83	81 19 21 57 98	43 34 00 57 50	20 67 30 37 14	39 84 53 12 90	78 65 10 23 09
20 95 76 51 15	97 32 97 58 43	39 53 93 17 32	03 16 16 65 24	34 21 10 91 88	58 22 98 68 41	54 15 20 20 14	46 70 52 78 60	06 90 74 56 50	76 31 58 65 83
37 42 58 13 23	44 82 67 97 98	54 38 79 75 11	75 72 16 32 68	74 59 37 13 10	34 52 36 07 27	98 37 33 09 60	58 31 70 14 75	49 07 71 35 11	95 06 72 41 00
31 51 93 66 48	35 56 04 55 20	29 19 24 55 97	32 79 45 92 58	76 37 11 35 79	05 68 08 06 87	64 66 88 25 52	14 82 30 95 52	01 07 97 96 23	64 99 36 57 51
26 19 62 67 67	70 61 22 01 70	18 84 36 99 94	00 57 54 28 92	45 72 40 45 06	60 23 19 24 02	03 63 69 83 81	80 85 48 37 86	47 97 87 19 83	66 94 71 49 11
42 55 29 27 78	76 81 16 63 13	04 93 73 21 04	84 91 77 67 10	50 00 21 92 50	96 91 83 66 19	00 59 90 66 01	65 48 83 28 97	72 80 43 01 65	55 64 58 40 56
90 27 74 26 17	61 75 82 44 09	42 16 28 32 77	98 36 41 39 73	97 45 85 85 08	15 25 93 29 16	19 84 12 57 24	61 33 44 76 24	54 70 39 17 47	08 99 19 55 89
91 78 11 76 19	19 09 29 73 14	94 09 50 52 42	00 57 41 82 55	34 52 10 83 68	17 58 82 71 68	72 43 90 64 99	10 49 51 53 27	52 04 26 28 62	10 70 98 42 50
53 42 75 49 95	48 74 48 88 94	52 58 34 96 85	31 34 96 82 31	47 55 84 13 48	86 83 54 45 38	32 46 69 85 86	24 47 32 60 28	29 45 36 58 21	30 84 54 36 35
09 48 74 77 55	89 30 05 25 52	93 01 06 07 66	98 09 08 65 43	51 91 90 05 91	79 45 73 93 80	89 71 17 57 63	79 51 34 10 11	12 97 06 05 00	96 77 28 11 35
23 43 73 69 25	89 55 18 26 26	87 69 88 38 12	40 83 94 06 23	96 55 93 14 80	49 89 86 62 40	81 15 86 62 40	81 15 69 78 10	18 40 92 93 88	13 01 90 70 82
59 47 86 32 53	50 76 46 09 82	07 81 75 38 35	53 87 71 45 94	64 24 70 27 23	45 39 48 71 66	76 34 85 81 21	04 93 90 43 54	92 77 48 04 07	70 52 35 99 32
80 73 71 10 81	52 63 88 33 59	35 11 58 95 64	10 50 63 30 30	07 31 89 22 05	92 53 64 22 75	68 24 20 99 94	21 95 33 19 10	23 01 49 45 26	34 34 81 38 89
67 44 30 85 08	18 69 78 50 15	88 24 02 62 46	36 89 70 02 89	78 50 98 86 01	99 41 50 17 32	32 35 95 10 22	34 50 81 80 34	12 13 53 83 62	86 07 50 83 86
35 15 98 11 06	18 64 51 44 49	37 38 41 40 68	14 65 00 37 23	87 63 85 50 66	87 40 93 92 43	88 84 79 42 86	15 16 07 30 59	92 54 78 72 92	34 17 73 57 56
66 17 63 96 90	44 57 17 04 05	85 84 64 44 05	34 06 31 13 81	92 30 29 21 61	86 48 92 62 86	04 86 51 39 73	61 17 22 69 09	03 39 10 59 24	06 79 60 36 84
53 52 48 13 33	85 33 38 32 53	77 67 16 86 82	50 81 83 25 84	53 65 33 12 88	89 52 47 59 61	52 00 88 05 98	80 62 64 78 59	33 74 08 06 67	41 77 42 65 24

FIGURE 8.8
Choosing the starting point in a random numbers table



	1	2	3	(4)	5
	38 01 08 18 62	82 52 01 82 29	02 56 28 19 24	88 42 92 63 07	23 99 90 93 57
	51 10 40 21 24	04 69 90 71 43	04 78 84 81 84	41 31 82 31 79	40 79 15 65 18
1	92 12 24 41 22	72 73 42 19 31	84 53 15 16 78	98 77 86 76 75	66 51 70 90 93
	94 72 67 55 42	52 52 26 41 89	32 38 14 58 97	71 94 93 90 49	66 42 05 69 12
	77 75 72 87 20	86 70 64 02 44	89 24 08 35 53	32 96 00 84 78	48 68 39 83 83
	92 44 11 50 85	05 70 08 70 64	91 81 58 48 16	61 87 48 52 08	60 42 80 59 20
	60 04 91 78 89	71 40 77 32 66	11 30 10 01 21	49 12 88 73 47	68 54 94 32 12
2	28 39 28 16 75	92 57 77 21 95	56 93 73 19 17	94 62 18 76 31	00 85 74 86 15
	88 49 94 80 45	16 20 72 31 64	74 04 31 00 86	97 79 33 98 04	55 26 34 15 70
	71 23 62 84 00	35 01 41 52 70	05 91 02 35 24	53 74 60 11 41	36 34 18 08 46
	96 96 31 54 02	00 91 92 76 35	15 68 62 95 24	32 12 73 38 93	77 48 20 37 37
	42 24 86 51 17	60 92 31 00 55	68 99 02 84 40	43 90 67 66 07	93 58 14 66 19
3	48 04 03 20 10	64 51 11 11 69	31 07 84 90 36	84 56 50 31 14	58 67 15 93 17
	24 82 46 95 57	73 54 42 99 51	33 72 12 89 86	63 44 34 78 78	62 23 04 30 78
	81 67 50 87 94	68 85 73 36 83	04 80 31 52 66	70 04 32 61 56	87 67 45 06 85
	13 45 91 94 98	03 88 43 86 42	98 65 79 38 10	91 12 81 98 30	31 10 49 95 83
	72 24 96 81 87	52 68 73 61 17	51 94 47 58 01	13 88 40 38 70	51 11 02 00 63
4	55 05 71 44 11	66 04 57 07 14	92 20 82 92 33	30 08 96 22 15	50 11 40 49 63
	92 36 97 30 14	88 41 90 80 35	07 75 80 26 05	94 14 31 80 07	55 41 14 57 90
	89 92 58 84 08	73 41 65 61 95	43 97 81 33 05	74 67 22 23 00	86 26 66 99 63
	26 80 83 98 13	77 10 83 11 03	00 44 16 60 42	30 88 02 35 74	26 31 51 32 71
	10 71 47 27 12	75 45 51 26 23	19 59 86 21 70	98 76 96 40 12	97 70 77 57 74
(5)	~~44 94 81 62 78~~	~~00 77 55 27 14~~	~~52 71 25 82 30~~	52 02 54 04 07	51 23 05 30 59
	33 85 26 45 29	22 81 84 43 83	11 60 71 38 45	93 07 22 30 42	99 30 52 21 40
	45 50 56 50 40	26 05 25 93 64	78 17 59 58 83	80 47 43 71 41	03 06 18 79 54
	92 53 64 22 75	68 24 20 99 94	21 95 33 19 10	23 01 49 45 26	34 34 81 38 89
	99 41 50 17 32	32 35 95 10 22	34 50 81 80 34	12 13 53 83 62	86 07 50 83 86
6	82 40 93 92 43	88 84 79 42 86	15 16 07 30 59	92 54 78 72 92	34 17 73 57 56
	86 48 92 62 86	04 86 51 39 73	61 17 22 69 09	03 39 10 59 24	06 79 60 36 84
	89 52 47 59 01	52 00 88 05 98	80 62 64 78 59	33 74 08 06 67	41 77 42 65 24
	22 21 03 21 90	20 37 19 57 62	99 37 27 35 26	12 68 43 81 53	71 92 33 99 26
	90 53 60 07 99	17 18 66 37 53	74 41 09 90 62	44 56 94 44 36	31 74 10 57 63
7	26 37 79 96 33	88 52 34 17 95	31 23 24 58 77	75 88 64 08 53	05 81 86 00 75
	43 02 16 18 53	51 79 03 90 34	30 34 88 89 36	85 70 92 05 82	01 57 58 98 83
	20 95 76 51 15	97 32 97 58 43	39 53 93 17 32	03 16 16 65 24	34 21 10 91 88

we see that the first number in the leftmost column is 30, so we choose individual number 30 in the population. The next number (98) does not apply because only 90 persons are in the population. Our next choice is 52, we ignore 93, and then we choose 80. Proceeding to the next block down, we choose 23 and 12, ignore 92, choose 3 and 33. We continue down the column and proceed to any additional columns we need, ignoring the numbers 91–99 and 00 and any numbers we have already selected, until we get a sample of 40.

We have probably said enough about the use of a random numbers table. We turn now to specific probability sampling techniques.

Simple random sampling. Simple random sampling is exactly the process just described: Every member of the population has an equal chance of being selected. Such an approach is easy when the population is small and all of its members are known. For example, one of us authors once used it in a study to evaluate the quality of certain teacher training institutes during the summer of 1992 (Cole & Ormrod, 1995). Fewer than 300 people had attended the institutes,

and we knew who and where they all were. But for very large populations—for instance, all 10-year-olds or all lawyers—simple random sampling is neither practical nor, in many cases, possible.

Stratified random sampling. Think of grades 4, 5, and 6 in a public school. This is a *stratified population*. It has three different layers (*strata*) of distinctly different types of individuals. In stratified random sampling, the researcher samples equally from each one of the layers in the overall population.

If we were to sample a population of fourth-, fifth-, and sixth-grade children in a particular school, we would assume that the three strata are roughly equal in size (i.e., there are similar numbers of children at each grade level), and so we would take equal samples from each of the three grades. Our sampling method would look like that in Figure 8.9.

Stratified random sampling has the advantage of guaranteeing equal representation of each of the identified strata. It is, of course, most appropriate when those strata are roughly equal in size in the overall population as well.

Proportional stratified sampling. In the simple stratified random sampling design just described, all strata of the population are essentially equal in size. But now we come to a different situation. Consider a community that has, for example, 1,000 Jewish people, 2,000 Catholics, and 3,000 Protestants. Let's imagine a survey situation. A local newspaper publishes a section dealing with interfaith church news, religious events, and syndicated articles of interest to the religious community in general. The editor wants to obtain certain information and opinions from the paper's readers and so decides to conduct a survey.

In this situation, the editor chooses his sample in accordance with the proportions of each religious group in the paper's readership. For every Jewish person, there should be two Catholics and three Protestants. In this situation, the people are not obviously segregated into the different strata, so the first step is to identify the members of each stratum and then select a random sample from each one. Figure 8.10 schematically represents this type of sampling.

Cluster sampling. Sometimes the population of interest is spread out over a large area. It may not be feasible to make up a list of every person living within the area and, from the list,

FIGURE 8.9
Stratified random sampling design

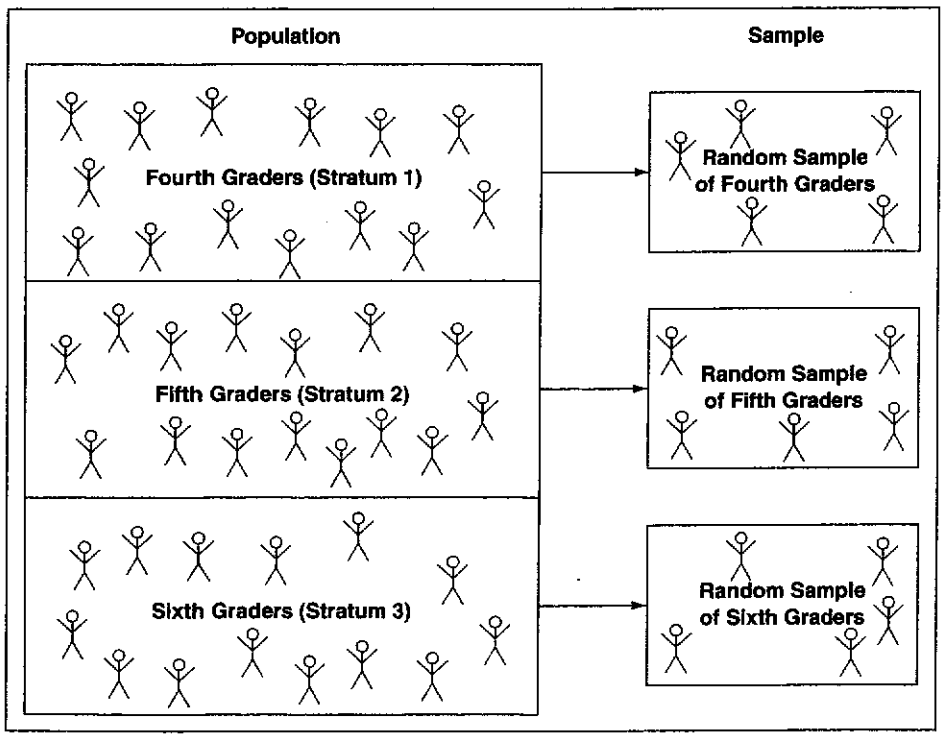


TABLE 8.3 Population characteristics and probability sampling techniques appropriate for each population type

Population Type	Example	Appropriate Sampling Technique(s)
1. Population is generally a homogeneous group of individual units.	A particular variety of flower seeds, which a researcher wants to test for germination potential.	• Simple random sampling • Systematic sampling of individual units (when large populations of human beings are involved)
2. Population contains definite strata that are approximately equal in size.	A school with six grade levels: kindergarten, first, second, third, fourth, and fifth.	• Stratified random sampling
3. Population contains definite strata that appear in different proportions within the population.	A community in which residents are Catholic (25%), Protestant (45%), Jewish (15%), Muslim (5%), or nonaffiliated (10%).	• Proportional stratified sampling
4. Population consists of discrete clusters with similar characteristics. The units within each cluster are as heterogeneous as units in the overall population.	Travelers in the nation's 20 leading air terminals. (It is assumed that all air terminals are similar in atmosphere, purpose, design, etc. The passengers who use them differ widely in such characteristics as age, gender, national origin, socioeconomic status, and belief system, with such variability being similar from one airport to the next.)	• Cluster sampling • Systematic sampling (of clusters)

Nonprobability Sampling

In **nonprobability sampling**, the researcher has no way of predicting or guaranteeing that each element of the population will be represented in the sample. Furthermore, some members of the population have little or no chance of being sampled. Following are three common forms of nonprobability sampling.

Convenience sampling. Convenience sampling—also known as *accidental sampling*—makes no pretense of identifying a representative subset of a population. It takes people or other units that are readily available—for instance, those that arrive on the scene by mere happenstance.

Convenience sampling may be quite appropriate for some research problems. For example, suppose you own a small restaurant and want to sample the opinions of your patrons on the quality of food and service at your restaurant. You open for breakfast at 6 a.m., and on five consecutive weekdays you question the first 40 patrons who arrive. Customers who have on one occasion expressed an opinion are eliminated on subsequent arrivals. The opinions you eventually get are from 36 men and 4 women. It is a heavily lopsided poll in favor of men, perhaps because the people who arrive at 6 a.m. are likely to be in certain occupations—laborers, construction workers, truck drivers—that are predominantly male. The data from this convenience sample give you the thoughts of robust, hardy men about your breakfast menu—that's all. Yet such information may be all you need for your purpose.

Not all research data need to be collected through careful, thoughtful sampling procedures. But without such safeguards, the conclusions drawn from the research may not be trustworthy.

Quota sampling. Quota sampling is a variation of convenience sampling. It selects respondents in the same proportions that they are found in the general population, but not in a random fashion. Let's consider a population in which the number of African Americans equals the number of European Americans. Quota sampling would choose, say, 20 African Americans and 20 European Americans, but without any attempt to select these individuals randomly from the overall population.

For example, suppose you are a reporter for a television station. At noon, you position yourself with microphone and television camera beside Main Street in the center of a particular city. As people pass, you interview them. The fact that people in the two categories may come in clusters of two, three, or four is no problem. All you need are the opinions of 20 people from each category. This type of sampling regulates only the size of each category within the sample; in every other respect, the selection of the sample is nonrandom and, in most cases, convenient.

FIGURE 8.11
Cluster sampling design

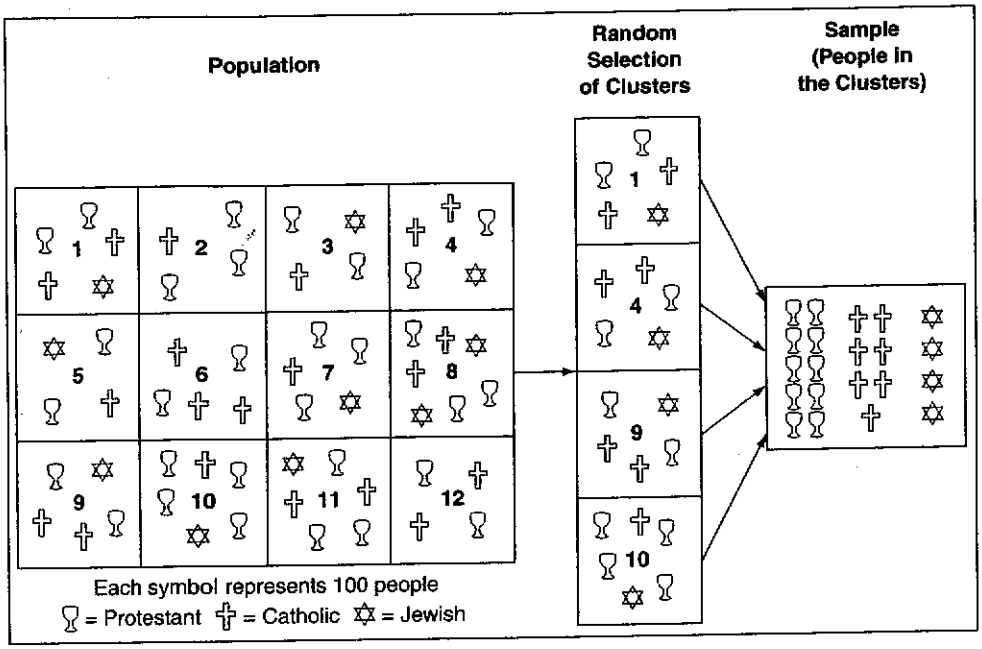
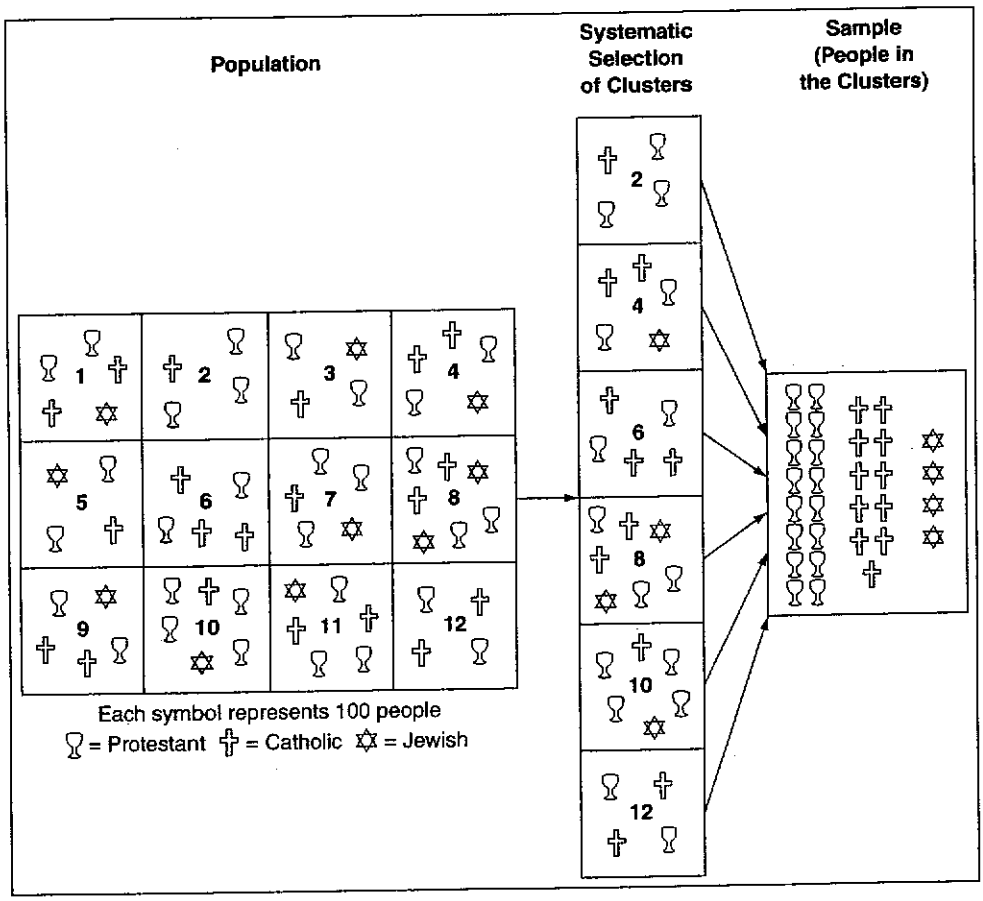


FIGURE 8.12
Systematic sampling design



A sampling design should not be chosen blindly or willy-nilly. Each of the designs just described is uniquely suited to a particular kind of population, and you should therefore consider the nature of your population when selecting your sampling technique. The design diagrams in Figures 8.9 through 8.12 can assist you in your choice.

Purposive sampling. In purposive sampling, people or other units are chosen, as the name implies, for a particular *purpose*. For instance, we might choose people who we have decided are "typical" of a group or those who represent diverse perspectives on an issue.

Pollsters who forecast elections frequently use purposive sampling: They may choose a combination of voting districts that, in past elections, has been quite useful in predicting the final outcomes.

Purposive sampling may be very appropriate for certain research problems. However, the researcher should always provide a rationale explaining why he or she selected the particular sample of participants.

Sampling in Surveys of Very Large Populations

Nowhere is sampling more critical than in survey research. Frequently the researcher reports that x percent of people believe such-and-such, that y percent do so-and-so, or that z percent are in favor of a particular political candidate. *Such percentages are meaningless unless the sample is representative of the population about which inferences are to be drawn.*

A basic rule governs survey research: Nothing emerges from a long, involved study that is any better than the care, precision, and thought that went into the basic planning of the research design and the selection of the population. The results of a survey are no more trustworthy than the representativeness of the sample. Population parameters and sampling procedures are of paramount importance and become critical factors in the success of the study.

But now imagine that a researcher wants to conduct a survey of the country's *entire adult population*. How can the researcher possibly hope to get a random, representative sample of such a large group of people? The Survey Research Center of the University of Michigan's Institute for Social Research (1976) has used what it calls a *multistage sampling of areas:*

1. *Primary area selection.* The country is divided into small "primary areas," each consisting of a specific county, a small group of counties, or a large metropolitan area. A predetermined number of these areas are randomly selected.
2. *Sample location selection.* Each of the selected primary areas is divided into smaller sections ("sample locations"), such as specific towns. A small number of these locations is randomly selected.
3. *Chunk selection.* The sample locations are divided into even smaller "chunks" that have identifiable boundaries such as roads, streams, or the edges of a city block. Most chunks have 16 to 50 dwellings, although the number may be larger in large cities. Once again, a random sample is selected.
4. *Segment selection.* Chunks are subdivided into areas containing a relatively small number of dwellings, and some of these "segments" are, again, chosen randomly.
5. *Housing unit selection.* Approximately four dwellings are selected (randomly, of course) from each segment, and the residents of those dwellings are asked to participate in the survey. If a doorbell is unanswered, the researcher returns at a later date and tries again.

As you may have deduced, the approach just described is a multistage version of cluster sampling (see Figure 8.13). At each stage of the game, units are selected randomly. "Randomly" does *not* mean haphazardly or capriciously. Instead, a mathematical procedure is employed to ensure that selection is entirely random and the result of blind chance. This process should yield a sample that is, in all important respects, representative of the country's population.

Identifying a Sufficient Sample Size

A basic rule in sampling is: *The larger the sample, the better*. But such a generalized rule is not very helpful to a researcher who must make a practical decision about a specific research situation. Obviously, we need to provide more guidance here. Gay, Mills, and Airasian (2009, p. 133) have offered the following guidelines for selecting a sample size, which we'll refer to by the symbol N:

- For smaller populations, say, $N = 100$ or fewer, there is little point in sampling; survey the entire population.

> The house on the muddy back road, the apartment at the top of a long flight of stairs, the house with the growling dog outside must each have an opportunity to be included in the sample. People who live on back roads can be very different from people who live on well paved streets, and people who stay at home are not the same as those who tend to be away from home. If you make substitutions, such important groups as young men, people with small families, employed women, farmers who regularly trade in town, and so on, may not have proportionate representation in the sample. (Survey Research Center, 1976, p. 37)

In research, **bias** is any influence, condition, or set of conditions that singly or in combination distort the data. Data are, in many respects, delicate and sensitive to unintended influences. We talk about the hard facts, the solid truth, and yet every researcher soon learns that data are neither so hard nor so solid as the phrases might suggest. Data are highly susceptible to distortion.

Bias can creep into a research project in a variety of subtle and undetected ways. It can be easily overlooked by even the most careful and conscientious researcher. For instance, when conducting an interview, the researcher's personality may affect the interviewee's responses. In asking questions, the researcher's tone of voice or emphasis within a sentence may influence how a respondent replies.

Bias attacks the integrity of the facts. It is especially vicious when it enters surreptitiously into the research system and goes undetected. It can render suspect even the most carefully planned research effort.

Among the conditions that lead to bias is any influence that may have disturbed the randomness with which a sample population has been selected. Here we are talking about **sampling bias**.

The best way to appreciate sampling bias is to see it at work. Suppose that a researcher wants to conduct a survey of a certain city's population and decides to use the city telephone book as a source for selecting a random sample. She opens to a page at random, closes her eyes, puts her pencil down on the page, and selects the name that comes closest to the pencil point. You can't get much more random than that, she reasons. But the demon of bias is there. Her possible selections are limited to people who are listed in the phone book. People with very low income levels won't be adequately represented because some of them cannot afford telephone service. Nor will wealthy individuals be proportionally represented because many of them have unlisted numbers. And, of course, people who use only cell phones—people who, on average, are fairly young—aren't included in the phone book. Hence, the sample will consist of a greater percentage of people at middle-income levels and in older age-groups than exists in the city's overall population (e.g., Keeter, Dimock, Christian, & Kennedy, 2008).

As noted earlier in the chapter, studies involving Internet-based questionnaires are apt to be biased in favor of computer-literate individuals with easy access to the Internet. Studies involving mailed questionnaires frequently fall victim to bias as well, often without the researcher's awareness. Let's take a simple situation. Suppose a questionnaire is sent to 100 citizens, asking, "Have you ever been audited by the Internal Revenue Service (IRS) to justify your income tax return?" Of the 70 questionnaires returned, 35 are from people who indicate they have been audited, whereas 35 are from people who indicate they have never been audited. The researcher might therefore conclude that 50% of American citizens are likely to be audited by the IRS at one time or another.

The researcher's generalization may not be accurate. We need to consider how the nonrespondents—30% of the original sample—might be different from those who responded to the questionnaire. Many people consider an IRS audit to be a reflection of their integrity. Perhaps for this reason, some individuals in the researcher's sample may not have wanted to admit that they had been audited and so tossed the questionnaire into the wastebasket. If previously audited people were less likely to return the questionnaire than nonaudited people, the sample was biased, and thus the results didn't accurately represent the truth of the matter. Perhaps, instead of a 50-50 split, an estimate of 65% (people audited) versus 35% (people not audited) is more accurate. The data the researcher obtained do not enable the researcher to make such an estimate, however.

The examples just presented illustrate two different ways in which bias may creep into the research sample. In the case of the telephone survey, *sample selection* itself was biased because not

everyone in the population had an equal chance of being selected. In fact, people not listed in the telephone directory had *zero* chance of being selected. Here we see the primary disadvantage of nonprobability sampling, and especially of convenience sampling: People who happen to be readily available for a research project—those who are in the right place at the right time—are almost certainly *not* a random sample of the overall population.

In the example concerning IRS audits, *response rate*—and, in particular, potential differences between respondents and nonrespondents—was the source of bias. In that situation, the researcher's return rate of 70% was quite high. More often, however, the return rate in a questionnaire study is 50% or less, and the more nonrespondents there are, the greater the likelihood of bias. Likewise, in telephone surveys, a researcher won't necessarily reach certain people even with 10 or more attempts, and those who *are* eventually reached won't all agree to an interview (Witt & Best, 2008).

Nonrespondents to *mailed questionnaires* may be different from respondents in one or more ways (Rogelberg & Luong, 1998). They may have less interest in the topic being studied. They may have illnesses, disabilities, or language barriers that prevent them from responding. And on average, they have lower educational levels. In contrast, people who are hard to reach *by telephone* are apt to be young working adults who are *more* educated than the average individual (Witt & Best, 2008). To the extent that such characteristics affect how people respond to a survey, bias will exist in the data the survey yields.

Acknowledging the Probable Presence of Bias

It is almost impossible for people to live in this world without coming into contact with disease-bearing germs and other microorganisms. Likewise, in the research environment, the researcher cannot avoid having data contaminated by bias of one sort or another. What is unprofessional, however, is for the researcher to fail either to recognize or to acknowledge the likelihood of biased data in the study. When formulating conclusions about the data, a researcher must be sure to consider the effect that bias may have had in distorting the data.

In survey research, you should *always* report the percentages of people who have and have not consented to participate, such as those who have agreed and refused to be interviewed or those who have and have not returned questionnaires. Furthermore, you should be candid about possible sources of bias that result from differences between participants and nonparticipants. Rogelberg and Luong (1998) have suggested several strategies for identifying possible bias in questionnaire research; we list three especially useful ones here:

1. Carefully scrutinize the questionnaire for items that might be influenced by one's education level, interest in the topic, or other factors that frequently distinguish respondents from nonrespondents.
2. Compare the responses on questionnaires that were returned quickly with responses on those that were returned later, perhaps after a second reminder letter or after the deadline you imposed. The late ones may, to some extent, reflect the kinds of responses that nonrespondents would have given. Significant differences between the early and late questionnaires probably indicate bias in your results.
3. Randomly select a small number of nonrespondents and try to contact them by mail or telephone. Present an abridged version of your survey, and, if some people reply, match their answers against those in your original set of respondents.

One of us authors once used a variation on the third strategy in the study of summer training institutes mentioned earlier in the chapter (Cole & Ormrod, 1995). A research assistant had sent questionnaires to all attendees at one summer's institutes so that the institutes' leaders could improve the training sessions the following year, and she had gotten a return rate of 50%. She placed telephone calls to small random samples of both respondents and nonrespondents and asked a few of the questions that had been on the questionnaire. She obtained similar responses from both groups, leading the research team to conclude that the responses to the questionnaire were probably fairly representative of the entire population of institute participants.

Good researchers demonstrate their integrity by admitting, without reservation, that bias is omnipresent and may well have influenced their findings. Ideally, they point out precisely how bias may have infiltrated the research design. With this knowledge, other scholars can realistically appraise the research and judge its merits.

PRACTICAL APPLICATION Population Analysis for a Descriptive Study

Select a particular population and conduct an analysis of its structure and characteristics. Analyze the population you have chosen by completing the following checklist.

✔ CHECKLIST

Analyzing a Population

_____ 1. On the following line, identify the particular population you have chosen:

_____

_____ 2. Now answer the following questions with respect to the *structure of the population:*

	Yes	No
a. Is the population a relatively homogeneous group of individuals or other units?	_____	_____
b. Could the population be considered to consist generally of equal "layers," each of which is fairly homogeneous in composition?	_____	_____
c. Could the population be considered to be composed of separate homogeneous layers differing in size and number of units comprising them?	_____	_____
d. Could the population be envisioned as isolated islands or clusters of individual units, with the clusters being similar to one another in composition?	_____	_____

_____ 3. Through what means would you extract a representative sample from the total population? Describe your procedure on the following lines:

_____

_____

_____

_____ 4. Refer to Table 8.3. Is your sampling procedure appropriate for the characteristics of the population? _____ Yes _____ No

_____ 5. Have you guaranteed that your sample will be chosen by chance and yet will be representative of the population? _____ Yes _____ No

_____ 6. If the preceding answer is yes, explain how this will be done.

_____

_____

_____

_____ 7. Indicate what means will be employed to obtain the information you need from the sample.

_____

_____

- What procedures should I follow to obtain the necessary information? How should I implement those procedures?
- How do I get a sample that will truly be reflective of the entire population about which I am concerned?
- How can I collect my data in a way that minimizes misrepresentations and misunderstandings?
- How can I control for possible bias in the collection and description of the data?
- What do I do with the data once I have collected them? How do I organize and prepare them for analysis?
- Above all, in what ways might I reasonably interpret the data? What conclusions might I reach from my investigation?

A Sample Dissertation

We conclude the chapter by illustrating how questionnaires might be used in a correlational study to address the topic of violence in intimate relationships (e.g., husband and wife, boyfriend and girlfriend) in American society. The excerpts we present are from Luis Ramirez's doctoral dissertation in sociology completed at the University of New Hampshire (Ramirez, 2001).

Ramirez hypothesizes that violence between intimate partners—in particular, assault by one partner on the other—is, in part, a function of ethnicity, acculturation (e.g., adoption of mainstream American behaviors and values), criminal history, and social integration (e.g., feelings of connectedness with family and friends). He further hypothesizes that as a result of such factors, differences in intimate partner violence might be observed in Mexican Americans and non-Mexican Americans.

Ramirez begins Chapter 1 by discussing the prevalence of violence (especially assault) in intimate relationships. We pick up Chapter 1 at the point where he identifies his research questions and hypotheses. We then move into Chapter 2, where he describes his methodology. As has been true for earlier proposal and dissertation samples, the research report appears on the left-hand side, and our commentary appears on the right.

Dissertation Abstract 5

RESEARCH QUESTIONS

[T]he following questions will be addressed: What role does acculturation into American society have on intimate partner violence for Mexican Americans? What are the effects of a person's criminal history on intimate partner violence? What is the extent and the relation of criminal history to intimate partner violence, and is criminal history restricted to one type of crime or is it a more general tendency (violent versus property crimes)? Are crimes that are committed early in life more indicative of a pattern of crime as compared to crimes that begin later in life? Do people who assault their partners possess weak social bonds with the society they live in? Finally, this study will ask the question, "Are there differences between criminal history and bond to society for Mexican Americans and Non-Mexican Whites, and how do these factors affect intimate partner violence?"

Comments

To understand factors underlying violence in intimate partner relationships—his main research problem—the author identifies a number of subproblems, which he expresses here as research questions.

_____ 8. What are the weaknesses inherent in this method of obtaining the data?

_____

_____

_____ 9. What safeguards have you established to counteract any potential bias in your approach to data collection? Be specific.

_____

_____

_____

Interpreting Data in Descriptive Research

Data are of little or no value merely as data. In our discussion of descriptive research methods, we have primarily discussed the acquisition of data: how to obtain the data from the general population with appropriate techniques (observations, interviews, questionnaires, sampling) and how to protect those data against distortion of bias. We have been thinking of the process of data collection only.

At this juncture, we remind you of two basic principles of research:

1. The purpose of research is to seek the answer to a problem in the light of the data that relate to the problem.
2. Although collecting data for study and organizing it for inspection require care and precision, extracting meaning from the data—the interpretation of the data—is all-important.

A descriptive study is often a very "busy" research method: The researcher must decide on a population; choose a technique for sampling it; develop a valid means of collecting the desired information; minimize the potential entrance of bias into the study; and then actually collect, record, organize, and analyze all of the necessary data. The activities connected with descriptive research are complex, time-consuming, and occasionally distracting. Therein lies an element of danger. With all this action going on, it would not be surprising if the researcher lost sight of the problem and subproblems. But the problem and its subproblems are precisely the reason for the entire endeavor.

All research activity is subordinate to the research problem itself. Sooner or later, the entire effort must result in an interpretation of the data and a setting forth of conclusions, drawn from the data, to resolve the problem under investigation. Inexperienced researchers sometimes forget this fact. Activity for activity's sake is seductive. Amassing great quantities of data can provide a sense of well-being. Like Midas looking at his hoard of gold, researchers might lose sight of the ultimate demands that the problem itself makes on those data. Presenting the data in displays and summaries—graphs, charts, tables—does nothing more than demonstrate the researcher's acquisitive skills and consummate ability to present the same data in various ways. Descriptive research ultimately aims to solve problems through the *interpretation* of the data that have been gathered.

Some Final Suggestions

As we approach the end of the chapter, it is important to reflect on several issues related to descriptive research. Consider each of the following questions within the context of the research project you have in mind:

- Why is a description of this population and/or phenomenon valuable?
- What specific data will I need to solve my research problem and its subproblems?

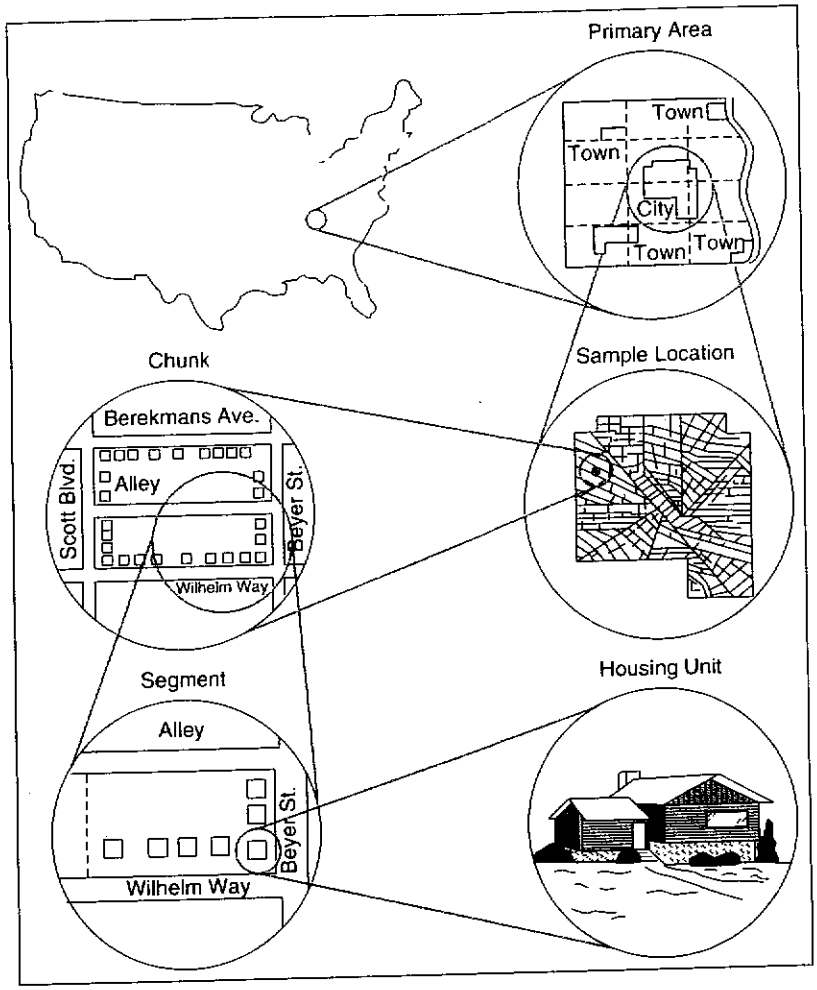


FIGURE 8.13
Multistage sampling

From the *Interviewer's Manual* (Rev. ed., p. 36) by the Survey Research Center, Institute for Social Research, 1976, Ann Arbor: University of Michigan. Reprinted with permission.

If the population size is around 500 (give or take 100), 50% should be sampled.
If the population size is around 1,500, 20% should be sampled.
Beyond a certain point (about $N = 5{,}000$), the population size is almost irrelevant and a sample size of 400 will be adequate.

Generally speaking, then, the larger the population, the smaller the percentage—but *not* the smaller the number!—one needs to get a representative sample.

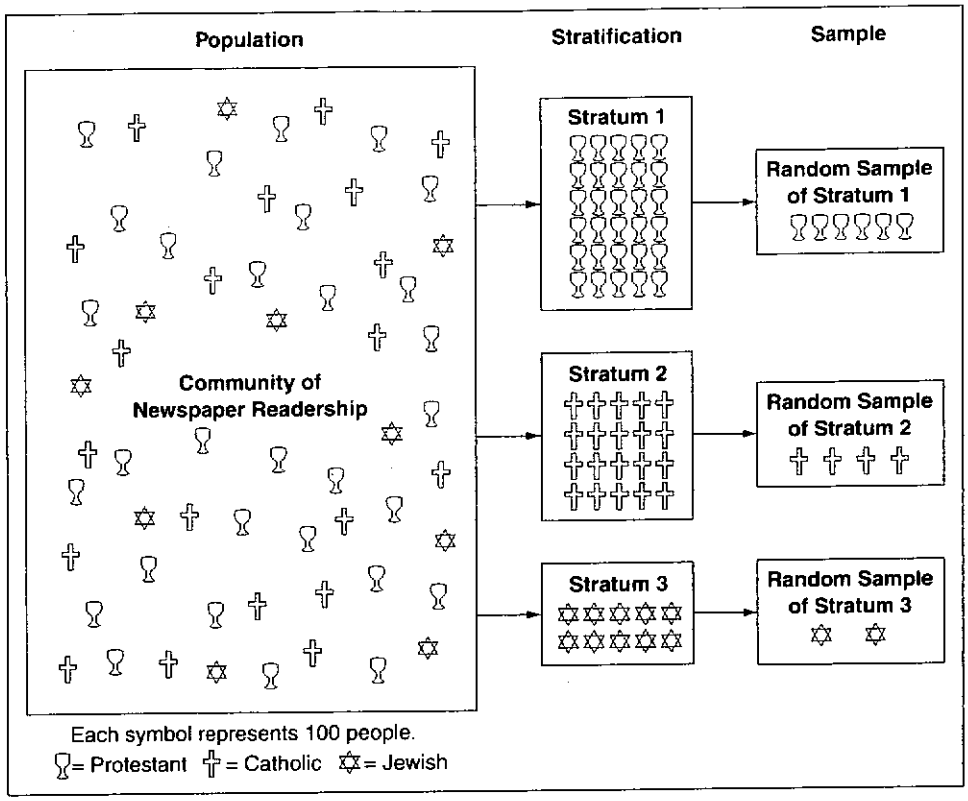
To some extent, the size of an adequate sample depends on how homogeneous or heterogeneous the population is—how alike or different its members are with respect to the characteristics of research interest. If the population is markedly heterogeneous, a larger sample will be necessary than if the population is fairly homogeneous. Important, too, is the degree of precision with which the researcher wishes to draw conclusions or make predictions about the population under study.

Statisticians have developed formulas for determining the desired sample size for a given population. Such formulas are beyond the scope of this book, but you can find them in many introductory statistics books (e.g., Lind, 2006; Triola, 2008) and on many Internet websites (e.g., search "calculating sample size").

Bias in Research Sampling

Look once again at the five steps in the University of Michigan's Survey Research Center procedure for obtaining a sample in a national survey. Notice the last sentence in the fifth step: "If a doorbell is unanswered, the researcher returns at a later date and tries again." The researcher does *not* substitute one housing unit for another; doing so would introduce *bias* into the sampling design. The center's *Interviewer's Manual* describes such bias well:

FIGURE 8.10
Proportional stratified sampling design



select a sample for study through normal randomization procedures. Instead, we might obtain a map of the area showing political boundaries or other subdivisions. We can then subdivide an expansive area into smaller *units*. For example, a city can be subdivided into precincts, clusters of city blocks, or school boundary areas; a state can be divided into counties or townships. In cluster sampling, it is important that the clusters be as similar to one another as possible, with each cluster containing an equally heterogeneous mix of individuals.

A subset of the identified clusters is randomly selected. The sample consists of the people within each of the chosen clusters. Using our example of community religious groups, let's assume the community is a large city that we have divided into 12 areas, or clusters. We randomly select clusters 1, 4, 9, and 10, and their members become our sample. This sampling design is depicted in Figure 8.11.

Systematic sampling. Systematic sampling involves selecting individuals—or perhaps clusters—according to a predetermined sequence. The sequence must originate by chance. For instance, we might create a randomly scrambled list of units that lie within the population of interest and then select every 10th unit on the list.

Let's take the cluster diagram presented in Figure 8.11. The population has 12 cells, or clusters. Half the cell numbers are odd, and the other half are even. Using the systematic sampling technique, we choose, by *predetermined sequence*, the clusters for sampling. Let's toss a coin. Heads dictates that we begin with the first odd-numbered cluster; tails dictates that we begin with the first even-numbered cluster. The coin comes down tails, which means that we start with the first even-numbered digit, which is 2, and select the systematically sequential clusters 4, 6, 8, 10, 12. Figure 8.12 shows the systematic sampling design as used for sampling clusters.

Table 8.3 identifies the various kinds of populations for which different probability sampling techniques might be appropriate in a research study.

If relations are found between these characteristics, it suggests that social agencies that deal with intimate partner violence need to adjust their policies and intervention procedures to better meet the characteristics of their clients. The focus of primary prevention could be put on the social bonding process, the criminal history of the individual, or the acculturation process in order to help solve future problems. Furthermore, a comparative study of intimate partner assault among ethnic groups could provide further clarification to a body of literature and research that has produced mixed results.

Here the author addresses the importance of the study, both pragmatic (results have potential implications for social policy and practice) and theoretical (results may shed light on inconsistencies in previous research studies).

[The author briefly reviews theoretical frameworks related to ethnicity and acculturation, criminal history, and control theory, which he then uses as a basis for his hypotheses.]

HYPOTHESES

The hypotheses are organized by the theoretical frameworks from which they have been derived, helping the reader connect them to rationales the author has previously provided.

The theoretical frameworks reviewed led to the following hypotheses:

Ethnicity and Acculturation

1. The rate of intimate partner violence is lower for Mexican Americans than Non-Mexicans.
2. The higher the acculturation into American Society, the higher the probability of assaulting a partner for Mexican Americans.

Notice how the hypotheses are single-spaced. Single-spaced hypotheses often appear in theses and dissertations, but check the guidelines at your own institution to see whether such formatting is desired.

Criminal History

3. Criminal history is more prevalent for Mexican Americans than for Non-Mexicans.
4. The more crimes committed in the past, the higher the probability of physically assaulting a partner.
5. Criminal history is more associated with an increased risk of intimate partner violence for Mexican Americans than Non-Mexicans.
6. Early onset crime is more associated with an increased risk of intimate partner violence than criminal behavior beginning later in life.
7. Previous violent crime is more associated with an increased risk of intimate partner violence than property crime.

Social Integration

8. Mexican Americans are more socially integrated than Non-Mexican Whites.
9. The more socially integrated an individual is, the lower the probability of physically assaulting a partner.
10. Social integration is more associated with a decreased risk of intimate partner violence for Mexican Americans than Non-Mexicans.

An in-depth review of the literature is postponed until Chapters 3 through 5, where the author also relates his own results to previous research findings. Although this is an unusual organizational structure, it works well in this situation, allowing the reader to connect results relative to each hypothesis to the appropriate body of literature.

A more detailed review of the literature will be presented in . . . following chapters. Literature for all hypotheses will be reviewed in their respective chapters.

Figure 1.1 is a diagramed representation of what I believe is the causal process that could affect intimate partner violence. It includes demographic and control variables, the main independent variables (acculturation, criminal history, social integration), and intimate partner violence. These variables will be described in detail in the next chapter.

Note the transition to the next chapter, which immediately follows.

FIGURE 1.1 Model of Intimate Partner Violence

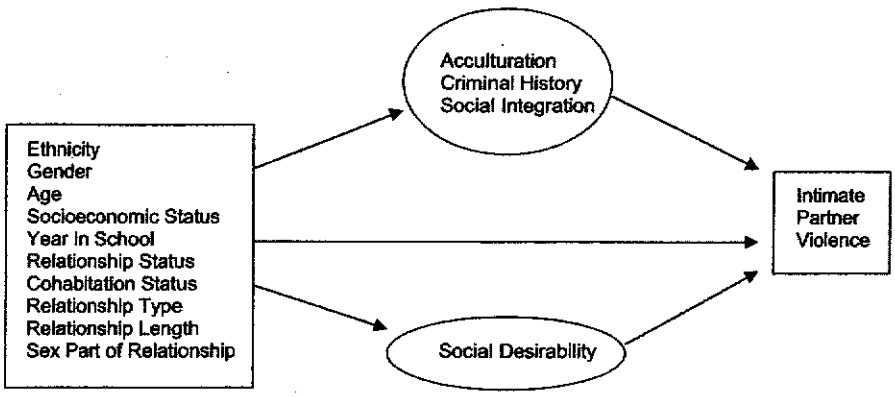


Figure 1.1 effectively condenses and summarizes the researcher's hypotheses. Also, it graphically demonstrates that four variables—acculturation, criminal history, social integration, and social desirability—are hypothesized to be mediating variables *in the relationship between demographics and violence.*

CHAPTER 2

Methods

Some style manuals suggest that an author include at least a small amount of text between two headings of different levels. For example, before beginning the "Sample" section, the author might provide an advance organizer, describing the topics he will discuss in the chapter and in what order.

Sample

The issues discussed in the previous chapter will be investigated using data from a sample of college students who have been or are currently in a dating or married relationship. A sample of college students is appropriate for this study for the following reasons: (1) The National Crime Victimization Survey found that the rates of non-lethal intimate partner violence was greatest for the 20 to 24 year age group, followed by the 16 to 19 age group, and then the 25 to 34 age group (Renison & Welchans, 2000). The majority of college students fall into the high-risk age categories. Sugarman and Hotaling (1989) identified eleven studies that provided rates for physical assault of dating partners and concluded the rates of assaulting a partner range from 20% to 59%. (2) College students make up about a third of the 18 to 22 year old population. College students are a sizable population in reference to the general population (about 15 million). (3) College students are in a formative period of their lives in relation to the habits that they develop with an intimate partner. These habits could surface in other intimate relations (O'Leary, Malone, & Tyree, 1994; Pan, Neidig, & O'Leary, 1994).

It is important to mention that a sample of college students is not a representative sample of the general population in the United States. This group generally has lower levels of criminal behavior, substance abuse, and marriage rates. Additionally, college students may be more socially integrated into society and are engaged in education as a tool for upward mobility. In short, this is a segment of society that plays by the rules.

The author clearly realizes that his sample (college students) is not representative of the entire U.S. population. He presents a good case that the sample is quite appropriate for his research questions. At the same time, he acknowledges that his sample has some shortcomings.

Data Collection

Six hundred and fifty questionnaires were passed out to students at The University of Texas at El Paso and Texas Tech University during the fall 1999, spring 2000, and summer 2000 semesters. Students who were enrolled in Sociology, Anthropology, and History classes [were] the respondents.

The author, whose home town is El Paso, has numerous acquaintances at both institutions and so can easily gain access to these students. He must, of course, seek approval from the internal review boards at the two institutions, as well as at the institution where he is completing his doctorate.

Respondents filled out the questionnaire (Appendix A) in a classroom setting. Each respondent received a booklet consisting of: (1) a cover sheet explaining the purpose of the study, the participant's rights, and the name of a contact person and telephone number for those who might have questions after the test session was over; (2) the demographic questions; (3) the instruments described in this section. The purpose, task demands, and rights were explained orally as well as in printed form at the beginning of each session. Respondents were told that the questionnaire would include questions concerning attitudes, beliefs, and experiences they may have had. They were guaranteed anonymity and confidentiality of their responses and they were told that the session would take an hour or slightly more. In actuality, the range of time that it took students to finish was between 30 minutes to 1 hour. All students were asked to sign a written consent form before completing their questionnaires. Students were also given instructions on how to properly fill out three scantron sheets before they were left to fill out the questionnaire at their own pace.

The author has combined his informed consent forms and questionnaires into a single booklet that he can easily distribute. Doing so is quite common in descriptive research, especially with adult samples, and increases the efficiency of data collection.

The author is using computer technology (scantron sheets) in his data collection. Given the nature of his sample (college students) and his sample size (576), this approach is reasonable.

A debriefing form was given to each participant as they turned in their questionnaire. It explained the study in more detail and provided names and telephone numbers of local mental health services and community resources, such as services for battered women. Students that voluntarily participated in the study were offered extra credit points by their professors.

Given the sensitive nature of some questionnaire items, the debriefing that follows data collection appropriately includes information about community resources for individuals who have been victims of partner violence.

The initial sample consisted of 650 respondents of which 576 chose to complete the questionnaire. Of these, 33 questionnaires were omitted because they were illegible or partially completed. Finally, of the 543 remaining questionnaires, 348 were selected for this study because they met the criteria of having no missing data for any specific question, were either Mexican American/Mexican National or Non-Mexican White, and had been in a heterosexual romantic relationship for a month or longer during the previous 12 months.

Here the author describes his criteria for including completed questionnaires in his data set. In essence, he is addressing the issue of admissibility of the data *(see Chapter 4).*

NOTE: Excerpt is from *The Relation of Acculturation, Criminal History, and Social Integration of Mexican American and Non-Mexican Students to Assaults on Intimate Partners* (pp. 3–4, 14–20) by I. L. Ramirez, 2001, unpublished doctoral dissertation, University of New Hampshire, Durham. Reprinted with permission.

For Further Reading

Alreck, P. L., & Settle, R. B. (2003). *The survey research handbook* (3rd ed.). New York: McGraw-Hill.

Berdie, D. R., Anderson, J. F., & Niebuhr, M. A. (1986). *Questionnaires: Design and use* (2nd ed.). Metuchen, NJ: Scarecrow Press.

Bourque, L. B., & Fielder, E. P. (2002). *How to conduct telephone surveys* (2nd ed.). Thousand Oaks, CA: Sage.

Bourque, L. B., & Fielder, E. P. (Eds.). (2002). *How to conduct self-administered and mail surveys* (2nd ed.) Thousand Oaks, CA: Sage.

Carpenter, J. (2002). Measuring social capital: Adding field experimental methods to the analytical toolbox. In J. Isham, T. Kelly, & S. Ramaswamy (Eds.), *Social capital and economic development: Well-being in developing countries* (pp. 119–135). Cheltenham, England: Edward Elgar Publishing.

Daniel, J. (2011). *Sampling essentials: Practical guidelines for making sampling choices*. Thousand Oaks, CA: Sage.

Delandshere, G., & Petrosky, A. R. (1998). Assessment of complex performances: Limitations of key measurement assumptions. *Educational Researcher, 27*, 14–24.

Fink, A. (2009). *How to conduct surveys: A step-by-step guide* (4th ed.). Thousand Oaks, CA: Sage.

Fowler, F. J., Jr. (2008). *Survey research methods* (4th ed.). Thousand Oaks, CA: Sage.

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust Web-based studies? A comparative analysis of six preconceptions about Internet questionnaires. *American Psychologist, 59*, 93–104.

Gubrium, J. F., & Holstein, J. A. (Eds.). (2001). *Handbook of interview research: Context and method*. Thousand Oaks, CA: Sage.

Henry, G. T. (1990). *Practical sampling*. Thousand Oaks, CA: Sage.

Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., & Couper, M. (2004). Psychological research online: Report of Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet. *American Psychologist, 59*, 105–117.

Litwin, M. S. (1995). *How to measure survey reliability and validity*. Thousand Oaks, CA: Sage.

Magnusson, D., Bergman, L. R., Rudinger, G., & Torestad, B. (Eds.). (1991). *Problems and methods in longitudinal research: Stability and change*. New York: Cambridge University Press.

Menard, S. (Ed.). (2007). *Handbook of longitudinal research: Design, measurement, and analysis*. Burlington, MA: Academic Press.

Neuman, W. L. (2011). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Boston: Allyn & Bacon. (See Chapter 10, "Survey Research.")

Oppenheim, A. N. (1992). *Questionnaire design, interviewing, and attitude measurement*. New York: St. Martin's Press.

Panter, A. T., & Sterba, S. K. (Eds.). *Handbook of ethics in quantitative methodology*. New York: Routledge.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher, 38*, 109–119.

Rogelberg, S. G., & Luong, A. (1998). Nonresponse to mailed surveys: A review and guide. *Current Directions in Psychological Science, 7*, 60–65.

Rosenfeld, P. (1993). Computer-administered surveys in organizational settings: Alternatives, advantages, and applications. *American Behavioral Scientist, 36*, 485–511.

Schuman, H., & Presser, S. (1982). *Questions and answers: Experiments in the form, wording, and context of survey questions*. San Diego, CA: Academic Press.

Schwarz, P. N., & Sudman, S. (1996). *Answering questions: Methodology for determining cognitive and communicative processes in survey research*. San Francisco: Jossey-Bass.

Wright, J., & Marsden, P. (2010). *Handbook of survey research* (2nd ed.). Bingley, England: Emerald Group Publishing.