CHAPTER 14

Survey Research

Ask and you shall know.

INSTRUCTIONAL OBJECTIVES

After studying this chapter, the student will be able to:

1. State the purpose of survey research.
2. Classify the four categories of surveys according to their scope and focus.
3. Distinguish between longitudinal and cross-sectional surveys.
4. Describe the different types of longitudinal surveys.
5. List the steps involved in carrying out a survey.
6. Explain the importance of probability sampling in survey research.
7. Define margin of error and use sample data to calculate the margin of error and the confidence interval around the population parameter.
8. Calculate the sample size needed to achieve a desired margin of error in a survey.
9. List some of the factors that influence a researcher's decision about sample size.
10. State the merits and disadvantages of the interview as a data-gathering technique.
11. Describe data-gathering techniques used in survey research.
12. Define focus group.
13. Write both open-ended and closed-ended survey questions.
14. List guidelines for conducting an interview.
15. State the advantages and disadvantages of the questionnaire as a data-gathering technique.
16. State some rules for writing items for a questionnaire.
17. List guidelines to follow for developing the format of a questionnaire.
18. Explain the advantages of pilot testing a questionnaire.
19. Explain the follow-up procedures a researcher should employ with a mailed or an electronic survey.

20 Write a cover letter for a questionnaire.

21 Outline procedures for dealing with nonrespondents after follow-up procedures have been used.

22 Discuss the procedures for assessing the score validity and reliability of questionnaires.

23 Explain the data analyses that are appropriate for survey data.

24 Explain the statistics used with cross-tabulations.

25 Describe issues affiliated with electronic surveys.

In **survey research**, investigators ask questions about peoples' beliefs, opinions, characteristics, perceptions, and behavior. The survey questionnaire is used widely as a source of data in studies in sociology, business, psychology, political science, and education. It also provides information used for nonresearch planning and policy setting in the areas of government, business, health, and education. For example, the U.S. Census Bureau conducts a monthly survey for the Bureau of Labor Statistics that focuses on measuring labor force participation and unemployment. These data are used to produce the monthly unemployment figures for the nation as well as for the individual 50 states. Surveys are taken of consumer choices, use of health services, numbers of women and minority faculty in universities, and so on. A survey researcher may want to investigate associations between respondents' characteristics such as age, education, social class, race, and their current attitudes toward a particular issue. Typically, survey research does not make causal inferences, but rather describes the distributions of variables in a specified group.

Although researchers want to gather information about the characteristics of populations, they usually study a smaller group (i.e., a sample) carefully drawn from the population and then use the findings from the sample to make inferences about the population.

The range of topics covered by surveys, and surveying techniques, have increased significantly in the past several years. Hardly a week goes by that you are not exposed through the news media to the results of a survey. The well-known Gallup poll, for instance, surveys public opinion on a variety of issues. Market researchers ask what products you purchase or might purchase; political pollsters ask for whom you are likely to vote; television networks want to know what shows you watch.

Many doctoral dissertations, and a large volume of the research published in educational journals, involve survey methods. Public school districts also conduct surveys to gather data such as average teacher load, number of preschool children in the community, number of students who participate in extracurricular activities, opinions of parents and students, and data on myriad other topics.

THINK ABOUT IT 14.1

Which of the following questions would best be answered by survey methods?

1. Do voters in our school district think we should raise taxes in order to build new classrooms?
2. Do people who have taken driver education have fewer accidents than people who have not?
3. What do voters consider the most important issues in the upcoming election?
4. What do school principals consider to be the major problems in their schools?
5. Does dividing second-grade math students into ability groups produce greater math achievement than conducting math instruction in a single group?

Answers

Questions 1, 3, and 4 would be suitable for survey research. Ex post facto would be appropriate for question 2, and experimental research would be suitable for question 5.

TYPES OF SURVEYS

Before initiating survey research, the investigator must determine the most appropriate format for the proposed investigation. Surveys are classified according to their focus and scope (i.e., census and sample surveys) or according to the time frame for data collection (i.e., longitudinal and cross-sectional surveys). Becoming familiar with the options enables the researcher to select the method that will provide the most useful data.

SURVEYS CLASSIFIED ACCORDING TO FOCUS AND SCOPE

A survey that covers the entire population of interest is referred to as a **census**. An example is the U.S. Census undertaken by the federal government every 10 years. In research, however, "population" does not refer to all the people of a country; rather, **population** refers to the entire group of individuals to whom the findings of a study apply. The researcher defines the specific population of interest. It is often difficult, or even impossible, for researchers to study very large populations. Hence, they select a smaller portion, a **sample**, of the population for study. A survey that studies only a portion of the population is known as a **sample survey**.

Surveys may be confined to simple tabulations of **tangibles** such as what proportion of children ride school buses or average class enrollment. The most challenging type of survey is one that seeks to measure **intangibles** such as attitudes, opinions, values, or other psychological and sociological constructs. In such a study, you must bring to bear not only the skills involved in proper sampling, but also the skills involved in identifying or constructing appropriate measures and employing the scores on these measures to make meaningful statements about the constructs involved. If you classify surveys on the basis of their scope (i.e., census versus sample) and their focus (i.e., tangibles versus intangibles), four

categories emerge: (1) a **census of tangibles**, (2) a **census of intangibles**, (3) a **sample survey of tangibles**, and (4) a **sample survey of intangibles**. Each type has its own contributions to make—and its own inherent problems.

A Census of Tangibles

When you seek information about a small population, such as a single school, and when the variables involved are concrete, there is little challenge in finding the required answers. If a school principal wants to know how many children ride the school bus or how many teachers have master's degrees, a simple count will provide the needed information. Because the study covers the entire population, the principal can have all the confidence characteristic of perfect induction. Well-defined and unambiguous variables are being measured, and as long as the enumeration is accurate and honest the principal can say, without much fear of contradiction, "Sixty-five percent of the present faculty has master's degrees." The strength of a census of this type lies in its irrefutability. Its weakness lies in its confinement to a single limited population at a single point in time. The information provided by such a census may be of immediate importance to a limited group, but typically such surveys add little to the general body of knowledge in education.

A Census of Intangibles

Suppose the school principal now seeks information about student achievement or aspirations, teacher morale, or parents' attitudes toward school. The task will be more difficult because this census deals with constructs that are not directly observable but must be inferred from indirect measures. Test scores and responses to questionnaires serve to approximate constructs such as knowledge and attitudes. The National Study of School Evaluation publishes an opinion inventory designed to measure student, teacher, and parent attitudes and opinions about schools. Administering the inventory to all students, teachers, or parents in the school system would represent a census of intangibles.

Another example of this type of census is the achievement-testing program carried out by most schools. All children are tested, and the test scores are used to compare their performance with national norms, their own previous performance, and so on.

The value of a census of intangibles is largely a question of the extent to which the instruments used actually measure the constructs of interest. Reasonably good instruments are available for measuring aptitude and achievement in a variety of academic areas. Many other variables remain very difficult to measure. Because researchers lack instruments that can measure meaningfully the constructs involved, many important questions in education have not been answered satisfactorily. Such variables as teacher success, student motivation, psychological adjustment, and leadership have been difficult to define or operationalize.

A Sample Survey of Tangibles

The expense involved in carrying out a census that seeks information about large groups is often prohibitive. Therefore, researchers employ sampling, and when sampling is done well, the inferences made concerning the population can be quite formative in describing, exploring, or explaining an issue under study.

A classic example of a sample survey of tangibles is the report on equality of educational opportunities commissioned by the U.S. Department of Health, Education, and Welfare. This study—sometimes called the Coleman Report, after James Coleman who developed it—was conducted in response to Section 402 of the Civil Rights Act of 1964, which directed the Commissioner of Education to conduct a survey of inequalities in educational opportunities among various groups in the United States. The sample survey included more than 600,000 children in grades 1, 3, 6, 9, and 12, in approximately 4000 schools. The schools were considered generally representative of all U.S. public schools, although there was some intentional over-representation of schools with minority group populations. From the survey data the researchers concluded that 65 percent of African Americans attended schools in which more than 90 percent of students were African American, and 80 percent of Caucasians attended schools enrolling more than 90 percent Caucasian. When comparisons were made concerning class size, physical facilities, and teacher qualifications, relatively little difference was found among schools serving different racial and ethnic groups. However, these variables did differ between metropolitan and rural areas and between geographic regions. Students disadvantaged in regard to these variables appeared to be rural children and those in the South, regardless of race.

A Sample Survey of Intangibles

Public opinion polls are examples of studies measuring intangible constructs. Opinion is not directly observable, but must be inferred from responses made by the subjects to questionnaires or interviews. Opinion polling began in the 1930s and has expanded tremendously. When respondents have freely revealed their preferences before elections, for instance, pollsters have been quite accurate in inferring public opinion, from which they have predicted subsequent election results. These polls provide excellent examples of the usefulness of sample statistics in estimating population parameters. However, if people who support one candidate are reluctant to reveal their preference, whereas people who support the other candidate feel free to do so, considerable error is introduced into the poll results. For example, people are more willing to say they will vote against an incumbent than for him or her. A classic example occurred before the 1948 presidential election, when several polls showed Dewey leading the incumbent, Truman, but with many people indicating they were undecided. Newspapers had already printed headlines proclaiming Dewey the winner, but Truman won the election. Apparently, most of those who indicated they were undecided actually voted for Truman. Respondents are also reluctant to reveal a choice that may appear to be based on self-interest, prejudice, or lack of knowledge about issues.

How someone is going to vote is an intangible, but what he or she marks on a ballot is tangible. Cable network news stations such as MSNBC and CNN have done quite well in predicting how states will vote when only a few precincts have reported, because they can use tangible measures of a sample (i.e., how some ballots have been marked) to predict the vote of a population. Therefore, the risks are only those involved in estimating population parameters from sample statistics. However, pollsters who estimate how a population *will* vote

on the basis of how people *say* they will vote have the additional hindrance of measuring what is intangible at the time the measurements are made. Surveys of intangibles are limited by the fact that the data researchers collect are only indirectly measuring the variables of concern. The severity of this limitation depends on how well the observations measure the intangible variable.

The same survey may study tangibles and intangibles at the same time. The survey on equality of educational opportunities asked students to answer questionnaires and administered intelligence and achievement tests in order to make inferences about social class, ability, and achievement, as well as the relationship of these variables to each other and to tangible variables in the study.

THINK ABOUT IT 14.2

What type of survey is illustrated in the following examples?

1. Randomly selected teachers are asked how many years experience they have teaching.
2. The superintendent of School District 205 has Iowa Test of Basic Skills scores for all second-graders in his district.
3. The state superintendent of schools' list of enrollment in each of the state's 1,143 schools.
4. Some students at Jefferson School are given a physical fitness test in order to estimate the fitness of all the students in the school.

Answers

1. Sample survey of tangibles
2. Census of intangibles
3. Census of tangibles
4. Sample survey of intangibles

SURVEYS CLASSIFIED ACCORDING TO THE TIME DIMENSION

Surveys are also classified according to the duration of data collection: longitudinal surveys, which study changes across time, and cross-sectional surveys, which focus on a single point in time.

Longitudinal Surveys

Longitudinal surveys gather information at different points in time in order to study changes over an extended period. Three different designs are used in longitudinal survey research: panel studies, trend studies, and cohort research.

Panel Studies In a **panel study**, the *same* subjects are surveyed several times over an extended period of time. For example, a researcher studying the development of quantitative reasoning in elementary school children would select

a sample of second-graders and administer a measure of quantitative reasoning. This same group would be followed through successive grade levels and tested each year to assess how quantitative reasoning skills develop over time. Researchers have studied how age affects IQ by measuring the same individuals as adolescents, at college age, at middle age, and older. Because the same subjects are studied over time, researchers can see the changes in the individuals' behavior and investigate the reasons for the changes. For example, in Terman's (1926) classic panel study of intelligence, he followed exceptionally bright children to maturity.

Trend Studies A **trend study** differs from a panel study in that *different* individuals randomly selected from the same general population are surveyed at intervals over a period of time. For example, researchers who have studied national trends in mathematics achievement sample middle school students at various intervals and measure their math performance. Although the same individuals are not tested each time, if the samples from the population of middle school students are selected randomly, the results each time can be considered representative of the middle school population from which the student samples were drawn. Test scores from year to year are compared to determine if any trends are evident. Another example of a trend study is the survey on alcohol, tobacco, and other drug use among Indiana youth conducted annually since 1991 by the Indiana Prevention Resource Center. This continuing trend study permits researchers to evaluate the prevention and enforcement efforts directed at the teenage population and to plan future programs.

Cohort Studies In a **cohort study**, a *specific* population is followed over a length of time with different random samples studied at various points. Whereas trend studies sample a general population that changes in membership over time, a cohort study samples a specific population whose members do not change over the duration of the survey. Typically, a cohort group has age in common. For example, a school system might follow the high school graduating class(es) of 2012 over time and ask them questions about higher education, work experiences, attitudes, and so on. From a list of all the graduates, a random sample is drawn at different points in time and data are collected from that sample. Thus, the population remains the same during the study, but the individuals surveyed are different each time.

Cross-Sectional Surveys

Cross-sectional surveys study a cross section (sample) of a population at a single point in time. In a longitudinal study of vocabulary development, for example, a researcher would compare a measure of first-grade students' vocabulary skills in 2006 with one when they were fourth-grade students in 2009 and seventh-grade students in 2012. A cross-sectional study would compare the vocabulary skills of a sample of children from grades 1, 4, and 7 in 2012. The cross-sectional survey is the method of choice if you want to gather data at one point in time.

THINK ABOUT IT 14.3

How would you administer a questionnaire to assess changes in students' political attitudes during college with a (a) cross-sectional approach, (b) panel study, (c) trend study, and (d) cohort study?

Answers

a. In the cross-sectional study, you would select a random sample from each of the four levels and administer the questionnaire to them at the same time.
b. Panel, trend, and cohort studies are all longitudinal. In all three, you first randomly select a sample of freshmen from your population of interest. In a panel study, you assess your original sample and study the same individuals again when they are sophomores, juniors, and seniors.
c. In the trend study, you select a random sample of sophomores from the population. A year later, you select a random sample of juniors, and then in the final year you select a random sample of seniors.
d. The cohort study would differ from the trend study in that the subsequent samples are selected only from the population who were enrolled as freshmen when the study began and does not include students who transferred in later.

Longitudinal surveys are more time consuming and expensive to conduct because the researcher must keep up with the subjects and maintain their cooperation over a prolonged period. Cross-sectional surveys, in contrast, do not require years to complete. Hence, they are less expensive. A major disadvantage of the cross-sectional method is that chance differences between samples may bias the results. You may, by chance, select a sample of first-graders who are more mature than average and a sample of fourth-graders who are less mature than average, with the result that the difference between the groups appears much smaller than it is in reality. However, researchers can usually obtain larger samples for cross-sectional studies than for longitudinal studies, and these larger samples mitigate the problem of chance differences.

SURVEY TECHNIQUE

The survey permits you to gather information from a large sample of people relatively quickly and inexpensively. Conducting a good survey, however, is not as easy as it might initially appear. It requires careful planning, implementation, and analysis if it is to yield reliable and valid information.

SIX BASIC STEPS INVOLVED IN SURVEY RESEARCH

1. *Planning*. Survey research begins with a question that the researcher believes can be answered most appropriately by means of the survey method. For example, "How do elementary teachers feel about retaining students?" and "What is the extent of tobacco use among high school students in this district?" are questions that a survey could answer. The research question in survey research typically concerns the beliefs,

preferences, attitudes, or other self-reported behaviors of people (i.e., respondents) in the study.

2. *Defining the population*. One of the first important steps is to define the population under study. To whom will you distribute the survey? The population may be quite large, or it may be rather limited. For instance, the population might be all elementary teachers in the United States or all elementary teachers in the state of Hawaii. You might further restrict the population to all first-year male elementary teachers in the state of Hawaii. Defining the population is essential for selecting the appropriate subjects and determining to whom the results can be generalized.

   Once the population has been defined, the researcher must obtain or construct a complete list of all individuals in the population. This list, called the **sampling frame**, can be very difficult and time consuming to construct, although such a list may already be available.

3. *Sampling*. Because researchers generally cannot survey an entire population, they select a *sample* from that population. It is very important to select a *sample* that will provide results similar to those that would have been obtained if the entire population had been surveyed. In other words, the sample must be representative of the population. The extent to which this happens depends on the way subjects are selected. The sampling procedure that is most likely to yield a representative sample is some form of probability sampling (see Chapter 7). Probability sampling, based on the concept of random selection, permits you to estimate how far sample results are likely to deviate from the population values.

4. *Constructing the instrument*. A major task in survey research is constructing the instrument that will be used to gather the data from the sample. Two basic types of data-gathering instruments are interviews and questionnaires.

5. *Conducting the survey*. Once a data-gathering instrument is prepared, it must be field tested (i.e., in a pilot test) to determine if it will provide the desired data. Also included in this step are training the users of the instrument, interviewing subjects or distributing questionnaires to them, and verifying the accuracy of the data gathered.

6. *Processing the data*. The last step includes coding the data, statistical analysis, interpreting the results, and reporting the findings.

Many considerations are involved in implementing the foregoing steps. The balance of this chapter discusses these considerations in detail.

DATA-GATHERING TECHNIQUES

The two basic data-gathering techniques in survey research are interviews and questionnaires. Interviews involve some form of direct contact between the participants in the sample group and the interviewer (i.e., the researcher or someone trained by the researcher), who presents the questions to each person in the sample group and records their responses. A questionnaire is sent to all members of the sample group, who record and return their responses to the questions.

Personal Interviews

In a **personal interview**, the interviewer reads the questions to the participant in a face-to-face setting and records their responses. One of the most important aspects of the interview is its flexibility. The interviewer has the opportunity to observe the participant and the total situation in which he or she is responding. Questions can be repeated or their meanings explained in case they are not understood by the participants. The interviewer can also press for additional information when a response seems incomplete or not entirely relevant.

A greater response rate is another obvious advantage of the personal interview. The term **response rate** refers to the proportion of the selected sample that agrees to be interviewed or returns a completed questionnaire. With interviews, response rates are very high—perhaps 90 percent or better. Personal contact increases the likelihood that the individual will participate and will provide the desired information. With questionnaires, personal contact is missing, and participants are more likely to not complete and return the questionnaire. This results in many *nonresponses* and a lower rate of return. The low response rate typical for a mailed or electronic questionnaire (less than 30 percent is common) not only reduces the sample size, but also may bias the results (Fowler, 2002). However, an interviewer can receive responses to all or most of the questions. Missing data represent a serious problem for the questionnaire.

Another advantage is the control that the interviewer has over the order in which questions are considered. In some cases, it is very important that participants not know the nature of later questions because their responses to these questions may influence earlier responses. This problem is eliminated in an interview because the participant does not know what questions are imminent and cannot change answers previously given. In cases where participants have reading comprehension issues or there is a language barrier, interviews rather than questionnaires are the preferred information-gathering technique.

The main disadvantage of the personal interview is that it is more expensive than other survey methods. The selection and training of the interviewers, their salary, and their travel to the interview site or use of technology such as Adobe Connect for e-interviews, make this procedure costly. It takes a great deal of time to contact potential participants, set up appointments, and actually conduct the interview. Another disadvantage is the possibility of **interviewer bias**, which occurs when the interviewer's own feelings, attitudes, or gender, race, age, or other characteristics influence the way questions are posed and responses interpreted.

Another problem is **social desirability bias**, which leads participants to try to "please" the interviewer by providing socially acceptable responses that they would not necessarily give on an anonymous questionnaire. That is, participants may say what they think the interviewer wants to hear. For example, in preference polls in elections involving minority candidates, the proportion of respondents who said they would vote for the minority candidates was often higher than the proportion of votes these candidates actually received in the election.

Focus Groups

A specific category of interviews is the **focus group**. Several participants are interviewed at the same time. An advantage of a focus group is that participants respond not only to the researcher, but also to other participants and their responses. The interaction between participants usually reveals more about the subjects' point of view than would be the case with a researcher-dominated interview. The focus group's interaction enables the researcher to see how subjects incorporate the viewpoints of the others in structuring their own understandings.

This method can provide the researcher with insight into how disagreements are, or are not, resolved. Sometimes the researcher can report a final consensus.

Focus groups are often used in qualitative research. The researcher invites participants who are interested in the same general topic to assemble to discuss it. They are assured that they will be free to express themselves in their own words and to respond not only to the researcher, but also to other participants and their responses.

In quantitative research, it is very difficult and often very expensive to assemble individuals who will gather in the same place at the same time to respond to the quantitative researchers' predetermined questions. Among those willing to bear the expense of assembling people to focus on a predetermined topic are manufacturers of consumer products who want to pinpoint the product characteristics that lead to high sales. Also, politicians often use focus groups to determine positions that might help them become elected or to be re-elected.

Telephone Interviews

The telephone interview is popular, and studies show that it compares quite favorably with face-to-face interviewing. In fact, the past 50 years have seen a gradual replacement of face-to-face interviewing with telephone interviewing as the dominant mode of survey data collection in the United States (Holbrook, Green, & Krosnick, 2003). Its major advantages are lower cost and faster completion, with relatively high response rates. The average response rate may reach 80 percent or higher (Neuman & Kreuger, 2003). Telephone interviews can be conducted over a relatively short time span with people scattered over a large geographic area. For example, national polling organizations often use the telephone to obtain nationwide opinions among voters near election time. Both landline phones and cell phones permit a survey to reach people who would not open their doors to an interviewer, but who might be willing to talk. Another advantage is that respondents have a greater feeling of anonymity—and hence there may be less interviewer bias and less social desirability bias than with personal interviews.

The main disadvantage of the telephone interview is that there is less opportunity for establishing rapport with the respondent than in a face-to-face situation. It takes a great deal of skill to carry out a telephone interview so that valid results are obtained. The interviewer often finds it difficult to overcome the suspicions of the surprised respondents. We recommend that you identify yourself and explain that you are doing a survey for educational purposes only.

An advance letter that informs the potential respondents of the approaching call is sometimes used to deal with this problem, but the letter can induce another problem: The recipient has time to think about responses or prepare a refusal to participate when the call comes.

Another limitation of telephone interviews is that complex questions are sometimes difficult for respondents to follow. The interviewer may not realize that they misunderstand the questions. It is best that interview questions be concise, with a limited number of options. Furthermore, the phenomenon of multitasking may affect the quality of telephone interviews. Without the interviewer's knowing, the respondent may be watching television or checking their iPad while answering the survey questions. Telephone interviews can be very time consuming. If the sample is very large, a researcher will need a number of people to help with the interviews. We recommend that telephone interviews be relatively brief.

Another disadvantage is that households without telephones, those with unlisted numbers, and those with only cell phones may be difficult to include in a survey, which may bias results. Neuman and Kreuger (2003) state that approximately 95 percent of the population can be reached by telephone. A technique known as "random-digit dialing" solves the problem of unlisted numbers (although it does not reach households without a telephone). In random-digit dialing, a computer randomly generates a list of telephone numbers based on all possible numbers thought to be in use in an area. Because of the random determination, this technique ensures that every household with telephone service has an equal chance of being included in the sample. Random-digit dialing has greatly improved sampling in telephone surveys. See Fowler (2002) for a thorough discussion of random-digit dialing.

Other limitations of telephone surveys arise from new technology that may make it increasingly difficult to reach potential respondents by phone. Services such as caller identification, phone number blocking, and similar options enable residential phone customers to have much greater control over incoming calls. People may simply ignore calls from the unfamiliar number of the surveyor, and telephone response rates may continue to drop.

Computer-Assisted Telephone Interviewing (CATI)

Computer and telecommunications technology has been applied to telephone surveys. Wearing earphones, the interviewer sits at a computer while it dials a randomly selected telephone number (i.e., through random-digit dialing or from a database). When the respondent answers, the interviewer reads the first question that appears on the computer screen and types the answer directly into the computer. The computer program displays the next screen containing the next question, and so on through the entire survey. Using CATI saves a great deal of time. The surveyor can fill in forms on a computer screen or type answers to open-ended questions very quickly. The major advantage is that CATI software immediately formats responses into a data file as they are keyed in, which saves the researcher time usually spent in coding and manually transferring responses from paper into the computer for analysis.

Conducting the Interview

Whether the interview is conducted in person or by telephone, your main job is to ask the questions in such a way as to obtain valid responses and to record the responses accurately and completely. Your initial task is to create an atmosphere that will put the respondent at ease. After introducing yourself, briefly state the purpose of the interview, but avoid giving too much information about the study, which could bias the respondent. It is well to begin the interview with fairly simple, nonthreatening questions.

You also have the responsibility of keeping the respondent's attention focused on the task and keeping the interview moving along smoothly. This can best be done if you are thoroughly familiar with the questions and their sequence, so that you can ask the questions in a conversational tone without constantly pausing to await the next question. Of course, you must refrain from expressing approval, surprise, or shock at any of the respondent's answers.

Interviews can be more or less structured. In a less-structured interview, the same questions are asked of all respondents but the interview is more conversational, and the interviewer has more freedom to arrange the order of the questions or to rephrase them. If comparable data are to be obtained, however, the interviewer must standardize the procedure by using a structured interview schedule. A structured interview schedule contains specific questions in a fixed order, to be asked of all respondents, along with transition phrases and **probes** (questions used to clarify a response or that push a little further into a topic). For example, if the respondent starts to hedge, digress, or give irrelevant responses, or if he or she has obviously misinterpreted the question, then the interviewer may use a fixed probe such as "Explain your answer a little further" or "Can you tell me a little more about that?" Another important technique besides the probe is the **pause**. A good interviewer needs listening skills and must at times be quiet until the respondent answers. In less-structured interviews, any marked deviations from the protocol should be documented so that the information can be taken into account when analyzing the participant's response. In using probes, take care not to suggest or give hints about possible responses. It takes less training time to teach interviewers to administer a structured interview than it does an unstructured one because everything they need to say or do is contained in the interview schedule. For this reason, the structured interview is the most widely used format for large studies with numerous interviewers.

Training the Interviewer

It is essential that potential interviewers receive training before being assigned to conduct interviews. Quality of interviewers is probably one of the least appreciated aspects of survey research. "Interviewers have a great deal of potential for influencing the quality of the data they collect" (Fowler, 2002, p.117).

Certain aspects of interviews need to be standardized and therefore should always be included in interviewer training: (1) procedures for contacting respondents and introducing the study, (2) instructions on asking questions so that interviewers ask all questions in a consistent and standardized way, (3) procedures for probing inadequate answers in a nondirective way,

(4) procedures for recording answers to open-ended and closed-ended questions, and (5) rules for handling the interpersonal aspects of the interview in a nonbiasing way (Fowler, 2002). To be able to answer respondents' questions, interviewers should also know the purpose of the project, its sponsorship, the sampling approach used, and the steps that will be taken with respect to confidentiality.

Interviewer trainees should be provided with written manuals on interviewing procedures. They should observe interviews being conducted by trained individuals, and they should be supervised when conducting practice interviews. In practice interviews, the interviewees should be individuals drawn from the same population that will be used in the research project.

Mailed Questionnaires

The direct one-on-one contact with subjects in a personal interview is time consuming and expensive. Often, much of the same information can be obtained by means of a **questionnaire** mailed to each individual in the sample, with a request that he or she complete and return it by a given date. Because the questionnaire is mailed, it is possible to include a larger number of subjects as well as subjects in more diverse locations than is practical with an interview.

A **mailed questionnaire** has the advantage of guaranteeing confidentiality or anonymity, thus perhaps eliciting more truthful responses than would be obtained in a personal interview. In an interview, participants may be reluctant to express unpopular or politically incorrect points of view or to give information they think might be used against them later. The mailed questionnaire also eliminates the problem of interviewer bias.

A disadvantage of the mailed questionnaire is the possibility that respondents will misinterpret the questions. It is extremely difficult to formulate a series of questions whose meanings are crystal clear to every participant. The investigator may know exactly what is meant by a question, but because of poor wording or different meanings of terms the respondent makes a significantly different interpretation.

Another important limitation of mailed questionnaires is the low return rate. It is easy for the individual who receives a questionnaire to lay it aside and simply forget to complete and return it. A low response rate limits the generalizability of the results of a questionnaire study. It cannot be assumed that nonresponse is randomly distributed throughout a group. Studies have shown that there are usually systematic differences in the characteristics of respondents and non-respondents to questionnaire studies. Response rate is often higher among the better educated, the more conscientious, and those more interested in or generally more favorable to the issue involved in the questionnaires. The goal in a questionnaire study is a 100 percent return rate, although a more reasonable expectation may be a 40 percent to 75 percent return rate.

A number of factors have been found to influence the rate of returns for a mailed questionnaire, including (1) length of the questionnaire, (2) cover letter, (3) sponsorship of the questionnaire, (4) attractiveness of the questionnaire, (5) ease of completing it and mailing it back, (6) interest in the content, (7) use of a monetary incentive, and (8) follow-up procedures. We discuss these factors in detail later.

Electronic Mail Surveys

Researchers have begun to use e-mail to deliver questionnaires. Dillman (2007) found that e-mail surveys have the advantage of prompter returns, lower item nonresponse, and more complete answers to open-ended questions. **Electronic** mail surveys can be completed at the respondent's pace. Research shows that some of the factors found to be important for regular mail surveys are also important for e-mail surveys. For example, people who received a prior e-mail notification about the survey were more likely to respond; also, surveys addressed individually to a person (rather than being part of a mailing list) had higher response rates.

The main disadvantage is that they are appropriate only when the researcher has e-mail addresses for all members of a finite population such as all elementary teachers in a given school district or all members of a local union or fraternal lodge.

Internet Surveys

The Internet has become a popular platform for survey research. The questionnaire is placed on a web site constructed by the investigator or housed on a paid survey-related site such as Survey Monkey. Respondents answer the questions and submit the questionnaire online. **Web-based surveys** have a number of advantages. They have the potential of reaching large populations and permit the collection of larger amounts of data than is possible with traditional survey methods. They can be conducted quickly and easily and are less expensive than mailed surveys. The cost advantage increases as the size of the sample increases.

Another important benefit is in the processing of survey data. Web-based surveys can significantly reduce the amount of time, effort, and cost associated with getting the data into a system for analysis. Furthermore, because they are available 24 hours a day, respondents can reply when and where they choose. In a University of Colorado survey, 55 percent of respondents cited ease of use as one of the things they liked most about answering a web survey (Cook, Heath, & Thompson, 2000).

A limitation of Internet surveys is that samples are restricted to those with access to the technology and who choose to respond. The multitude of potential responses in a web survey does not overcome the problem of sampling error due to lack of representativeness. You need to be sure that the survey is reaching the desired respondents.

To generate enthusiasm and maximize response, you should send an introductory letter separate from the instrument that explains what the survey is about, requests cooperation, and provides an incentive for completing the survey. The problem, however, is that the introductory letter is usually sent by e-mail and may be deleted before it is ever read by the potential respondent.

Meister and Melnick (2003) conducted an Internet survey of first- and second-year teachers to identify their concerns in the areas of classroom management, time management, communication with parents, and academic preparation. Ten flyers, each containing a request for new teachers to visit the web site and respond to the survey, were sent to 1000 principals, located in all 50 states. They were asked to distribute the flyers to first- and second-year

teachers. A second mailing, via e-mail, was sent to 500 principals. They were asked to forward the e-mail to any first- or second-year teachers in their institution. A total of 273 teachers from 41 states responded. Sixty percent of the respondents were elementary teachers, 27 percent high school, and 13 percent middle school. Rural, suburban, and urban school districts were all represented. The responses were recorded in a database as teachers answered items on the web site.

The findings showed that new teachers were less confident of their knowledge and skills in the areas of discipline, time management, and communication. They reported needing assistance in handling disruptive students and those with special needs. Eighty-four percent reported that they felt sometimes "overwhelmed" by the paperwork and other noninstructional demands on their time. One in four teachers did not feel well prepared by their student teaching experience, especially in the areas of reading and language arts. Forty percent of beginning elementary teachers responded that they were not prepared to teach reading.

This research example illustrates the advantages and disadvantages of Internet surveys. The survey had the potential of reaching teachers over a wide geographic area and permitted them to respond at their convenience. The fact that the data were put into the system ready for analysis saved time and effort for the researchers. The disadvantage is that the response rate appears to be low, but actually we cannot calculate the response in the usual way. There is no way of knowing how many teachers actually received a flyer or the e-mail message about the survey. DeVaus (2002) discusses a variety of applications for web-based surveys, as well as tips on designing the questionnaire.

Directly Administered Questionnaires

A **directly administered questionnaire** is given to a select group assembled at a specific place for a specific purpose. Examples include surveying freshmen or their parents attending summer orientation at a university. Surveys at universities are often administered in classrooms or in residence halls.

The main advantage of direct administration is the high response rate, which typically is close to 100 percent. Other advantages are the low cost and the fact that the researcher is present to provide assistance or answer questions. The disadvantage is that the researcher is usually restricted in terms of where and when the questionnaire can be administered. Also, when a population is limited or is selected for convenience (e.g., parents of freshmen in a specific university), the results of the survey will be equally limited in terms of generalizability.

As we have seen, several data-collection methods exist for survey research, and researchers choose the one best suited for their particular study and research questions posed. However, a multimode approach in which researchers use combinations of methods in the same study is quite common. In fact, Fowler (2002) states that mixing modes is one of the best ways to minimize survey nonresponse because it enables researchers to reach people who are inaccessible via a single mode. Table 14.1 provides a summary of the advantages and disadvantages of the various data-collection methods.

Table 14.1 Comparison of Data-Collection Methods

Method	Advantages	Disadvantages
Personal interview	Researcher is present Can establish rapport Flexibility High response rate Fewer incomplete answers Good for surveys on complicated issues Time for thoughtful answers	Time-consuming Expensive Interviewer bias Social desirability bias Need trained interviewers No anonymity Less safe for researchers
Telephone interview	Low cost Convenient High response rate Quick Greater safety for interviewers and respondents	Less rapport Individuals may refuse to talk Questions need to be short and simple Limited to those with telephones
Mailed questionnaire	Low cost Allows anonymity No interviewer bias Convenient Nonthreatening	Slow Low response rate and nonresponse bias Cannot clarify question Missing data Literacy required
Electronic or Internet questionnaires	Low cost Easy for researcher Convenient for respondent Allows anonymity Potential of quick response Time for thoughtful answers Low social desirability bias Greater safety for researcher and respondent	Limited to computer savvy Difficulty getting cooperation Potentially lower response rate Literacy requirement May not be able to identify respondents
Directly administered questionnaire	Low cost High response rate Researcher present	Not as flexible with respect to time and place

STANDARD ERROR OF THE SAMPLING PROPORTION

We have seen in Chapter 7 that, even with random sampling, there will always be some error in estimating a population parameter from sample statistics. The statistic most commonly reported in a sample survey is a proportion or a percentage of the sample that gives a particular response. The discrepancy between the known sample proportion and the unknown population value is referred to as **sampling error**. The first step in assessing how much sample results are likely to deviate from the population values is to calculate the standard error of the sampling proportion.

CALCULATING THE STANDARD ERROR

Using the obtained sample proportions, you can calculate the variance and then the standard error of the proportion. The variance of a proportion is pq, where p is the proportion agreeing or having a certain characteristic, and $q(1 - p)$ is the proportion not agreeing or not having the characteristic. If the

variance for the proportion is *pq,* the formula for the standard error of the proportion is

$$\text{SE}(\sigma_p) = \sqrt{\frac{\text{Var}}{n}} = \sqrt{\frac{pq}{n}} \tag{14.1}$$

where

$\text{SE or } \sigma_p$ = standard error of the proportion
p = proportion agreeing
q = proportion not agreeing $(1 - p)$
n = size of the sample

For example, assume you survey a random sample of 100 parents of prekindergarten children in a school district and ask if they are in favor of full-day kindergarten in the district. You find that 80 say "Yes" to this question and 20 say "No." The standard error of the proportion is

$$\text{SE} = \sigma_p = \sqrt{\frac{(.80)(.20)}{\sqrt{100}}} = \sqrt{\frac{.1600}{100}} = \sqrt{.0016} = .04$$

From probability theory, we say that we are 95 percent confident that the sample statistic is within 1.96 standard errors of the parameter.

In our example, standard error of .04 multiplied by 1.96 gives us .0784. Rounding .0784 to .08 and expressing it as a percent, we arrive at 8 percent. Thus, we would report that 80 percent of parents of prekindergarten children are in favor of a full day with a **margin of error** of 8 percent. This means that out of 100 trials, 95 would contain the population proportion, which is thought to be between 72 and 88 percent. Walker (2005) has an extensive discussion pertaining to the standard error of the proportion and its uses in educational research and, more generally, in social science research.

THINK ABOUT IT 14.4

A community survey asked the question, "Are you in favor of having a Wal-Mart in our community?" From the 899 completed surveys, it was found that 63 percent of the residents answered "Yes" to this question. Calculate the margin of error and set up the 95 percent confidence interval for these data.

Answer

$$\sigma_p = \sqrt{\frac{(.63)(.37)}{899}} = \sqrt{\frac{.2331}{899}} = \sqrt{.000259} = .0161$$

$1.96 \times 0161 = 0316$, or a 3 percent margin of error. The confidence interval would be 63 ± 3. We would be 95 percent confident that the percent in the population favoring a Wal-Mart in the community is between 60 and 66 percent. In surveys conducted by the major polling organizations, a margin of error of 3 percent is generally considered acceptable.

CALCULATING SAMPLE SIZE

Early on, researchers must decide how large a sample to draw. We know that sampling error decreases as the size of the sample increases. Once you decide on an acceptable margin of error, you can determine the sample size needed to achieve that margin of error by applying Formula 14.2.

$$n = \left(\frac{\frac{1}{E}}{\sqrt{pq}} \right)^2 (z)^2 \qquad (14.2)$$

where

n = sample size needed
E = desired margin of error
pq = variance of hypothesized proportion
z = score of confidence level

For example, if we want a ±3 percent margin of error at the 95 percent confidence level ($z = 1.96$) for a hypothesized $p = .5$, $q = .5$, the required n would be

$$n = \left(\frac{\frac{1}{.03}}{\sqrt{(.5)(.5)}} \right)^2 (1.96)^2 = \left(\frac{1}{.06} \right)^2 (3.8416) = (277.7783)\ (3.8416)$$

$$n = 1067.1131$$

If the hypothesized population proportions are .50/.50, then 1068 subjects would be needed to yield a margin of error of 3 percent. Remember that this is 1068 usable responses. If you are using a mailed questionnaire and anticipate a 50 percent return, you would need to mail out twice that number of questionnaires. If a 10 percent margin of error would be acceptable, then only 97 respondents would be needed.

$$n = \left(\frac{\frac{1}{10}}{\sqrt{(.5)(.5)}} \right)^2 (1.96)^2 = \left(\frac{1}{.2} \right)^2 (3.8416) = (25)(3.8416) = 96.04$$

Using $p = .50$ is a typical, safe way to calculate the needed sample size because pq is at its maximum possible value when p and q both equal .50 and $pq = .25$. Using population estimates of $p = q = .50$ in Formula 14.2, you can be confident that the margin of error will be as small as, or smaller than, the value specified. If there are compelling reasons, based on theory or the scholarly literature, to hypothesize unequal population proportions for p and q, these may be substituted for .50 and .50, respectively, and the result will be a lower required number of subjects. For instance, if 20 percent of the high school boys in a school district tried out for intramural sports last year, you might use $p = .20$ and $q = .80$ in calculating the number of subjects needed for a survey of this year's boys. Note that all these statistics are based on the assumption that the sample is a random sample of the population of interest. The credibility of the results is a function of the validity of that assumption. Table 14.2 shows the sample size needed to have a certain margin of error for a given p. Note that the largest sample is required when the desired margin of

Table 14.2 Minimum Sample Sizes Required for Various Margins of Error around the Parameter Estimation at the .95 Confidence Level

Maximum Margin of Error (%)	Value of p		
	.10 or .90[a]	.25 or .75	.50
1	3462	7212	9616
2	866	1803	2404
3	385	802	1069
5	139	289	385
10	35	73	97

[a]The sample variance, pq, is the same when $pq = (.10)(.90)$ as when $pq = (.90)(.10)$.

error is lowest and $p = .50$. As the acceptable margin of error increases and p varies from .50, the required sample size decreases.

As with all sampling, the representativeness of the sample is more important than its size. An unrepresentative sample leads to inappropriate conclusions regardless of its size. For example, a university wanted to determine the need for evening computer lab hours. It conducted a random survey of all full-time students enrolled at the university. However, the majority of the students taking evening classes and, thus, potentially more likely to use the computer lab in the evening were part-time, not full-time, students. A survey of full-time students could lead to inaccurate conclusions about the demand for evening computer lab hours.

Another factor influencing the decision about sample size is the variability of the population to be sampled. If the population is fairly homogeneous, then you can use a smaller sample than if the population is more heterogeneous A population of college students, for example, would be expected to show less variability than a population of adults in general, and, thus, a smaller sample could be used. Check published research studies to learn about the variability of population of interest and the sample size used in survey research on the population.

Sample size does not need to be a certain percentage of the population. When random sampling is used, a sample size that is only a small percentage of the population can represent the population well. The major public opinion polls, for example, do not use very large samples. The Nielsen rating service uses a sample of only approximately 5000 television-watching households in the United States to estimate the popularity of various programs. The Nielsen ratings are a major variable in determining which programs continue and which are cancelled.

CONSTRUCTING THE INSTRUMENT

FORMAT OF QUESTIONS

Once the overall research question has been determined, the next task is to construct an instrument that will provide the desired information. Because survey data consist of participants' responses to individual questions, it is essential to start with good questions. Two basic types of questions are used in survey instruments: closed-ended, or fixed alternative, and open-ended or free-response questions. Use **closed-ended questions** when all the possible relevant responses to a question

can be specified and the number of possible responses is limited. For example, in a survey of undergraduates, a question about class level would be closed-ended. The possible answers are known and are few in number: freshman, sophomore, junior, or senior. Other closed-ended questions might ask about residency status (in-state or out-of-state) or gender. **Open-ended questions** are used when there are a great number of possible responses or when the researcher cannot predict all of the possible responses. For example, a question about students' reasons for selecting a particular university would probably be open-ended. A question about students' college major would be open-ended because the researcher would probably not want to include the long list of possible majors. To be sure, both formats can be used in the same question—that is, a number of closed-ended responses can be followed by "Other" as the last possible response. However, the use of "Other" can make analysis more difficult because it allows participants the opportunity to deviate from discrete categories and provide numerous varied responses. For example, you might ask a professor the question:

1. What type of writing assignments do you typically require in your course? (circle as many as apply)
   a. Reports
   b. Themes or essays
   c. Research papers
   d. Take-home essay examinations
   e. Minute papers
   f. Other (please specify)

There are advantages and disadvantages to both question formats. The open-ended question permits a free response rather than restricting the respondent to a choice from among stated alternatives. Individuals are free to respond from their own frame of reference. Open-ended questions are easier to construct, but analyzing them is very tedious and time consuming. You must read and interpret each response and then develop a coding system that will permit a quantitative analysis of the responses. Some responses may be unclear, and you may be unsure how to classify or code the response. The responses to open-ended questions typically differ in length, and some respondents may give more than one response to a particular question. If asked the most important reason for choosing a certain university, a respondent might answer, "I chose University X because of its academic reputation and because it is in-state and less expensive for me." You must then decide whether to use either both answers or only the first one, assuming the first answer is the more important one. An option to consider with such a question is rank ordering using a combination of closed- and open-ended responses.

Closed-ended questions take more time to construct, but the responses are easier to tabulate. Responses to closed-ended questions can be coded directly into software spreadsheets, such as SPSS or Excel for analysis. Closed-ended questions can be answered more easily and quickly by respondents. A closed format also ensures that all subjects have the same frame of reference in responding and may also make it easier for subjects to respond to questions related to sensitive or private topics.

A limitation of the closed-ended question is that it provides little insight into whether respondents really have any information or any clearly formulated opinions about an issue. It is easier for the uninformed respondent to choose one of the suggested answers than to admit to lack of knowledge on an issue. For example, in response to the question "What is the effect of outsourcing on the U.S. economy?," the respondent who has little knowledge of outsourcing and the reasons for it could easily select a reasonable response from among the alternatives provided. In contrast, knowledgeable respondents or participants who possess well-informed opinions on the issue may dislike being restricted to simple response categories that do not permit them to qualify their answers. It is possible to obtain the benefits of both open-ended and closed-ended questions. A researcher can first use the open-ended format with a small sample to identify possible alternative responses to questions.

STRUCTURE OF QUESTIONS

We have discussed the advantages and disadvantages of closed-ended and open-ended questions in survey research. Now let us look at the ways these two formats can be used to structure questions for interviews and surveys.

1. **Completion**, or **fill-in, items** are open-ended questions to which respondents must supply their own responses in their own words. For example, "What is the major weakness you have observed in your students' preparation for college?"
2. **Checklists** are questions that present a number of possible responses respondents are asked to check those that apply. For example,

   What type of teaching aids do you use in your classes? (check as many as apply)

   _____ 1) CHALKBOARD
   _____ 2) OVERHEAD PROJECTOR
   _____ 3) COMPUTER PROJECTOR
   _____ 4) DVDs
   _____ 5) SMARTBOARD
   _____ 6) OTHER (please specify)

3. **Scaled items** ask respondents to rate a concept, event, or situation on such dimensions as quantity or intensity, indicating "how much"; on quality, indicating "how well"; or on frequency, indicating "how often." For example,

   How would you rate the writing skills of students you are teaching this semester? (check one)

   _____ 1) VERY POOR
   _____ 2) LESS THAN ADEQUATE
   _____ 3) ADEQUATE
   _____ 4) MORE THAN ADEQUATE
   _____ 5) EXCELLENT
   _____ 6) INSUFFICIENT INFORMATION

   or

How well prepared in basic math skills are the students who typically enroll in your course? (check one)

_____ 1) NOT AT ALL PREPARED
_____ 2) SOMEWHAT PREPARED
_____ 3) WELL PREPARED
_____ 4) EXTREMELY WELL PREPARED

4. **Ranking items** ask respondents to indicate their order of preference among a number of options. Rankings should involve no more than six options; otherwise it becomes too difficult for respondents to make item comparisons independently and not provide ties to certain items. An example of a ranking item follows:

   Please rank the order of difficulty your students have in reading each of the following materials, with 1 being the most difficult and 4 the least difficult.

   _____ 1) TEXTBOOKS
   _____ 2) OTHER REFERENCE BOOKS
   _____ 3) JOURNAL ARTICLES
   _____ 4) TECHNICAL REPORTS

5. **Likert-type items** let subjects indicate their responses to selected statements on a continuum from *strongly disagree* to *strongly agree*, as seen in Figure 14.1. An advantage of this type of item is that points can be assigned to the various responses and, thus, measures of central tendency, variability, correlation, and the like can be calculated. For example:

   Students who typically enroll in my course are prepared in basic math skills: (circle one)

   Strongly Disagree Disagree Agree Strongly Agree

THINK ABOUT IT 14.5

Assume you are conducting a survey of students in the local high school on the extent of alcohol and drug use among students in the school. Construct one item of each of the following types to include in this survey: (1) completion, (2) checklist, (3) scaled item, (4) ranking, and (5) Likert-type.

Suggested Answer:

Completion
How would you describe the extent of alcohol use in your school?

Checklist
Are you (check one)
_____ 1) Male
_____ 2) Female

Scaled item
On average, how often do the students that you know use any form of alcohol? (check one)
_____ 1) Twice a week
_____ 2) Once a week
_____ 3) Twice a month
_____ 4) Once a month

Ranking

How would you rank the following in terms of extent of use among students you know? (1 most common and 5 least common)

_____ 1) Beer

_____ 2) Wine

_____ 3) Liquor

_____ 4) Marijuana

_____ 5) Hard drugs

Likert-type item

Alcohol consumption is a serious problem among students in this school. (circle one)

Strongly Disagree Disagree Agree Strongly Agree

FIELD TESTING

Now let us examine the task of formulating good questions. Deciding how to word survey questions is a challenge. Before beginning to write a structured set of survey questions, it can be helpful to have a focus group—known in this instance as a pilot study—discuss the questions in a nonstructured form. You bring together a group of participants, possibly 10 percent of the sample size, representative of the study's population, to discuss the topics covered in the survey. A moderator keeps the discussion focused on a preset agenda and asks questions to clarify comments. Note that the pilot group is not used in the final sample to avoid bias from having been exposed previously to the survey questions.

A pilot study can be very valuable in questionnaire development. Discussions help the researcher understand how people talk about the survey issues, which aids in choosing vocabulary and phrasing questions. A pilot study can often suggest issues, concerns, or points of view about the topic that the researcher had not considered.

Respondents answer the questions and provide feedback to the researcher on any difficulties they have with the items. Pay attention to such comments as "I don't know what you mean here" and "More than one of these answers apply to me." Try to ascertain whether all respondents interpret the questions in the same way. See if different responses are given to different versions of the same question. Observations of the respondents as they fill out the questionnaire can also be enlightening. Spending an undue amount of time on a question or leaving a question blank and returning to it later can be clues that there are problems with some items.

The results of field tests can be used to clarify the items or perhaps to eliminate some. It is especially important to determine whether the questions will operate equally well for the different social classes and ethnic groups of the population to be studied. The following specific issues should be addressed by field testing:

1. Do the respondents seem comfortable with the questionnaire and motivated to complete it?
2. Are certain items confusing?
3. Could some items result in hostility or embarrassment to respondents?
4. Are the instructions clear?

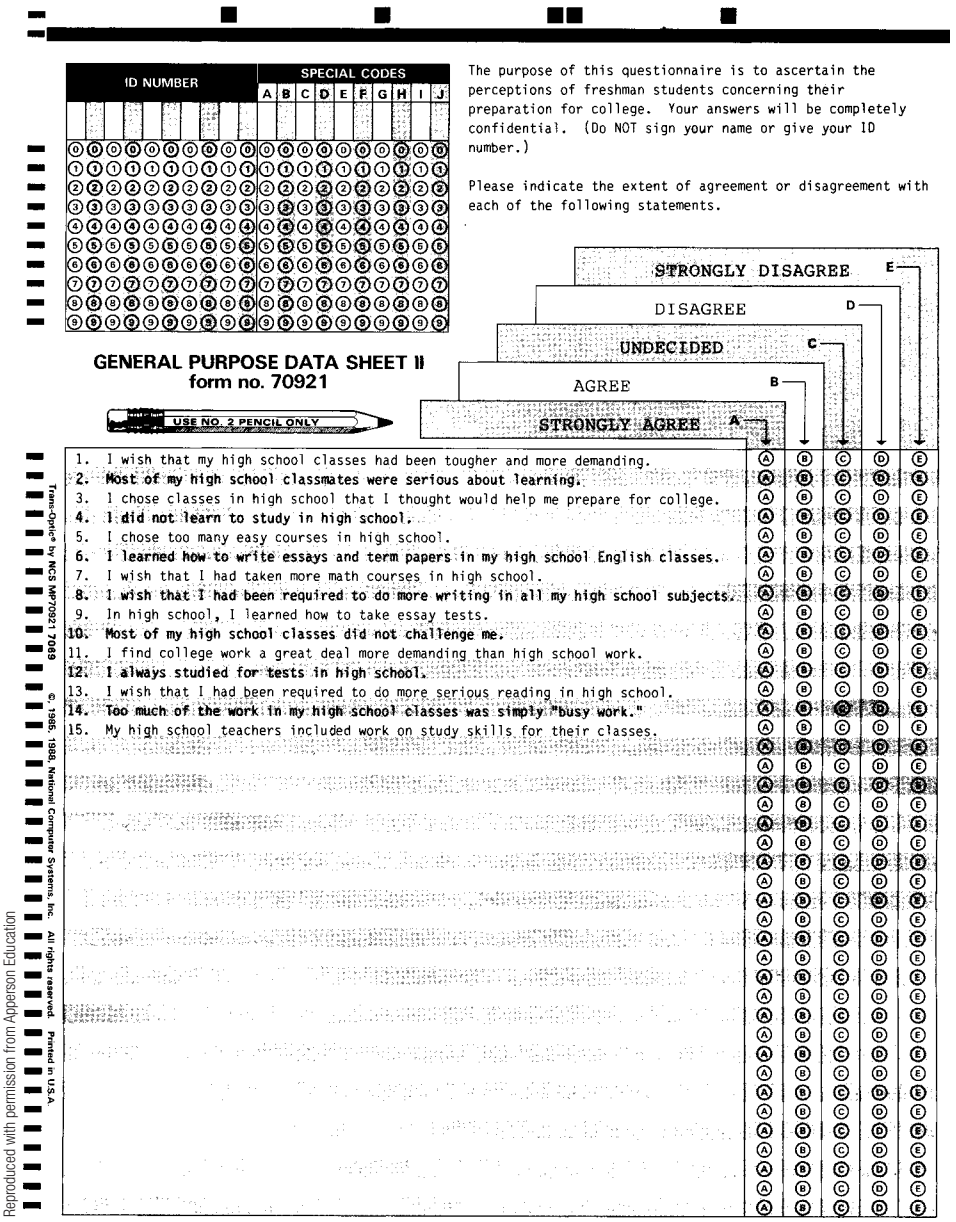


Figure 14.1 Example of a Scannable Sheet with Survey Statements

Source: Reproduced with permission from Apperson Education

5. How long will it take a respondent to complete the questionnaire?
6. Do all respondents interpret the items in the same way?

WRITING SURVEY QUESTIONS

A significant body of research has shown that changes in phrasing, the amount of information offered, and the choice of responses available to participants can influence the outcome of a survey. The way you ask a question may prescribe the response. It is possible for survey researchers with competing agendas to

come up with entirely different responses. Consider the following three ways to ask a question:

1. Would you support increased taxes to pay for full-day kindergarten?
2. Would you support an increase in your property taxes to pay for full-day kindergarten?
3. Would you support a 5 percent increase in your property taxes to pay for full-day kindergarten?

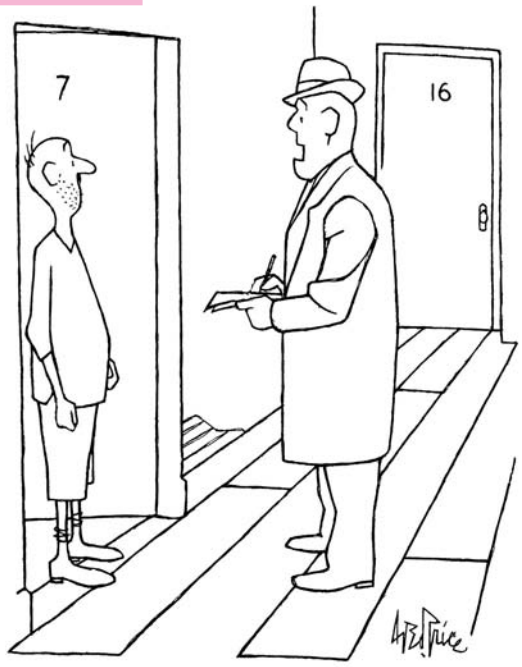
Individuals who might agree with the first general question would not necessarily agree when they consider that their property taxes would be increased. The following are basic guidelines for writing good questions:

1. *Questions should be short, simple, and direct*. Eliminate any words and phrases not essential to the clear meaning of the question. Short questions are easier to understand.
2. *Phrase questions so that they can be understood by every respondent*. The vocabulary should be nontechnical and written for participants with basic reading levels. Hence, a pilot study is important because similar participants can read the survey questions and give their interpretation of the content of each question. For example, questions using terms such as "authentic assessment" or "total quality management" may not be appropriate in a survey designed for working adults without a high school diploma. Also, be careful not to use slang, abbreviations, or acronyms that may not be familiar to all.
3. *Phrase questions so as to elicit unambiguous answers*. The question, "Did you vote in the last election?" is ambiguous because it does not specify which election. Quantify responses whenever possible. Words such as *often* and *sometimes* have different meanings for different people. For example, in a survey on how often undergraduates use the campus-based library for studying, the responses should be quantified (e.g., daily, five times per week, twice per week, and so on) instead of using responses such as usually, sometimes, and often.
4. *Phrase questions so as to avoid bias that may predetermine a respondent's answer*. The wording of a question should not point the respondent in a certain direction. Avoid stereotyped, prestige-carrying, emotionally loaded, or superlative words. Some words have such emotional appeal that they tend to bias questions regardless of how they are used. For example, the wording "Have you exercised your American right and registered to vote?" would undoubtedly bias the question. Simply asking, "Are you registered to vote?" would be preferable. Dillman (2007) states that words such as freedom, equality, private enterprise, justice, and honesty have strong positive appeal in our culture. Words such as bureaucratic, socialist, boss, and government planning have strong negative appeal; avoid such words if possible.
5. *Avoid questions that might mislead because of unstated assumptions*. The frame of reference for answering the questions should be clear and

consistent for all respondents. If any assumptions must be made before respondents give an answer, then also include questions designed to inquire into these assumptions. For example, in a survey designed for high school seniors, the question "Do you think your high school has prepared you adequately for college?" assumes the student is going to college and knows what is required in the way of preparation. The question, "Have you registered to vote for the next presidential election?" assumes that the high school student is 18 years of age, which may not be true.

6. *Avoid leading questions, which imply a desired response*. For example, "What do you think of the biased coverage of the Afghanistan War by the cable news networks?" is a leading question.
7. *Avoid questions that may elicit embarrassment, suspicion, or hostility in the respondent*. Questions should not put the respondent on the defensive. For example, people often resent questions about their age, income, religion, or educational status. Instead of asking a subject's age, the researcher can ask for his or her year of birth. People seem less concerned about giving their year of birth than about giving their age. The question "Do you have a high school diploma?" might be asked instead as "What grade had you completed when you left school?" In fact, it is best to avoid personal questions entirely unless such information is essential to the research.
8. *Avoid* ***"double-barreled" questions***, *which attempt to ask two questions in one*. For example, "Do you feel that the university should provide basic skills courses for students and give credit for those courses?" is a double-barreled question. When a respondent answers such a question, the researcher does not know whether the answer applies to both parts of the question or just to one. A yes answer to the preceding question may mean either that the respondent believes the university should offer basic skills courses and give credit for them or that it should offer the courses, but not give credit for them. You can identify a double-barreled question by noting "and," "or," or some other conjunction in the wording.
9. *Make sure the alternatives to each questionnaire item are exhaustive—that they express all the possible alternatives on the issue*. For example, "What is your marital status?" should include not only the alternatives "married" and "single" but also "unmarried cohabiting," "widowed," "divorced," and "separated." In developing the alternatives for questionnaire items designed to identify attitudes or opinions on issues, it is a good idea first to present the questions in an open-ended form to a small sample of respondents. Their answers can then be used as alternatives in the final product. On questions with a wide variety of possible responses, always include the alternative "other," along with a request that the respondent explain that choice. The question "What is your position in the school system?" might be followed by the alternatives "administrator," "teacher," "librarian," and "other (please specify)."
10. *Keep the questionnaire as brief as possible so that it requires a minimum of the respondents' time*. Respondents are much more likely to

PICTURE THIS



"Next question: I believe that life is a constant striving for balance, requiring frequent tradeoffs between morality and necessity, within a cyclic pattern of joy and sadness, forging a trail of bittersweet memories until one slips, inevitably, into the jaws of death. Agree or disagree?"

complete and return a short questionnaire. The researcher must make an effort to eliminate all unnecessary items, especially those whose answers are available from other sources. All the items of a questionnaire should serve a research problem function; that is, they should elicit data needed to test the hypotheses or answer the questions of the research study. For example, you can eliminate a question that asks the respondent's age in a study where this information is not needed in the data analysis.

11. *Make sure the respondents have the information necessary to answer the questions.* Avoid questions dealing with experiences or topics you know are unfamiliar to your sample.

See Fowler (2002) for additional suggestions on writing survey questions.

THINK ABOUT IT 14.6

Which of the basic guidelines for writing good questions are violated in the previous cartoon?

Answer
1, 2, 3, 8, 10, 11

USING A MAILED QUESTIONNAIRE

As we discussed previously, it is not always practical to obtain survey data using an interview format. In these cases questionnaires are used, and quite often they are mailed to respondents. A well-constructed questionnaire is an important factor influencing response rate. We have already offered guidelines for writing the questions. In this section, we discuss the overall arrangement of the questions, the cover letter, the follow-up, and other factors that contribute to the success of a questionnaire.

DIRECTIONS

It is very important to begin with precise directions that tell the respondents exactly what to do. Indicate how and where they are to mark their responses; for example, "Please indicate your response to the following questions by placing an X in the box next to the answer of your choice." If the format changes within the questionnaire, include new directions for that section.

ORDER OF QUESTIONS

The questions must be arranged in appropriate order. This is important because it can influence the respondents' interest in completing the questionnaire. The very first question should be especially interesting and easy enough for all respondents to interpret and answer. If respondents are motivated to answer the first question, they are more likely to continue with the questionnaire. The first question should seek worthwhile information that is related clearly to the topic under consideration. For this reason, rarely begin a questionnaire with questions relating to age, gender, education, occupation, ethnic origin, marital status, and the like. Participants may regard these questions as irrelevant or as an invasion of privacy and decide not to continue with the questionnaire. The first few questions should be of the closed-ended type, which the respondent can complete quickly, instead of open-ended ones that may require a long written response.

It is advisable to group together questions that are similar in content. For example, in a questionnaire asking university faculty about the basic academic skills of their students, all questions related to reading would be placed together. Questions pertaining to writing would appear together, followed later by questions related to mathematics skills. Within the content areas, group items according to the type of question. For example, the questions requiring a simple yes or no answer would be placed together, as would items requiring respondents to rank or indicate their extent of agreement or disagreement.

Within each topic area, arrange the questions in psychological order. A logical or psychological arrangement contributes to better thought-out responses on behalf of the respondents. For example, first ascertain whether respondents were satisfied with working conditions before asking them to recommend changes. If both general and specific questions are asked on a topic, place the general questions first. Objective items on an issue or situation should precede subjective questions. Questions that are less likely to be objectionable should precede those that may be more so. Participants are sometimes reluctant to answer questions about attitudes,

preferences, motives, behavior, personal feelings, and the like, but if objective questions can be used first to clarify and specify the situation, it may be easier for individuals to respond. For example, a researcher who wanted to survey students about the extent of marijuana usage might begin by asking more objective questions first, such as "How would you describe marijuana usage in your school: no problem, slight problem, moderate problem, or serious problem?" This could be followed by the questions, "Do you think the frequency of marijuana smoking has decreased, stayed about the same, or increased this year?" and "Do you know students who use marijuana?"; then perhaps "Do you ever smoke marijuana?" This principle of placing less objectionable questions before more objectionable ones implies that items dealing with demographic data such as age, gender, and occupation should be placed at the end of the questionnaire rather than at the beginning. The respondent will have fewer objections to giving this personal information after completing the questionnaire and understanding why such data would be relevant.

MATRIX SAMPLING

A procedure called **matrix sampling** is sometimes used when a survey is long and the accessible population is large. This technique involves randomly selecting respondents, each of whom is administered a subset of questions randomly drawn from the total set of items. The practical advantage of using matrix sampling is that less time is required for each individual to respond. This is an important because one obstacle to obtaining a high response rate is the unwillingness of some individuals to take the time to answer a long questionnaire.

FORMAT OF THE QUESTIONNAIRE

The questionnaire should be formatted so that it is attractive, easy for the participant to read and answer, and convenient for the researcher to code and score. To achieve these ends, incorporate the following suggestions into the design of the survey document:

1. *Number questions consecutively* throughout the questionnaire, with no repetitions or omissions. Having a unique number for each question avoids confusion in coding responses.
2. *Differentiate questions from response categories* by using regular type for the questions and uppercase letters for the potential responses. Put any specific directions for responding inside parentheses and in lowercase. For example,

   Do you favor setting standards in basic skills as a requirement for high school graduation? (check your response)

   _____ 1) NO
   _____ 2) YES

3. *Use numbers to identify the various response categories*. The numbers assigned to the options represent a form of precoding that will facilitate

data processing. Place a blank or box in front of the response options and ask the respondent to place an *X* or checkmark in the space. For example:

What is the highest level of education that you have completed? (check one)

_____ 1) GRADE SCHOOL
_____ 2) SOME HIGH SCHOOL
_____ 3) HIGH SCHOOL DIPLOMA
_____ 4) SOME COLLEGE
_____ 5) COLLEGE DEGREE
_____ 6) SOME GRADUATE WORK
_____ 7) GRADUATE DEGREE

A checkmark or *X* in front of item 5 indicates that the individual has completed college, and a count can easily be made for that category. Because of the varying lengths of the options, survey designers recommend placing the numbers for the response options at the beginning, to the left of the response categories rather than at the right, for ease of scoring.

4. *Be consistent in assigning numbers to the various response categories*. Always use the same number for the same response throughout the questionnaire. It is conventional to assign low numbers to negative responses and higher numbers to positive responses. For example, 1 is assigned to "no" and 2 is assigned to "yes"; 1 is assigned to "unfavorable" and 2 to "favorable." Whatever the number scheme, use it throughout because it is confusing for a participant to associate 1 with "no" in the first part of the questionnaire and then find 1 associated with "yes" in another part.
5. *Response categories should be arranged vertically rather than horizontally*. The vertical arrangement makes the questionnaire appear less crowded and eliminates the common error of checking the space on the wrong side of the answer, as might occur for the following:

   What is your present marital status?

   _____ 1) NEVER MARRIED _____ 2) MARRIED _____ 3) DIVORCED
   _____ 4) SEPARATED _____ 5) WIDOWED _____ 6) UNMARRIED, COHABITING

   The previous confusing arrangement could be improved in the following way:
   What is your present marital status? (check one)

   _____ 1) NEVER MARRIED
   _____ 2) MARRIED
   _____ 3) DIVORCED
   _____ 4) SEPARATED
   _____ 5) WIDOWED
   _____ 6) UNMARRIED, COHABITING

6. *Use contingency questions when not every question will be relevant to all respondents*. A **contingency question** is one whose relevancy depends on

the response to a prior question. For example, in a survey designed to assess faculty interest in using computer-assisted testing, the following question might be used:

1. Are you interested in using computer-assisted testing in your classes?
   1) NO (if NO, please go to question 6)
   2) YES
   3) UNDECIDED
2. If YES, would you use the computer-assisted testing for
   1) QUIZZES
   2) TESTS
   3) BOTH

In the preceding example, the second question is a contingency question because its relevance is contingent on the answer given in the first question. Faculty members who indicated in question 1 that they were not interested in computer-assisted testing would not even have to read the next four questions, but could proceed to question 6, the next relevant question. Thus, contingency questions save time for the respondent and provide more accurate data for the researcher. Contingency questions can be set off by arrows, indents, or enclosed boxes.

7. *Keep the questionnaire as short as possible*.

PREPARING THE COVER LETTER

Researchers often find it useful to send an introductory letter to potential respondents in advance of the questionnaire. This procedure alerts the subject to the study so that he or she is not overwhelmed by the questionnaire process. In any case, a **cover letter** sent to the respondent by name and title should accompany the questionnaire.

Figure 14.2 shows a cover letter with the important parts labeled. The cover letter introduces potential respondents to the questionnaire and "sells" them on responding. It includes the following elements:

1. *The purpose of the study*. The first paragraph of the letter should explain the purpose of the study and its potential usefulness. It is helpful to relate the importance of the study to a reference group with which the potential respondent may identify. For example, a cover letter with a questionnaire for graduate students should stress the importance of the data for improving graduate education at the university.
2. *A request for cooperation*. The letter should explain why the potential respondent was included in the sample and should make an appeal for the respondent's cooperation. Respondents should be made to feel that they can make an important contribution to the study.
3. *The protection provided the respondent*. The letter must not only assure the respondents that their responses will be confidential, but also explain how that confidentiality will be maintained. To facilitate the follow-up procedure necessary for a high return rate, use identification numbers

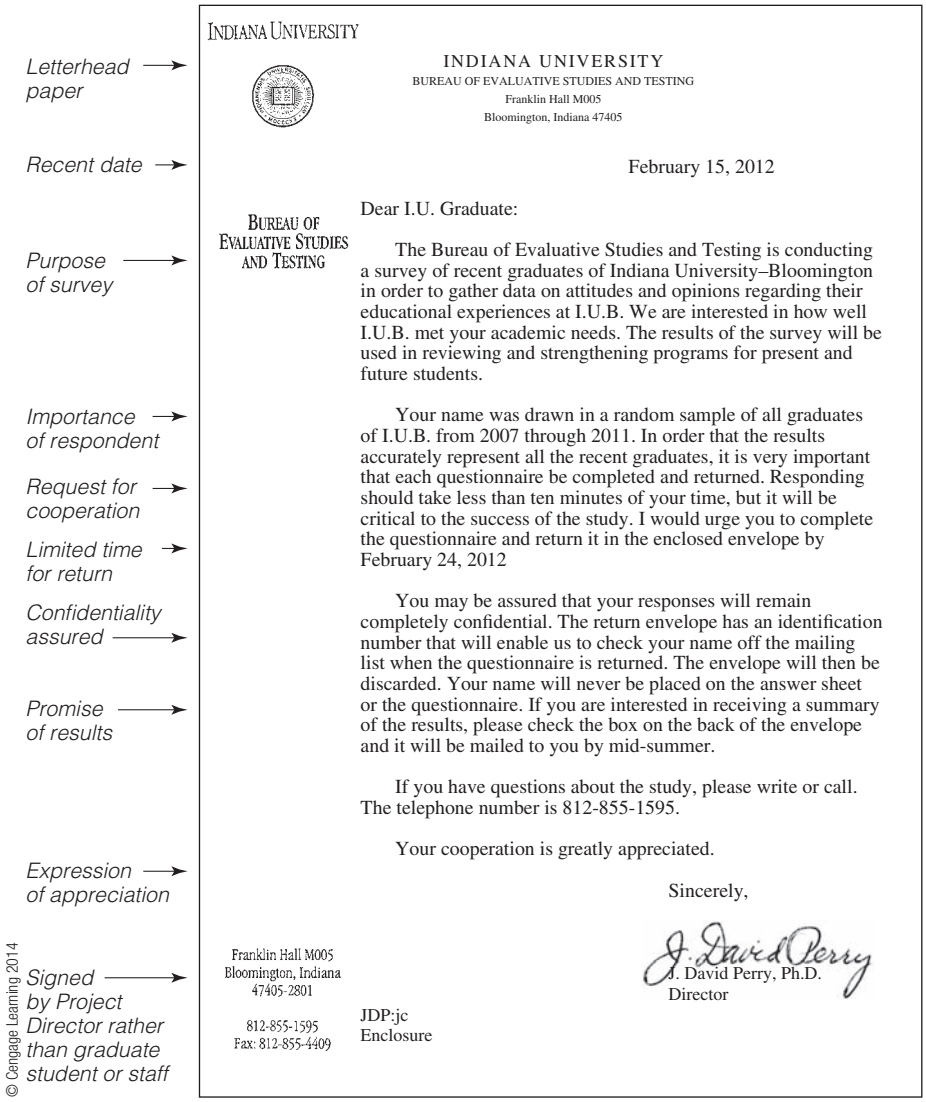


Figure 14.2 Example of a Cover Letter for a Survey

on the questionnaires. Lack of identification numbers compounds the problem of nonresponse bias because there is no way to know who has responded and who has not, making follow-up procedures very difficult. If identification numbers are used, respondents must be told that the numbers are there simply to let the researcher check the respondents' names off the sampling frame list when the questionnaires are returned. The respondents must be assured that their names will never be placed on the questionnaires themselves; thus, there will be no way to associate specific responses with particular individuals. If the researcher intends to destroy the questionnaires immediately after the responses have been added to an analytical spreadsheet, this information should be conveyed in the letter to reassure the respondents of their anonymity.

You may prefer not to use any identification system at all, especially when the topic is sensitive. In this case, we recommend that you include in the package a postcard that the respondent can mail separately to indicate that the questionnaire has also been mailed. This postcard contains a message that the questionnaire has been returned and has a place for the respondent to write his or her name or a coded identifier. You can keep a record of the returned questionnaires this way.

4. *Sponsorship of the study*. The signature on the letter is important in effecting the return of the questionnaire. If the study is part of a doctoral dissertation, it is helpful if a person well-known to the respondents, such as the head of a university department or the dean of the college, signs or countersigns the letter. Such a signature is likely to be more effective than that of an unknown graduate student. If there is a sponsor for the study, such as a foundation or an agency, mention this. Use university or agency letterhead.
5. *Promise of results*. An offer may be made to share the findings of the study with interested respondents. Instruct them on how to request the results. One method is to print the request on the back of the return envelope, with a check box and a place for the respondent's name and address.
6. *Appreciation.* Include an expression of appreciation for the participant's assistance and involvement with the study.
7. *Recent date on the letter*. Date the cover letter near the day of distribution. A potential respondent will not be impressed by a letter dated several weeks before receipt.
8. *Request for immediate return*. It is also important to urge immediate return of the questionnaire. If a time period such as 2 weeks or 1 month is suggested, the respondent may lay the questionnaire aside and, despite good intentions, forget about it. A questionnaire that fails to receive attention within a week is not likely ever to be returned.

Although all of the preceding elements should be included, the cover letter should be as brief as possible and ideally kept to a single page. Enclose the letter in an envelope along with the questionnaire. Always include a self-addressed, stamped return envelope for the respondent's use. This is indispensable for a good return rate.

MAXIMIZING RESPONSE RATES

A variety of factors may influence how much attention a potential participant pays to the request to respond to a survey such as was there *pre-notification* that the survey had been sent?; was the *appearance* of the survey appealing?; did the survey request appear to be *personalized*?; was there an *incentive* for responding to the survey?; was the *sponsorship* from someone perceived as a legitimate source?; did the request explain why the information derived from the survey might be *useful*?; and was it *easy to respond* to the survey and return it at no cost to the participant?

A participant's intention to complete a survey is comprised of various elements such as the *incentive* that a participant may receive if they complete a

survey, the *perceived length* of the survey, one's *attitude* toward research, a participant's perceived *loss of privacy* via completion of a survey, and a participant's *available free time* to engage in survey research.

Completion of a survey is generally predicated on three areas: a participant's *incentive* to complete the survey, the *perceived length* of the survey, and a participant's *attitude* toward research.

Finally, maximizing the response rate by having participants return completed surveys is centered on the prevailing components of *attitude* toward research and participant *incentive*.

Research shows that response rates for most U.S. national surveys of all types have been declining during the past four decades (Cook, Heath, & Thompson, 2000). Helgeson, Voss, and Terpening (2002) conceptualize the process of responding to a request to complete a survey as involving four phases through which the researcher must move the potential respondent.

1. *Attention*, the entry point to the response process, is critical for maximizing the probability of a response. If you gain the respondent's attention, then you have moved the respondent into the process and increased the likelihood that there will be a follow-through to the end.
2. *Intention*, the next phase, involves the potential respondent's estimating the time and effort that must be expended to comply with the researcher's request for information and making a decision to continue with the survey.
3. *Completion* is the phase in which the potential respondent moves from consideration of the survey completion process to the physical and mental activity necessary for actual survey completion.
4. *Return* involves sealing the completed survey into the return envelope and mailing it—a relatively low-cost action compared to earlier phases of the decision process.

We will now summarize some aspects of survey design that may affect the potential respondent at each phase.

MONETARY INCENTIVE

A great deal of research has shown that a token monetary incentive consistently increases response rate. Researchers believe that the payment creates a feeling of obligation on the part of the recipient and a need to reciprocate that is satisfied by completing and returning the questionnaire. Monetary incentives are most effective when prepaid rather than promised and when included with the first mailing of the questionnaire rather than with a follow-up. Among several survey design factors examined, Helgeson, Voss, and Terpening (2002) found that the one-dollar incentive included with the mailed survey had the most effect on returns. Jobber, Saunders, and Mitchell's (2004) research not only supports the use of a prepaid monetary incentive, but also shows that the response rate increases as the value of the incentive increases. They concluded from their studies that the inclusion of any incentive, regardless of amount, raises the response rate by an average of 15 percent. The size of the incentive had an additional effect at the rate of 2 percent per dollar within the observed range of incentives.

Newby, Watson, and Woodliff (2003) also found that monetary incentives effectively increased return rates. Of course, offering payment is not always

possible because even a token amount can greatly increase the cost of the survey. If the sample is not too large, however, it is an option worth considering.

FOLLOW-UPS

To reach the maximum percentage of returns in a mailed questionnaire survey, planned **follow-up** mailings are essential. The steps typically taken in the follow-ups are explained here.

First Reminder

If the questionnaire has not been returned in a week or 10 days after the initial mailing, send a postcard or an e-mail reminder to the respondent. This follow-up serves as a polite reminder that a questionnaire was sent earlier and that the response is very important to the study. Urge respondents to complete and return the questionnaire immediately. Of course, express thanks to those who may have already returned the questionnaire. An offer can be made to send another questionnaire to those who may have misplaced or never received it. Usually, the follow-up reminder brings in a relatively large number of responses.

Second Follow-Up

This follow-up, which should be sent approximately 3 weeks after the original mailing, involves a letter, another copy of the questionnaire, and an addressed return envelope. The letter should first tell nonrespondents that their questionnaires have not been received and should reiterate the usefulness of the study. Emphasize that a replacement questionnaire is enclosed, and make a strong appeal to the respondents to complete and return it. Tell them not to respond a second time if they have already mailed the questionnaire.

Third Follow-Up

The third and final follow-up is sent out 6 or 7 weeks after the initial mailing. It is similar to the second, having both a letter and a replacement questionnaire. Many researchers send this follow-up by certified mail. If a researcher has 75 percent to 90 percent returns after three follow-ups, he or she may be ready to terminate the survey and to declare the remaining subjects nonrespondents. The researcher must decide whether the responses obtained through further follow-up efforts would be worth the cost and time involved. It is sometimes suggested that the researcher include in the third follow-up a postcard on which subjects could indicate that they do not wish to participate in the survey and will not be returning the questionnaire. Such a procedure facilitates definite identification of nonrespondents.

DEALING WITH NONRESPONSE

Nonresponse can be a serious problem in survey research. Nonresponse can bias survey data, especially when it is nonrandom and if it is in some way correlated with the variables measured in the study. Nonrespondents may differ systematically from respondents. Research shows that respondents tend to differ from nonrespondents in characteristics such as education, intelligence, motivation, and interest in the topic of the survey. A survey with low response rate can thus be biased even though the researcher started out by mailing questionnaires

to a representative sample. Recent studies, however, suggest that the effect of nonresponse may not be as pronounced as once thought, and that low response rates may not necessarily indicate bias (McCarty, 2003). Teitler, Reichman, and Sprachman (2003) investigated the costs and benefits of improving response rates for a difficult-to-reach population. They concluded that efforts to improve response rate were beneficial in obtaining a representative sample, but there was a point of diminishing returns beyond which the benefits were marginal. The findings from these studies do not justify low response rates; there is no question that high response rates are preferable to low ones. They simply indicate that lower response rates do not necessarily translate into biased data.

What can a researcher do about nonrespondents? They cannot be ignored. If, after all follow-up attempts, response rate remain below approximately 75 percent, try to learn something about the characteristics of the nonrespondents and the extent to which they might differ from the respondents. There are several ways to do this:

1. *Compare respondents to population*. If you have access to information on characteristics of the population—age, gender, education, ethnicity, socioeconomic status, and so on—you can compare these characteristics of the respondents with those of the population from which the sample was selected. If your data show that your respondents closely resemble the population of interest, it is reasonable to assume that nonrespondents also resemble the parent population. Thus, further study of nonrespondents is not necessary.
2. *Compare early to late respondents*. Research has shown that nonrespondents are often similar to late respondents. Thus, a second way to estimate the possible responses of nonrespondents is to look at those of the late respondents. Prior to this step, however, categorize respondents into early and late groups and compare their responses to check for any important differences. If no important differences appear between early and late respondents, you can assume the respondents are an unbiased sample of the recipients and can thus generalize to the total group.
3. *Interview a sample of nonrespondents*. A more systematic approach is to interview either personally or by telephone a small random sample (i.e., perhaps 10 percent) of the nonrespondents. A sample of nonrespondents drawn for comparison purposes is sometimes called a "double-dipped sample." Using the questionnaire as an interview schedule, the investigator gathers responses from the random sample of nonrespondents. The mean responses or the proportion of responses of nonrespondents can be compared to those of respondents to determine whether the two groups differ meaningfully. If no important differences are found when the responses of the initial respondents are compared with those of the interview sample, the researcher can reasonably assume that the respondents represent an unbiased sample of all who received the questionnaire. The data can be pooled and generalizations made regarding the total sample and the population. Without such a check, the researcher has no way of knowing if the respondents are different and therefore biased. Using a double-dipped sample is the preferred method of checking for bias because the direction and the extent of bias that is due to nonresponse can be directly assessed. However, it is costly and time consuming.

VALIDITY

Attention should be given to the validity of interviews and questionnaires—that is, whether they are really measuring what they purport to measure (i.e., the measurement is meaningful). *Face validity,* which was discussed in Chapter 9, can be important in survey research. Subjects should perceive questions to be relevant. Consider the following questions:

1. The recipe for angel food cake calls for half a cup of sugar. You are only making half a cake. How much sugar should you use?
2. A pilot began a flight with half a tank of fuel. During this flight, he used half of his fuel supply. How much is left?

Both ask basically the same question. Candidates for pilot training would consider question 2 appropriate and question 1 inappropriate. The reverse would be true for culinary arts students. Question 2 has face validity for pilot training candidates. Question 1 has face validity for culinary arts students. Subjects are less inclined to complete and return a questionnaire they perceive as being inappropriate.

Construct validity can be assessed by having some colleagues who are familiar with the purpose of the survey examine it to judge whether the items are appropriate for measuring what they are supposed to measure and whether they are a representative sample of the behavior domain under investigation.

Criterion-related validity can be based on the relationship of survey responses to other variables. Direct observation of behavior, for example, has been a criterion used to validate surveys. After responses were obtained, observations were made to determine whether the actual behavior of the subjects agreed with their expressed attitudes, opinions, or other answers. If you find agreement between survey responses and actual behavior, you have some evidence for the criterion-related validity of the survey. Other data sources such as third parties may also be used as criteria.

Five potential problems may influence the validity of a questionnaire:

1. Respondents often report what they *think* is true or what they *wish* were true, whether or not that is the reality. Direct observation of a random sample of respondents is a good way to learn whether what they report is or is not reality.
2. Respondents may give unreasonable responses that are more socially acceptable than what is actually the case. Take this into account when interpreting survey results. If you read that a survey found that 40 percent of adult Americans would give a kidney transplant to a total stranger, ask yourself if 40 percent of adult Americans would yield a prime parking spot to a total stranger.
3. Respondents may give a response that they think the researcher wants to hear.
4. In interviews, respondents with little or no interest in the topic of the research may give off-hand thoughtless responses just to get the investigator "off their back."

5. Respondents who are not sure their anonymity is assured may give "safe" responses.

For problems 1 through 4, direct observation of the behavior of a random sample of respondents is an excellent way to assess validity. Problem 5 can be avoided by making sure respondents know that their anonymity is safeguarded.

RELIABILITY

Recall from Chapter 9 that in order to be useful, a measure must first elicit score **reliability** (consistency). Reliability indexes range from 0 (absolutely no consistency) to 1.0 (perfectly consistent). The strength of the reliability index sets the upper limit of the validity of a measure.

A procedure for assessing the score reliability of an interview procedure is to have two or more interviewers ask the same subjects identical questions and then assess the consistency of the responses that the interviewers report. With questionnaires, internal consistency may be checked by building some redundancy into the instrument—items on the same topic may be rephrased and repeated in the questionnaire or interview. Thus, the more consistent the responses, the higher the reliability.

Such procedures are often expensive and time consuming, and somewhat impractical because it is not easy to find subjects willing to repeat the questionnaire or interview. Another problem with this approach is that some answers to questions dealing with less stable aspects of behavior may legitimately change over time. However, it is important to assess consistency because if you employ an inconsistent procedure, you compromise your data.

STATISTICAL ANALYSIS IN SURVEYS

Some surveys do not require complex statistical analyses. Data analysis may simply consist of determining the frequencies and percentages of responses for the survey questions. For example, a survey of library resources may report the number of fiction books, the number of nonfiction books, and so on. A survey of people's attitudes on an issue may report the number and percentage of the respondents who gave each response, such as "strongly agree," "agree," "disagree," and so on. It is useful to convert numbers to percentages in order to discuss the proportion responding in a certain way and to make comparisons. Consider the hypothetical frequency data in Table 14.3, based on a survey of 440 students concerning their opinions of a proposed increase in student activity fees at the university they attend. Looking at the raw frequencies, you might say that upperclassmen (66 juniors and seniors) and graduate students (66) are equally opposed to the increase in fees. However, if you calculate *percentages* based on the total number of students *in each group,* you can see that more than twice the proportion of graduate students disapprove of the increase compared to juniors and seniors: 66 of the 88 graduate students (75 percent) disapprove, compared to 66 of 192 (34 percent) juniors and seniors.

Thus, the table shows that graduate students are more likely to disapprove of the activity fee increase than are undergraduates. To avoid misinterpretation, always make sure that the total numbers for different groups are presented in tables such as Table 14.3.

Table 14.3 Students' Attitudes toward Increase in Activity Fee

	Approve	No Opinion	Disapprove	Total
Freshmen and sophomores	60	68	32	160
Juniors and seniors	80	46	66	192
Graduate students	12	10	66	88
Total	152	124	164	440

CONTROLLING VARIABLES IN A SURVEY ANALYSIS

Consider the hypothetical results of a survey of attitudes toward a new library tax for improving and expanding the county public library. Table 14.4 shows that 63 percent (150/240) of city residents favor the library tax compared with 37 percent (96/260) of county residents. It appears from these data that there is a relationship between place of residence and attitude toward the library tax. A chi-square test will show whether there is a statistically significant relationship between the two variables (see Chapter 7 for discussion of chi square). Table 14.5 shows the calculation of chi square for these data. The expected frequencies for each cell are shown in parentheses.

Referring to Table A.4 of the Appendix with 1 degree of freedom, you can see that the chi square of 32.83 is statistically significant at the .01 level. You might conclude that there is a statistically significant association between place of residence and attitude toward the library tax. A more cautious observer, however, might point out that the city (i.e., the location of a major state university) has a greater proportion of educated people, and that it may be educational level rather than place of residence that accounts for the favorable attitude toward the library tax.

To explore this alternative explanation, you can control for the variable of educational level by holding it constant and then observing whether the

Table 14.4 Attitudes toward Library Tax by Residence

	City	County	Total
Favor	150	96	246
Oppose	90	164	254
Total	240	260	500

Table 14.5 Observed and Expected Frequencies for Attitude Data

	City	County	Total
Favor	150 (118)	96 (128)	246
Oppose	90 (122)	164 (132)	254
Total	240	260	500

$$\chi^2 = \frac{(150-118)^2}{118} + \frac{(96-128)^2}{128} + \frac{(90-122)^2}{122} + \frac{(164-132)^2}{132}$$

$$\chi^2 = 32.83$$

Table 14.6 Attitudes toward Library Tax Related to Education, Not to Residence

	College-Educated		
	City	County	Total
Favor	144	80	224
Oppose	36	20	56
Total	180	100	280
	Non-College-Educated		
	City	County	Total
Favor	6	16	22
Oppose	54	144	198
Total	60	160	220

statistically significant association between the first two variables continues to exist. The simplest way to hold the variable constant is to divide the subjects into separate groups, each having a different value on that variable, and look at the crosstabs for each of these groups separately. In this case, you could categorize the 500 respondents into college educated and non-college educated (assuming this information is available) and examine the relationship within the two separate groups.

The data in Table 14.6 show that the alternative explanation is correct: If college education is controlled by holding it constant, there is no statistically significant association between the variables "place of residence" and "attitude toward the tax." Among the college educated, 80 percent (144/180) of the city residents favor the tax and so do 80 percent (80/100) of the county residents. Among the non-college educated, 10 percent (6/60) of the city residents favor the tax, as do 10 percent (16/160) of the rural residents.

Suppose the data had been as shown in Table 14.7. Although educational level has been held constant within each table, the association between the variables "residence" and "attitude" is evident. Among the college educated, 67 percent (80/120) of the city residents favor the tax, compared with 33 percent (20/60)

Table 14.7 Attitudes toward Library Tax Related to Residence, Not to Education

	College-Educated		
	City	County	Total
Favor	80	20	100
Oppose	40	40	80
Total	120	60	180
	Non-College-Educated		
	City	County	Total
Favor	70	76	146
Oppose	50	124	174
Total	120	200	320

of the rural residents. Among the non-college educated, 58 percent of the city residents favor the tax, compared with 38 percent of the rural residents. In this case, there *is* something about the city and county respondents, other than their education, that leads them to respond differently pertaining to the library tax.

You might also want to investigate gender differences in responses to a survey. In this case, the crosstabs would show the frequency of responses to the questions for males and females, separately. Social class differences could be examined by categorizing the subjects into separate groups on the basis of appropriate criteria and looking at the responses for each of the groups. Tables that show frequencies of different groups' responses are often the best way to illustrate the associations among the variables of a survey. These tables are called cross tabulations, or **crosstabs**, because they permit comparisons across groups. The simplest crosstabs contain two variables with two categories for each variable. More complex forms are possible, such as 2×3, 2×4, 3×3, and so on. Cross tabulations are used most often with categorical or nominal data.

STATISTICS FOR CROSSTABS

Crosstabs are a widely used way to graphically show the differences in responses among various groups and whether or not an association may exist between the variables under consideration. When both variables in the cross tabulation are measured at the nominal level, the chi-square test may be used to determine whether a systematic association exists between the two variables. Chi-square, however, indicates *only* whether the variables are associated or are independent; it does not reveal the extent to which they are related. For example, the value of χ^2 for the "college-educated" versus the "non-college educated" data in Table 14.6 is 17.09, which is statistically significant at the .01 level. Thus, you know that there is a statistically significant association between place of residence and attitude toward the tax among the college educated, but you do not know how strongly the two variables are related.

To learn the extent of the relationship, you must calculate a coefficient of correlation. A coefficient that is frequently used with nominal data in 2×2 tables is the *phi coefficient* (φ). The phi coefficient is a mathematical simplification of the Pearson product moment coefficient for 2×2 tables. Thus, phi has a value of 0 when no relationship exists, +1.00 when a perfect positive relationship exists, and −1.00 when a perfect negative relationship exists. The phi coefficient for the college educated in Table 14.6 is .32. The phi coefficient is interpreted like any other Pearson coefficient. The .32 indicates a moderate positive relationship between place of residence and attitude toward the library tax among the college educated.

For tables larger than 2×2, an appropriate measure of relationship is the kappa statistic (κ) (Siegel & Castellan, 1988, p. 284). If there is a perfect relationship between the variables, k will equal 1.00. If agreement between variables is exactly what would be expected through chance, κ equals 0. If agreement is less than what would be expected by chance, κ will be a negative number. If the two variables in the cross tabulation are ordinal, statistics such as the Kendall coefficient of concordance (W) or the gamma statistic can be used to indicate the strength of the relationship between variables. You can find additional, comprehensive discussion of various association-based procedures for nominal and

ordinal variables in Chapter 9 of Siegel and Castellan's (1988) text. Further, a detailed discussion of the influence of weighting and design effect with national longitudinal data, derived via complex cluster sampling design and employing cross tabulations and chi-square, can be found in Walker and Young (2003).

RESEARCH IN THE PUBLIC EYE

In April 2012, *Education Week* reported on a study commissioned by Safe Kids Worldwide and Johnson & Johnson. Five hundred and sixteen athletes ages 8 to 18, 750 parents, and 752 coaches completed an online survey about sports injury prevention. Only 52 percent of coaches felt they were very knowledgeable; four in 10 had never had sports safety training, although eight in 10 felt such knowledge was very important. Forty-seven percent said they did not have time to focus on injury prevention.

Among parents, 49 percent said youth sports injuries were treated by a coach or adult on site, and nearly a third of parents said the injuries were severe enough to require immediate medical attention or a trip to the doctor. Seventy-five percent of the student athletes and 92 percent of the parents said they entrust coaches with player safety. But 39 percent of the coaches said they had felt pressured by a parent to let an injured athlete continue to play. Thirty-one percent of the young athletes believed that "good players should keep playing their sport even if they are hurt, unless a coach or adult makes them stop."

Eighty-eight percent of coaches said they had seen players injured, with 21 percent indicating that broken bones and 16 percent indicating that concussions had been sustained. Seventy-six percent of coaches wished they had concussion prevention training, and 735 wanted physical exams and heat illness prevention training. The majority of coaches wanted free online training modules.

A study reported by the *Washington Post* in February 2011 found that neither the most popular students nor the least popular are likely to be bullies at school, but rather the kids who are in the middle of the social hierarchies. A sociology professor from the University of California at Davis surveyed middle school and high school students in North Carolina for several years. He found moderate gender differences, with girls less physically aggressive and more indirectly aggressive and less likely to bully than are boys. However, overall rates of aggression were generally equivalent. Other findings from the study included:

- Youths from single-parent households are no more aggressive than others.
- Students whose parents have low levels of education are significantly less aggressive.
- While Latino students are more aggressive toward their same-gender classmates, there are no differences between Caucasians African Americans, and other minority students.
- Grade in school matters only for physical aggression.
- Academic achievement and sports participation have little effect on aggression, with the latter modestly increasing overall and verbal aggression.
- Pubertal development, generally thought to increase aggression, has no effect.

Discuss these two survey studies. What questions do you believe the researchers were trying to answer? Based on what you know about these studies, how you would classify each one. What do you think the advantages and disadvantages might be of trying to answer questions about these issues via a survey? How well do you believe the headlines reflect the survey findings?

SUMMARY

The survey is a widely used research method for gathering data ranging from physical counts and frequencies to attitudes and opinions. Surveys are classified according to their focus, scope, and duration of data collection. They should involve careful planning, unbiased sampling of a population, thoughtful development of data-gathering instruments, and careful analysis of the results.

If researchers choose some type of probability sampling, they can infer population values from the sample results. The usual procedure is to set up an interval or range that has a high probability of including the population value. The width of this interval is a function of the risk they are willing to take of being in error and the sample size. The interval narrows as the acceptable probability of error increases and as sample size increases. You can also use this procedure to estimate the sample size needed for any desired margin of error.

Interviews and questionnaires are the major means of data collection for a survey. Both procedures involve questioning selected subjects, but each has unique advantages and disadvantages. It is important that the instruments derive scores that are as valid and reliable as possible. Various follow-up procedures have proved effective in increasing returns from mailed questionnaires.

Cross tabulations provide an excellent way to show the associations existing among the variables in a survey.

KEY CONCEPTS

census
census of intangibles
census of tangibles
checklists
closed-ended question
cohort study
completion (fill-in) items
contingency question
cover letter
cross-sectional survey
crosstabs
directly administered questionnaire
double-barreled question
electronic mail survey
field test
focus group
follow-up
intangibles
interviewer bias
Likert-type item
longitudinal survey
mailed questionnaire
margin of error
matrix sampling
nonresponse
open-ended question
panel study
pause
personal interview
population
probe
questionnaire
ranking items
reliability
response rate
sample
sample survey
sample survey of intangibles
sample survey of tangibles
sampling error
sampling frame
scaled items
social desirability bias
survey research
tangibles
trend study
web-based survey

EXERCISES

1. Suggest a research question that can best be answered by means of a survey. Write two open-ended items, two closed-end items, and two Likert-type scale items that would provide data relevant to the question.

2. What data-gathering technique would you use for each of the following surveys?
   a. A survey of a sample of high school teachers throughout the state concerning the use of mandatory competency exams for high school graduation

b. A survey of the opinions of people in a major metropolitan area on the way juveniles who commit violent crimes are currently handled in the state's court system

c. A survey of the opinions of a sample of professors on the campus of State University about the use of pass/fail grades instead of letter grades (you want a very high response rate)

d. A freshman survey of certain noncognitive characteristics that might relate to academic achievement in the first year of college

e. A survey of 500 people from throughout the United States about their attitudes toward the administration's immigration policy

3. How does the proportion of yes/no responses influence the sample size needed?

4. A CNN poll reported the following: In a recently conducted survey of the American public, 37 percent of the respondents said they approve of the president's performance. What else would you want to know before you interpreted this report?

5. How would you interpret the following report? "A poll of 1,000 randomly selected registered voters in Minnesota found that 45 percent favored using state lottery profits for education. Figures from this poll are subject to a sampling error of 3 percent. The confidence level is 95 percent."

6. For each of the following three sample sizes, calculate the 95 percent margin of error for the population proportion. Assume the sample proportion is .40 for each. What effect does an increase in sample size have on the width of the interval? Why?

n of sample A = 100
n of sample B = 1000
n of sample C = 10,000

7. A school superintendent wants to determine what proportion of the 5000 middle school and high school students in his district use iPads for learning. He will distribute a questionnaire to a random sample of the students. How many students will he need to sample in order to have a 95 percent margin of error of ±5 percent in his estimate?

8. A national polling organization wants to be able to predict the outcome of the presidential election to within 3 percent. How large of a random sample will be needed to achieve this level of precision? Assume a 95 percent confidence level.

9. A survey had an initial response rate of 51 percent. What would you suggest to the researcher for dealing with this moderately low response rate?

10. The administration at State University wants to estimate the incoming freshman class's interest in a computer science major. Because it does not have the financial resources to survey all 5000 freshmen, the administration surveys a random sample of 500. It finds that 100 students report they are interested in such a major.

a. Calculate the margin of error in this survey.

b. What is the best estimate of the number of freshmen who would be interested in majoring in computer science?

11. Which of the following would be biased samples of the population of college students at a large university?

a. A random sample of students entering the library on Friday evening

b. A random sample of students registered for classes

c. A random sample of students buying season tickets for basketball

d. A sample composed of students who volunteered for a project after seeing a notice on the university's homepage

12. Assume you are conducting a survey to determine how elementary school teachers in a district feel about the policy of retaining students.

a. Write two closed-ended questions for this survey.

b. Write two open-ended questions for this survey.

c. Write a contingency question.

13. On the basis of the time of data collection, classify each of the following surveys:

a. Terman's study of adults who were intellectually gifted as children
b. A comparison of math achievement in public middle schools in the United States in 2002, 2007, and 2012
c. A follow-up of the 2007 graduates of the Duke University School of Business
d. A survey of reading achievement at different grade levels in a school system in 2012

14. Which of the following is an advantage of the longitudinal type of survey?
a. More intensive individual study
b. Providing data for different age groups at the same time
c. Prompt data gathering
d. No sampling errors

15. *The New York Times* editors invited readers to fill out a detailed questionnaire related to attitudes about crime and return it to the *Times*. The *Times* subsequently printed an article about the widespread fear of crime among U.S. citizens, especially older people. How would you evaluate this survey?

16. A graduate student is planning to use a survey to gather data for her dissertation, but is unsure whether to use mailed questionnaires or telephone interviews. Examine the following list of considerations and indicate whether a questionnaire or an interview would be more appropriate.
a. She is on a tight financial budget for her data collection.
b. Her goal is a 90 percent response rate from the sample.
c. Some questions may involve sensitive issues.
d. She hopes to minimize the time needed for coding and organizing data and getting it ready for computer analysis.

ANSWERS

1. Answers will vary. *Example*: How do adults in a community feel about the building of a new elementary school?

2. **a.** Mailed questionnaire
b. Telephone interview
c. Personal interview
d. Directly administered questionnaire
e. Telephone interview

3. The nearer the proportions are to .50/.50, the larger the sample size needed.

4. You would want to know the size of the sample, how it was drawn, the width of the interval around the estimated population parameter (the margin of error), and the confidence level.

5. The .95 confidence interval for the proportion of registered voters who favor using lottery funds for education is between 42 and 48 percent.

6. **Sample A**

$$\sigma_p = \sqrt{\frac{(.40)(.60)}{100}} = \sqrt{\frac{.24}{100}}$$

$$= \sqrt{.0024} = .04899$$

$.04899 \times 1.96 = .0960$ or .10 margin of error

Sample B

$$\sigma_p = \sqrt{\frac{(.40)(.60)}{1000}} = \sqrt{\frac{.24}{1000}}$$

$$= \sqrt{.0024} = .0155$$

$.0155 \times 1.96 = .03$ margin of error

Sample C

$$\sigma_p = \sqrt{\frac{(.40)(.60)}{10{,}000}} = \sqrt{\frac{.24}{10{,}000}}$$

$$= \sqrt{.000024} = .004899$$

$.0049 \times 1.96 = .01$ margin of error

An increase in sample size decreases the width of the interval. As n increases, the sample error decreases and hence the size of the interval.

7. Use Formula 14.2. Assume $p = q = .50$:

$$n = \left(\frac{\frac{1}{E}}{\sqrt{pq}}\right)^2 (1.96)^2 = \left(\frac{\frac{1}{.05}}{\sqrt{(.5)(.5)}}\right)^2 (1.96)^2$$

$$= \left(\frac{1}{.10}\right)^2 = (1.96)^2 = 100\,(3.8416) = 384$$

8. Assume $p = q = .50$:

$$n = \left(\frac{\frac{1}{E}}{\sqrt{pq}}\right)^2 (1.96)^2 = \left(\frac{\frac{1}{.03}}{\sqrt{(.5)(.5)}}\right)^2 (1.96)^2$$

$$= \left(\frac{1}{.06}\right)^2 (1.96)^2 = (277.78)(3.8416) = 1067$$

9. The researcher should follow up with a postcard reminder and then another mailing or two of the questionnaire. After follow-up efforts have been completed, the researcher should try to interview some of the remaining nonrespondents to find out about their characteristics and to obtain their responses in order to determine if they differ significantly from the respondents.

10. a. $p = .20$

$q = .80$ $\quad \sigma_p = \sqrt{\frac{pq}{n}}$

$$\sigma_p = \sqrt{\frac{(.20)(.80)}{500}}$$

$$= \sqrt{.00032} = .0179$$

margin of error $= 1.96 \times .0179 = .035$
or $\pm 3.5\%$

b. Between 16.5 and 23.5 percent or between 825 and 1175 students would be interested in a computer science major.

11. Samples a, c, and d would not be representative of the population of college students at a large university.

12. Answers will vary.

13. a. Longitudinal panel study
b. Longitudinal trend study
c. Longitudinal cohort study
d. Cross-sectional survey

14. a

15. Those who completed and returned the questionnaire would not be a representative sample of all U.S. citizens. A number of factors would bias this sample, such as socioeconomic level, educational level, or sufficient interest in the topic to complete the questionnaire.

16. a. Mailed questionnaire
b. Telephone interview
c. Mailed questionnaire
d. Telephone interview

REFERENCES

Cook, C., Heath, F., & Thompson, R. (2000). A meta-analysis of response rates in web or Internet-based surveys. *Educational and Psychological Measurement, 60*(6), 821–836.

DeVaus, D. (2002). *Conducting surveys using the Internet*. Thousand Oaks, CA: Sage.

Dillman, D. A. (2007). *Mail and Internet surveys: The tailored design method* (2nd ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Fowler, F. J. (2002). *Survey research methods* (3rd ed.). Thousand Oaks, CA: Sage.

Helgeson, J., Voss, K., & Terpening, W. (2002). Determinants of mail survey response: Survey design factors and respondent factors. *Psychology and Marketing*, *19*, 303–328.

Holbrook, A., Green, M., & Krosnick, J. (2003). Telephone versus face-to-face interviewing of national probability samples with long questionnaires. *Public Opinion Quarterly*, *67*, 79–125.

Jobber, D., Saunders, J., & Mitchell, V. (2004). Prepaid monetary incentive effects on mail survey research. *Journal of Business Research*, *57*(4), 347–350.

McCarty, C. (2003). Differences in response rates using most recent versus final dispositions in telephone surveys. *Public Opinion Quarterly*, *67*, 396–406.

Meister, D., & Melnick, S. (2003). National new teacher study: Beginning teachers' concerns. *Action in Teacher Education*, *24*(4), 87–94.

Neuman, W. L., & Kreuger, L. (2003). *Social work research methods: Qualitative and quantitative approaches*. Boston: Allyn & Bacon.

Newby, R., Watson, J., & Woodliff, D. (2003). SME survey methodology: Response rates, data quality, and cost effectiveness. *Entrepreneurship Theory and Practice*, *28*, 163–172.

Siegel, S., & Castellan, N. J., Jr. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York: McGraw-Hill.

Strauss, V. (February 8, 2011). Most, least popular kids less likely to bully, study says. Retrieved from www.washingtonpost.com/answersheet.

Teitler, J., Reichman, N., & Sprachman, S. (2003). Costs and benefits of improving response rates for a hard-to-reach population. *Public Opinion Quarterly*, *67*, 126–138.

Terman, L. M. (1926). The mental and physical traits of a thousand gifted children. In *Genetic studies of genius* (Vol. 1). Stanford, CA: Stanford University Press.

Toporek, B. (April 25, 2012). Survey: Many coaches misinformed about youth sports safety risks. Education Week.

Walker, D. A. (2005). The standard error of a proportion for different scores and test length. *Practical Assessment, Research and Evaluation*, *10*(5), 1–9.

Walker, D. A., & Young, D. Y. (2003). Example of the impact of weights and design effects on contingency tables and chi-square analysis. *Journal of Modern Applied Statistical Methods*, *2*(2), 425–432.