9

CHAPTER

Using Survey Research

Gordon Allport (1954) characterized an attitude as "probably the most distinctive and indispensable concept in contemporary social psychology" (p. 43). Since Allport's assessment, attitudes have transcended social psychology to become important in our everyday lives. We are surrounded by issues related to attitudes and their measurement. Pollsters and politicians are constantly measuring and trying to change our attitudes about a wide range of issues (such as abortion, the war on terrorism, and tax cuts). How and where we obtain information on these issues are also changing.

On November 4, 2008, a historic election took place in the United States. For the first time in history an African American was elected to the office of President of the United States. Not only did the 2008 election reflect a change in America's willingness to vote for an African American candidate, it also reflected a change in how many citizens obtained their information on the candidates and the important political issues underlying the election. According to a 2009 survey conducted by the Pew Research Center, 74% of Internet users relied on the Internet to participate in or get information about the presidential election. More interestingly, there was a major increase in the percentage of adults in general as well as Internet users who obtain political news over the Internet (see Figure 9-1 for these trends).

The increased reliance on Internet sources for political news was true for a wide range of demographic groups. For example, the percentage of adults who sought political information online increased among all age groups from 2004 to 2008, with the greatest net increase among 18 to 24 year olds (a 21% increase). The increase was evident among all income groups measured (with the greatest increase among those earning less than $30,000 per year) and among Democrats (a 10% increase), Republicans (a 9% increase), and independents (a 3% increase). Additionally, the Pew survey found that Obama supporters were more likely than opponent McCain supporters to engage in a variety of online political activities. For example, Obama supporters were more likely to use social networks (25%)
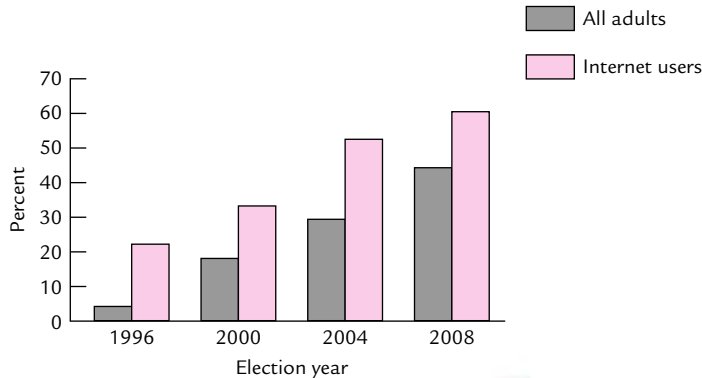
FIGURE 9-1   Trends in the use of the Internet to obtain political news.
SOURCE: http://pewresearch.org/pubs/1192/internet-politics-campaign-2008. Based on data provided at the Web site.

than McCain supporters (16%), and were more likely to post political content online (26% and 15% for Obama and McCain supporters, respectively).

Surveys are a widely used research technique. You may have participated in a survey yourself, or (perhaps more likely) you may have been the recipient of survey results. If you have answered a few questions from a local political party during election time, you have participated in a survey. Even those annoying questions on warranty registration cards that come with most products qualify as a survey of sorts. You are typically asked about your age, income, interests, magazines to which you subscribe, and so on. If you answered those questions and mailed back the card, you took part in a survey.

Even if you rarely participate in surveys, you are still likely to have encountered survey results. Political polls designed to gauge people's attitudes on key issues and candidates come out almost daily during election time. Polls about the U.S. president's approval rating, wars, and health care issues come out several times over the course of a year.

Because survey research is highly visible, you should understand the "ins and outs" of this important research technique. If you plan to use a survey technique in your own research, you should know about proper questionnaire construction, administration techniques, sampling techniques, and data analysis. Even if you never use survey techniques, understanding something about them will help you make sense out of the surveys that you are exposed to every day.

SURVEY RESEARCH

Before we discuss survey techniques, note the difference between the *field survey* and the observational techniques described in Chapter 8. In both naturalistic observation and participant observation, you simply observe behaviors and make copious notes about them. You do not administer any measures to your participants. Consequently,

you can only speculate about the motives, attitudes, and beliefs underlying the observed behaviors. In a field survey, you directly question your participants about their behavior (past, present, or future) and their underlying attitudes, beliefs, and intentions. From the data collected, you can draw inferences about the factors underlying behavior.

The inferences that you can draw from a field survey are limited by the fact that you do not manipulate independent variables. Instead, you acquire several (perhaps hundreds of) measures of the behaviors of interest. This purely correlational research strategy usually does not permit you to draw causal inferences from your data (see Chapter 4). For example, finding that political conservatism is a good predictor of voter choices does not justify concluding that political conservatism *causes* voter choices.

Instead, you use the field survey to evaluate specific attitudes such as those concerning issues surrounding nuclear disarmament, political candidates, or foreign imports. You also can use the field survey to evaluate behaviors. For example, you could design a questionnaire to determine which household products people use.

Surveys also have another important use: predicting behavior. Political polls often seek to predict behavior. Attitudes about political candidates are assessed, and then projections are made about subsequent voter behavior.

When you conduct survey research, you must ensure that your participants are treated ethically. One major ethical issue concerns whether and how you will maintain the *anonymity* of your participants and the *confidentiality* of their responses. Maintaining anonymity means that you guarantee there will be no way for the participants' names to be associated with their answers. This might be accomplished by instructing participants to mail back their questionnaires and informed-consent forms separately. No coding scheme would be used that would allow you to match up individual participants and their questionnaires. However, sometimes you may wish to code the questionnaires and informed-consent forms so that you can match them up later. You might do this, for example, if a participant has second thoughts about participating after the questionnaire has been returned. If so and you have promised your participants that their responses will remain anonymous, you must take steps to ensure that only authorized personnel associated with the research project can gain access to the code and only for the stated purpose.

Maintaining confidentiality means that you do not disclose any data in individual form, even if you know which participants filled out which questionnaires. If you promise your participants that their responses will remain confidential, ethical practice dictates that you report only aggregate results.

QUESTIONS TO PONDER

1. What are some of the applications of survey research?
2. Why is it important to know about survey methods, even if you do not intend to conduct surveys?
3. How does a field survey differ from other observational methods?
4. What are anonymity and confidentiality and why are they important?

DESIGNING YOUR QUESTIONNAIRE

The first step in designing a questionnaire is to clearly define the topic of your study. A clear, concise definition of what you are studying will yield results that can be interpreted unambiguously. Results from surveys that do not clearly define the topic area may be confusing. It is also important to have clear, precise operational definitions for the attitudes or behaviors being studied. Behaviors and attitudes that are not defined precisely also may yield results that are confusing and difficult to interpret.

Having a clearly defined topic has another important advantage: It keeps your questionnaire focused on the behavior or attitude chosen for study (Moser & Kalton, 1972). You should avoid the temptation to do too much in a single survey. Tackling too much in a single survey leads to an inordinately long questionnaire that may confuse or overburden your participants. It also may make it more difficult for you to summarize and analyze your data (Moser & Kalton, 1972). Your questionnaire should include a broad enough range of questions so that you can thoroughly assess behavior but not so broad as to lose focus and become confusing. Your questionnaire should elicit the responses you are most interested in without much extraneous information.

The type of information gathered in a questionnaire depends on its purpose. However, most questionnaires include items designed to assess the characteristics of the participants, such as age, sex, marital status, occupation, income, and education. Such characteristics are called *demographics*. Demographics are often used as *predictor variables* during analysis of the data to determine whether participant characteristics correlate with or predict responses to other items in the survey. Other, nondemographic items also can be included to provide predictor variables. For example, attitude toward abortion might be used to predict voter preference. In this case, attitude toward abortion would be used as a predictor variable.

In addition to demographics and predictor variables, you will have items designed to assess the behavior of interest. For example, if you were interested in predicting voter preference, you would include an item or items on your questionnaire specifically to measure voter preference (e.g., asking participants to indicate candidate preferences). That item, or a combination of several items, would constitute the *criterion variable*.

The questions to which your participants will respond are the heart of your questionnaire. Take great care to develop questions that are clear, to the point, and relevant to the aims of your research. The time spent in this early phase of your research will pay dividends later. Well-constructed items are easier to summarize, analyze, and interpret than poorly constructed ones. The next section introduces several popular item formats and offers suggestions for writing good questionnaire items.

Writing Questionnaire Items

Writing effective questionnaire items that obtain the information you want requires care and skill. You cannot simply sit down, write several questions, and use those first-draft questions on your final questionnaire. Writing questionnaire items involves

writing and rewriting items until they are clear and succinct. In fact, having written your items and assembled your questionnaire, you should administer it to a pilot group of participants matching your main sample in order to ensure that the items are reliable and valid.

When writing questionnaire items, you may choose among several popular types. Here we discuss the open-ended, restricted, partially open-ended, and rating-scale item types.

Open-Ended Items   **Open-ended items** allow the participant to respond in his or her own words. The following example might appear in a survey like the Pew Internet use survey:

How often did you use the Internet to get political news for the 2008 presidential election?

The participant writes an answer to the question in the space provided immediately below. Such information may be more complete and accurate than the information obtained with a restricted item (discussed next). A drawback to the open-ended item is that participants may not understand exactly what you are looking for or may inadvertently omit some answers. Thus, participants may fail to provide the needed information. Another drawback to the open-ended item is that it can make summarizing your data difficult. Essentially, you must perform a content analysis on open-ended answers. All of the methods and rules that we discussed in Chapter 8 would come into play. It may be tempting to interpret open-ended responses rather than just summarize them, running the risk of misclassifying the answers.

Restricted Items   **Restricted items** (also called *closed-ended items*) provide a limited number of specific response alternatives. A restricted item with ordered alternatives lists these alternatives in a logical order, as shown in this item adapted from the Pew survey:

How often did you use the Internet to get political news during the 2008 presidential election campaign?

__ Very often

__ Sometimes

__ Not too often

__ Never

Note how the alternatives for this question go from very often to never. Participants would respond by checking the blank space to the left of the desired answer. However, other methods for recording choices can be used with restricted items. For example, you could use a number to the right of each alternative and have participants circle the numbers corresponding to their choices.

Use unordered alternatives whenever there is no logical basis for choosing a given order, as shown in this example from the Pew survey:

Do you think that the political information you obtained from the Internet during the 2008 presidential election campaign was generally accurate or inaccurate?

__ Accurate

__ Inaccurate

__ Neither

__ Don't know

Because there is no inherent order to the alternatives, other orders would serve just as well. For example, you just as easily could have put "Inaccurate" before "Accurate."

By offering only specific response alternatives, restricted items control the participant's range of responses. The responses made to restricted items are therefore easier to summarize and analyze than the responses made to open-ended items. However, the information that you obtain from a restricted item is not as rich as the information from an open-ended item. Participants cannot qualify or otherwise elaborate on their responses. Also, you may fail to include an alternative that correctly describes the participant's opinion, thus forcing the participant to choose an alternative that does not really fit.

Partially Open-Ended Items   **Partially open-ended items** resemble restricted items but provide an additional, "other" category and an opportunity to give an answer not listed among the specific alternatives, as shown in this example adapted from the Pew survey:

In what capacity did you most use the Internet during the 2008 presidential election campaign?

__ Post political content online

__ Engage politically on an online social network

__ Share political videos, pictures, or audio content

__ Sign up for online political updates

__ Donate money online

__ Other (Specify) _____

Dillman (2000) offers several suggestions for formatting restricted and partially open-ended items. First, use a boldface font for the stem of a question and a normal font for response category labels (as we have done in the previous examples). This helps respondents separate the question from the response categories that follow. Second, make any special instructions intended to clarify a question a part of the question itself. Third, put check boxes, blank spaces, or numbers in a consistent position throughout your questionnaire (e.g., to the left of the response alternatives). Fourth, place all alternatives in a single column. Other tips offered by Dillman (2000) for constructing and formatting questionnaire items are summarized in Table 9-1.

TABLE 9-1   Suggestions for Writing Good Survey Items

SUGGESTION	EXAMPLE
Use simple rather than complex words.	Use "work" rather than "employment."
Make the stem of a question as short and easy to understand as possible, but use complete sentences.	"Would you like to study in America?"
Avoid vague questions in favor of more precise ones.	Use "How many years have you lived in your current house?" rather than "Years in your house."
Avoid asking for too much information. Respondents may not have an answer readily available.	Use a list of ordered alternatives rather than an open-ended question when asking how often the respondent does something.
Avoid "check all that apply" questions.	Instead of "check all that apply," list each item separately and have respondent indicate liking/disliking for each.
Avoid questions that ask for more than one thing.	Instead of asking "Would you like to study and then live in America?" ask "Would you like to study in America?" and "Would you like to live in America?" separately.
Soften the impact of potentially sensitive questions.	Instead of asking "Have you ever stolen anything?" ask "Have you ever taken anything without paying for it?"

SOURCE: After Dillman, 2000.

Rating Scales   A variation on the restricted question uses a rating scale rather than response alternatives. A rating scale provides a graded response to a question:

How much confidence do you have that the political news you obtained from the Internet during the 2008 presidential campaign was accurate?

1	2	3	4	5	6	7	8	9	10
No confidence									A lot of confidence

There is no set number of points that a rating scale must have. A rating scale can have as few as 3 and as many as 100 points. However, rating scales commonly do not exceed 10 points. A 10-point scale has enough points to allow a wide range of choice while not overburdening the participant. Scales with fewer than 10 points also are used frequently, but you should not go below 5 points. Many participants may not

want to use the extreme values on a scale. Consequently, if you have a 5-point scale and the participant excludes the end points, you really have only three usable points. Scales ranging from 7 to 10 points leave several points for the participants to choose among, even if participants do avoid the extreme values.

You also must decide how to label your scale. Figure 9-2 shows three ways that you might do this. In panel (a), only the end points are labeled. In this case, the participant is told the upper and lower limits of the scale. Such labeled points are called *anchors* because they keep the participant's interpretation of the scale values from drifting.

With only the end points anchored, the participant must interpret the meaning of the rest of the points. In Figure 9-2(b), all points are labeled. In this case, the participant knows exactly what each point means and may consequently provide more accurate information. In Figure 9-2(c), the scale is labeled at the end points and at the midpoint. This scale provides three anchors for the participant. This scale is a reasonable compromise between labeling only the end points and labeling all the points.

You may be wondering whether labeling each point changes the way that the participant responds on the scale. The answer seems to be a qualified no. When you develop a measurement scale, you are dealing with (1) the psychological phenomenon underlying the scale and (2) the scale itself. Labeling each point does not change the nature of the psychological phenomenon underlying the scale. You can assume that your scale, labeled at each point, still represents the phenomenon underlying the scale. In fact, researchers have sometimes expressed a misguided concern about such scale transformations (Nunnally, 1967). Minor transformations of a measurement
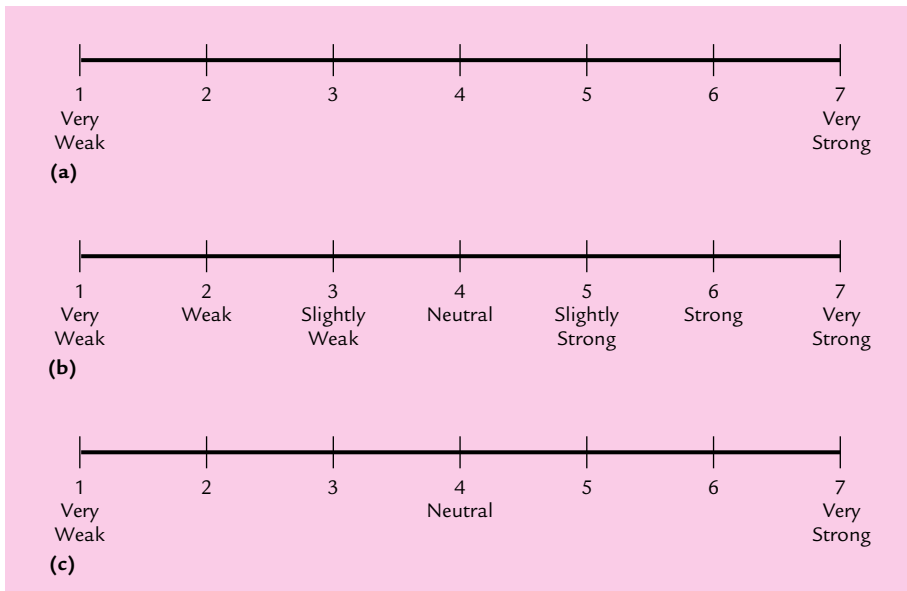

FIGURE 9-2   Three ways of labeling a rating scale: (a) end points only, (b) each point labeled, and (c) end points and midpoint labeled.

scale (such as labeling each point) probably do not affect its measurement properties or how well it represents the underlying psychological phenomenon being studied.

In the previous examples, participants respond by checking or circling the scale value that best represents their judgments. Alternative ways to format your scale give participants more flexibility in their responses. Figure 9-3 shows an example in which the end points are anchored and the participants are instructed to place a check or perpendicular line on the scale to indicate how they feel. To quantify the responses, you use a ruler to measure from an end point to the participant's mark. Your scale is then expressed in terms of inches or centimeters, and the resulting numbers are treated just like the numbers on a numbered scale.

Another variation on the rating scale is the *Likert scale*, which is widely used in attitude measurement research. A Likert scale provides a series of statements to which participants can indicate degrees of agreement or disagreement. Figure 9-4 shows two examples of formatting a Likert-scale item. In the first example, the attitude statement is followed by five blank spaces labeled from "Strongly Agree" to "Strongly Disagree." The participant simply checks the space that best reflects the degree of agreement or disagreement with each statement. The second example provides consecutive numbers rather than blank spaces and includes descriptive anchors only at the ends. Participants are instructed to circle the number that best reflects how much they agree or disagree with each statement. (For further information on Likert scaling, see Edwards, 1953).

A final note on rating scales is in order. Although rating scales have been presented in the context of survey research, be aware that rating scales are widely used in experimental research as well. Adapting rating scales to your particular research needs is a relatively simple affair. Anytime that your research calls for the use of rating scales, you can apply the suggestions presented here.

QUESTIONS TO PONDER

1. What are the steps involved in designing a questionnaire?

2. How do open-ended and restricted items differ, and what are the advantages and disadvantages of each?

3. What are the ways in which questionnaire items can be formatted?

4. What are some of the factors that you should pay attention to when constructing questionnaire items?

5. How do you design effective rating scales?

Very
Weak

Very
Strong

FIGURE 9-3   Rating scale formatted with no numbers. End points are labeled, and participants place marks on the line to indicate their responses.

Most political information on the Internet is accurate.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
_____	_____	_____	_____	_____

(a)

Most political information on the Internet is accurate.

Strongly Agree				Strongly Disagree
1	2	3	4	5

(b)

FIGURE 9-4   Samples showing Likert scales: (a) a standard Likert item on which the participant places a check in the blank under the statement that best reflects how he or she feels; (b) a five-point Likert scale using numbers that the participant circles.

Assembling Your Questionnaire

If your questionnaire is to be effective, its items must be organized into a coherent, visually pleasing format. This process involves paying attention to the order in which the items are included and to the way in which they are presented.

Dillman (2000) and Moser and Kalton (1972) agree that demographic items should *not* be presented first on the questionnaire. These questions, although easy to complete, may lead participants to believe that the questionnaire is boring. Dillman emphasizes the importance of the first question on a questionnaire. A good first question should be interesting and engaging so that the respondent will be motivated to continue. According to Dillman, the first question should apply to everybody completing the questionnaire, be easy so that it takes only a few seconds to answer, and be interesting. Of course, these rules are not carved in stone. If your research needs require a certain question to be presented first, that consideration should take precedence (Dillman, 2000).

Your questionnaire should have continuity; that is, related items should be presented together. This keeps your participant's attention on one issue at a time rather than jumping from issue to issue. Your questionnaire will have greater continuity if related items are grouped. An organized questionnaire is much easier and more enjoyable for the participant to complete, factors that may increase the completion rate. Continuity also means that groups of related questions should be logically ordered. Your questionnaire should read like a book. Avoid the temptation to skip around from topic to topic in an attempt to hold the attention of the participant. Rather, strive to build "cognitive ties" between related groups of items (Dillman, 2000).

The order in which questions are included on a questionnaire has been shown to affect the responses of participants. For example, McFarland (1981) presented

questions on a questionnaire ordered in two ways. Some participants answered a general question before specific questions, whereas others answered the specific questions first. McFarland found that participants expressed more interest in politics and religion when the specific questions were asked first than when the general questions were asked first. Sigelman (1981) found that question order affected whether or not participants expressed an opinion (about the popularity of the president), but only if the participants were poorly educated. Hence, question order may play a greater role for some participants than for others. Carefully consider your sample and the chosen topic when deciding on the order in which questions are asked.

The placement of items asking for sensitive information (such as sexual preferences or illegal behavior) is an important factor. Dillman (2000) suggests placing objectionable questions after less objectionable ones, perhaps even at the end of the questionnaire. Once your participants are committed to answering your questions, they may be more willing to answer some sensitive questions. Additionally, a question may not seem as objectionable after the respondent has answered previous items than if the objectionable item is placed earlier in the questionnaire (Dillman, 2000). You also should pay attention to the way that each page of your questionnaire is set up. There should be a logical "navigational path" (Dillman, 2000) that your respondent can follow. This path should lead the respondent through the questionnaire as if he or she were reading a book.

One way to accomplish this is to use appropriate graphics (e.g., arrows and other symbols) to guide respondents through the questionnaire. In fact, Dillman talks about two "languages" of a questionnaire. One language is verbal and relates to how your questions are worded. The other language is graphical and relates to the symbols and graphics used to guide respondents through the items on your questionnaire. Symbols and graphics can be used to separate groups of items, direct respondents where to go in the event of a certain answer (e.g., "If you answered 'No' to item 5, skip to item 7" could be accompanied by an arrow pointing to item 7), or direct respondents to certain pages on the questionnaire. Dillman suggests the following three steps for integrating the verbal and graphical languages into an effective questionnaire:

1. Design a navigational path directing respondents to read all the information on a page.

2. Create effective visual navigational guides to help respondents stay on the navigational path.

3. Develop alternate navigational guides to help with situations where the normal navigational guide will be interrupted (e.g., skipping items or sections).

QUESTIONS TO PONDER

1. Why is the first question on a questionnaire so important?

2. What does it mean that a questionnaire should have continuity? Why is continuity important?

3. What is a questionnaire's navigational path, and why is it important?

ADMINISTERING YOUR QUESTIONNAIRE

After you develop your questionnaire, you must decide how to administer it. You could mail your questionnaire to your participants, deliver your questionnaire via e-mail or post it on the Internet, telephone participants to ask the questions directly, administer your questionnaire to a large group at once, or conduct face-to-face interviews. Each method has advantages and disadvantages and makes its own special demands.

Mail Surveys

In a **mail survey**, you mail your questionnaire directly to your participants. They complete and return the questionnaire at their leisure. This is a rather convenient method. All you need to do is put your questionnaires into addressed envelopes and mail them. However, a serious problem called **nonresponse bias** occurs when a large proportion of participants fail to complete and return your questionnaire. If the participants who fail to return the questionnaire differ in significant ways from those who do return it, your survey may yield answers that do not represent the opinions of the intended population.

Combating Nonresponse Bias   To reduce nonresponse bias, you should develop strategies to increase your return rate. Dillman (2000) notes that the single most effective strategy for increasing response rate is to make multiple contacts with respondents. Dillman suggests making four contacts via mail. The first consists of a *prenotice letter* sent to the respondent a few days before the questionnaire is sent. The prenotice letter should inform the respondent that an important questionnaire will be coming in the mail in a few days. It also should inform the respondent what the survey is about and why the survey will be useful. The second mailing would deliver the questionnaire itself, accompanied by a cover letter. The cover letter should include the following elements in the order listed (Dillman, 2000): the specific request to complete the questionnaire, why the respondent was selected to receive the survey, the usefulness of the survey, a statement of confidentiality of the respondent's answers, an offer of a token of appreciation (if such an offer is to be made), an offer to answer questions, and a real signature.

The third mailing would take the form of a *thank you postcard* sent a few days or a week after the questionnaire was mailed. The postcard should thank the respondent for completing the questionnaire and remind the respondent to complete the questionnaire if not already done. The fourth contact provides a *replacement questionnaire*, sent 2 to 4 weeks after the original questionnaire and accompanied by a letter indicating that the original questionnaire had not been received. The letter also should urge the respondent to complete the replacement questionnaire and return it.

You may be able to increase your return rate somewhat by including a small token of your appreciation, such as a pen or pencil that the participant can keep. Some researchers include a small amount of money as an incentive to complete the questionnaire. As a rule, it is better to send the token along with the questionnaire rather than make the token contingent upon returning the questionnaire. One study found that 57% of respondents returned a survey questionnaire when promised $50

for its return whereas 64% returned the questionnaire when $1 was included with it (James & Bolstein, 1990).

Ironically, smaller rewards seem to produce better results than larger ones (Kanuk & Berenson, 1975; Warner, Berman, Weyant, & Ciarlo, 1983). Dillman (2000) suggests that a $1 token is preferred because it is easy to mail and seems to produce the desired results. Finally, monetary incentives work better than tangible rewards (Church, 1993).

A few factors that do *not* significantly affect response rate include questionnaire length, personalization, promise of anonymity, and inclusion of a deadline (Kanuk & Berenson, 1975). (For reviews of the research supporting these findings, see Kanuk & Berenson, 1975, and Warner et al., 1983.)

Internet Surveys

An increasingly popular method of administering questionnaires is to post them on the Internet. **Internet surveys** can be distributed via e-mail or listserves or posted on a Web site. Which method you use depends on the nature and purpose of your survey. E-mail surveys are easy to distribute but do not permit complex navigational designs (Dillman, 2000). Consequently, e-mail surveys are best for relatively short, simple questionnaires. Web-based surveys allow you to create and present more complex questionnaires that incorporate many of the design features discussed previously (Dillman, 2000). To aid you in the task of implementing a Web-based survey, commercial software packages are available that allow you to design sophisticated questionnaires for posting on a Web site. There is significant advantage to using the Internet to conduct a survey or recruit participants: You can reach a large body of potential participants with relative ease. Data can be collected quickly and easily, resulting in a large data set. You still need to consider the problem of nonresponse bias. As with the mail survey, you can combat this problem with prenotification. For an Internet survey a short text message to potential respondents is more effective than an e-mail notice (Bosnjak, Neubarth, Couper, Bandilla, & Kaczmirek, 2008).

There are also disadvantages to Internet surveys. As discussed in Chapter 6, a sample of respondents from the Internet may not be representative of the general population. According to a 2007 study by the U.S. Department of Commerce (2008), only 61.7% of households had access to the Internet in the home. Further, households with higher levels of education and income were more likely to have Internet access. Additionally, access was greater for Asians (75.5%) and Whites (67.0%) than Blacks (44.9%). Another disadvantage is that one must have the resources available to post a survey on the Internet. This requires computer space on a server and the ability to create the necessary Web pages or the resources to pay someone to create your net survey for you.

Despite the potential for biased samples in Internet surveys, there is evidence that the results obtained from Internet surveys are equivalent to the results obtained from paper-and-pencil surveys. Alan De Beuckelear and Flip Lievens (2009) conducted a survey across 16 countries using both Internet and paper-and-pencil deliveries. The results showed that in all of the countries the Internet and paper-and-pencil surveys returned equivalent results. De Beuckelear and Lievens (2009) concluded that data collected with the two methods could be combined because the two methods

produced such highly similar data. In another study, Christopher Fleming and Mark Bowden (2009) found that the sample demographics of an Internet and a mail survey on travel preferences did not differ significantly.

In both of the studies just cited, the topics of the surveys were not sensitive or controversial. There is some evidence that the equivalence of Internet and conventional methods may not apply to more sensitive topics (DiNitto, Busch-Armendariz, Bender, Woo, Tackett-Gibson, & Dyer, 2009). DiNitto, et al. conducted a survey over the Internet and by telephone asking men about sexual assault behaviors. The results showed that respondents in both types of survey reported sexual assault behavior. However, a wider variety of sexual assault behaviors were reported by respondents to the telephone survey.

So, where does this leave us? It would appear that Internet surveys may produce comparable results to other survey methods for nonsensitive issues. You can be reasonably confident that your Internet survey on such issues will yield data that are highly similar to data collected with more conventional methods. However, you must exercise more caution when surveying about sensitive behaviors. In the latter case, an Internet survey may produce results that differ from more conventional methods.

Telephone Surveys

In a **telephone survey**, you contact participants by telephone rather than by mail or via the Internet. You can ask some questions more easily over the telephone than you can in written form. Telephone surveys can be done by having an interviewer ask respondents a series of questions or by interactive voice response (IVR). Telephone surveys using live interviewers have lost popularity as new technologies have become available. IVR surveys involve respondents using a touch-tone telephone to respond to a series of prerecorded questions. Modern IVR technologies also allow respondents to provide verbal answers in addition to numeric responses.

Telephone surveys may not be the best way to administer a questionnaire. The plethora of "junk calls" to which the population is exposed has given rise to a backlash against telephone intrusions. Laws have been passed on the state and federal level protecting people from unwanted calls, making it more difficult to reach prospective respondents. These laws, combined with caller ID and answering machines (which allow residents to screen their calls), make the telephone a less attractive medium for surveys now than in the past.

Group-Administered Surveys

Sometimes you may have at your disposal a large group of individuals to whom you can administer your questionnaire. In such a case, you design your questionnaire as you would for a mail survey but administer it to the assembled group. For example, you might distribute to a first-year college class a questionnaire on attitudes toward premarital sex. Using such a captive audience permits you to collect large amounts of data in a relatively short time. You do not have to worry about participants misplacing or forgetting about your questionnaire. You also may be able to reduce any volunteer bias, especially if you administer your questionnaire during a class period. People may participate because very little effort is required.

As usual, this method has some drawbacks. Participants may not treat the questionnaire as seriously when they fill it out as a group as when they fill it out alone. Also, you may not be able to ensure anonymity in the large group if you are asking for sensitive information. Participants may feel that other participants are looking at their answers. (You may be able to overcome this problem by giving adjacently seated participants alternate forms of the questionnaire.) Also, a few participants may express hostility about the questionnaire by purposely providing false information.

A final drawback to group administration concerns the participant's right to decline participation. A participant may feel pressure to participate in your survey. This pressure arises from the participant's observation that just about everyone else is participating. In essence, a conformity effect occurs because completing your survey becomes the norm defined by the behavior of your other participants. Make special efforts to reinforce the understanding that participants should not feel compelled to participate.

Face-to-Face Interviews

Still another method for obtaining survey data is the **face-to-face interview**. In this method, you talk to each participant directly. This can be done in the participant's home or place of employment, in your office, or in any other suitable place. If you decide to use a face-to-face interview, keep several things in mind. First, decide whether to use a structured interview or an unstructured interview. In a *structured interview,* you ask prepared questions. This is similar to the telephone survey in that you prepare a questionnaire in advance and simply read the ordered questions to your participants. In the *unstructured interview,* you have a general idea about the issues to discuss. However, you do not have a predetermined sequence of questions.

An advantage of the structured interview is that all participants are asked the same questions in the same order. This eliminates fluctuations in the data that result from differences in when and how questions are asked. Responses from a structured interview are therefore easier to summarize and analyze. However, the structured interview tends to be inflexible. You may miss some important information by having a highly structured interview. The unstructured interview is superior in this respect. By asking general questions and having participants provide answers in their own words, you may gain more complete (although perhaps less accurate) information. However, responses from an unstructured interview may be more difficult to code and analyze later on. You can gain some advantages of each method by combining them in one interview. For example, begin the interview with a structured format by asking prepared questions; later in the interview, switch to an unstructured format.

Using the face-to-face interview strategy leads to a problem that is not present in mail or Internet surveys but is present to some extent in telephone surveys: The appearance and demeanor of the interviewer may affect the responses of the participants. Experimenter bias and demand characteristics become a problem. Subtle changes in the way in which an interviewer asks a question may elicit different answers. Also, your interviewer may not respond similarly to all participants (e.g., an interviewer may react differently to an attractive participant than to an unattractive one). This, too, can affect the results.

The best way to combat this problem is to use interviewers who have received extensive training in interview techniques. Interviewers must be trained to ask questions in the same way for each participant. They also must be trained not to emphasize any particular words in the stem of a question or in the response list. The questions should be read in a neutral manner. Also, try to anticipate any questions that participants may have and provide your interviewers with standardized responses. This can be accomplished by running a small pilot version of your survey before running the actual survey. During this pilot study, try out the interview procedure on a small sample of participants. (This can be done with just about anyone, such as friends, colleagues, or students.) Correct any problems that arise.

Another problem with the interview method is that the social context in which the interview takes place may affect a participant's responses. For example, in a survey of sexual attitudes known as the "Sex in America" survey (Michael, Gagnon, Laumann, & Kolata, 1994), some questions were asked during a face-to-face interview. Some participants were interviewed alone whereas others were interviewed with a spouse or other sex partner present. Having the sex partner present changed the responses to some questions. For example, when asked a question about the number of sex partners one had over the past year, 17% of the participants interviewed alone reported two or more. When interviewed with their sex partner present, only 5% said they had two or more sex partners. It would be most desirable to conduct the interviews in a standardized fashion with only the participant present.

A Final Note on Survey Techniques

Although each of the discussed techniques has advantages, the mail survey has been the most popular. The mail survey can reach large numbers of participants at a lower cost than either the telephone survey or the face-to-face interview (Warner et al., 1983) and produces data that are less affected by *social desirability effects* (answering in a way that seems socially desirable). For these reasons, consider mail surveys first.

After designing your questionnaire and choosing a method of administration, the next step is to assess the reliability and validity of your questionnaire. This is typically done by administering your questionnaire to a small but representative sample of participants. Based on the results, you may have to rework your questionnaire to meet acceptable levels of reliability and validity. In the next sections, we introduce you to the processes of evaluating the reliability and validity of your questionnaire.

QUESTIONS TO PONDER

1. What are the different ways of administering a questionnaire?
2. What are the advantages and disadvantages of the different ways of administering a questionnaire?
3. What is nonresponse bias and what can you do to combat it?
4. How do social desirability effects affect your decision about how to administer a questionnaire?

ASSESSING THE RELIABILITY OF YOUR QUESTIONNAIRE

Constructing a questionnaire is typically not a one-shot deal. That is, you don't just sit down and write some questions and magically produce a high-quality questionnaire. Developing a quality questionnaire usually involves designing the questionnaire, administering it, and then evaluating it to see if it does the job.

One dimension you must pay attention to is the reliability of your questionnaire. In Chapter 5, we defined *reliability* as the ability of a measure to produce the same or highly similar results on repeated administrations. This definition extends to a questionnaire. If, on testing and retesting, your questionnaire produces highly similar results, you have a reliable instrument. In contrast, if the responses vary widely, your instrument is not reliable (Rogers, 1995).

In Chapter 5, we described two ways to assess the reliability of a measure: the test–retest method and the split-half method. In the next sections, we discuss the application of these two methods when assessing the reliability of a questionnaire.

Assessing Reliability by Repeated Administration

Evaluating test–retest reliability is the oldest and conceptually simplest way of establishing the reliability of your questionnaire. You simply administer your questionnaire, allow some time to elapse, and then administer the questionnaire (or a parallel form of it) again to the same group of participants. Although this method is relatively simple to execute, you need to consider some issues before using it.

First, you must consider how long to wait between administrations of your questionnaire. An intertest interval that is too short may result in participants remembering your questions and the answers they gave. This could lead to an artificially high level of test–retest reliability. If, however, you wait too long, test–retest reliability may be artificially low. According to Tim Rogers (1995), the intertest interval should depend on the nature of the variables being measured, with an interval of a few weeks being sufficient for most applications. Rogers suggests that test–retest methods may be particularly problematic when applied to the following:

1. *Measuring ideas that fluctuate with time.* For example, an instrument to measure attitudes toward universal health care should not be evaluated with the test–retest method because attitudes on this topic seem to shift quickly.

2. *Issues for which individuals are likely to remember their answers on the first testing.*

3. *Questionnaires that are very long and boring.* The problem here is that participants may not be highly motivated to accurately complete an overly long questionnaire and therefore may give answers that reduce reliability.

Some of the problems inherent in using the *same* measure on multiple occasions can be avoided by using alternate or parallel forms of your questionnaire for multiple testing sessions. As noted in Chapter 5, the type of reliability being assessed with this technique is known as parallel-forms reliability (Rogers, 1995).

For the parallel-forms method to work, the two (or more) forms of your questionnaire must be equivalent so that direct comparison is meaningful. According to Rogers (1995), parallel forms should have the same number of items and the same response format, cover the same issues with different items, be equally difficult, use the same instructions, and have the same time limits. In short, the parallel versions of a test must be as equivalent as possible (Rogers, 1995).

Although the parallel-forms method improves on the test–retest method, it does not solve all the problems associated with multiple testing. Using parallel forms does not eliminate the possibility that rapidly changing attitudes will result in low reliability. As with the test–retest method, such changes make the questionnaire appear less reliable than it actually is. In addition, practice effects may occur even when alternate forms are used (Rogers, 1995). Even though you use different questions on the parallel form, participants may respond similarly on the second test because they are familiar with your question format.

Assessing Reliability With a Single Administration

Because of the problems associated with repeated testing, you might consider assessing reliability by means of a single administration of your questionnaire. As noted in Chapter 5, this approach involves splitting the questionnaire into equivalent halves and deriving a score for each half; the correlation between scores from the two halves is known as split-half reliability (Rogers, 1995). This technique works best when your survey is limited to a single specific area (e.g., sexual behavior) as opposed to multiple areas (sexual behavior and sexual attitudes).

Although the split-half method circumvents the problems associated with repeated testing, it introduces others. First, when you split a questionnaire, each score is based on a limited set of items, which can reduce reliability (Rogers, 1995). Consequently, the split-half method may underestimate reliability. Second, it is not clear how splitting should be done. If you simply do a first-half/second-half split, artificially low reliability may occur if the two halves of the form are not equivalent or if participants are less motivated to answer questions accurately on the second half of your questionnaire and therefore give inconsistent answers to your questions. One remedy for this is to use an odd–even split. In this case, you derive a score for the odd items and a score for the even items.

Perhaps the most desirable way to assess the split-half reliability of your questionnaire is to apply the Kuder–Richardson formula. This formula yields the average of all the split-half reliabilities that could be derived from splitting your questionnaire into two halves in every possible way. The resulting number (designated KR20) will lie between 0 and 1; the higher the number, the greater the reliability of your questionnaire. A KR20 of .75 indicates a "moderate" level of reliability (Rogers, 1995).

In cases in which your questionnaire uses a Likert format, a variation on the Kuder–Richardson formula known as *coefficient alpha* is used (Rogers, 1995). Like KR20, coefficient alpha is a score between 0 and 1, with higher numbers indicating greater reliability. Computation of this formula can be complex. For details, see a text on psychological testing (e.g., see Cohen & Swerdlik, 2010; Rogers, 1995).

Increasing Reliability

Regardless of the method you use to assess the reliability, there are steps you can take to increase the reliability of your questionnaire (Rogers, 1995):

1. Increase the number of items on your questionnaire. Generally, higher reliability is associated with increasing numbers of items. Of course, if your instrument becomes too long, participants may become angry, tired, or bored. You must weigh the benefits of increasing questionnaire length against possible liabilities.

2. Standardize administration procedures. Reliability will be enhanced if you treat all participants alike when administering your questionnaire. Make sure that timing procedures, lighting, ventilation, instructions to participants, and instructions to administrators are kept constant.

3. Score your questionnaire carefully. Scoring errors can reduce reliability.

4. Make sure that the items on your questionnaire are clear, well written, and appropriate for your sample (see our previous discussion on writing items).

QUESTIONS TO PONDER

1. What is meant by the reliability of a questionnaire and why is it important?
2. How do you assess reliability with repeated administrations?
3. How do you assess reliability with a single administration?
4. What steps can be taken to increase reliability?

ASSESSING THE VALIDITY OF YOUR QUESTIONNAIRE

In Chapter 5, we discussed the validity of a measure and described several forms of validity that differ in their method of assessment: content validity, criterion-related validity, construct validity, and face validity. As with other measures, a questionnaire must have validity if it is to be useful; that is, it must measure what it is intended to measure. For example, if you are designing a questionnaire to assess political attitudes, the questions on your test should tap into political attitudes and not, say, religious attitudes.

Here we review content validity, construct validity, and criterion-related validity as applied to a questionnaire (Rogers, 1995). In a questionnaire, *content validity* assesses whether the questions cover the range of behaviors normally considered to be part of the dimension that you are assessing. To have content validity, your questionnaire on political attitudes should include items relevant to all the major issues relating to such attitudes (e.g., abortion, health care, the economy, and defense). The *construct validity* of a questionnaire can be established by showing that the questionnaire's results agree with predictions based on theory.

Establishing the *criterion-related validity* of a questionnaire involves correlating the questionnaire's results with those from another, established measure. There are two ways to do this. First, you can establish *concurrent validity* by correlating your questionnaire's results with those of another measure of the *same* dimension administered at the same time. In the case of your questionnaire on political attitudes, you would correlate its results with those of another, established measure of political attitudes. Second, you can establish *predictive validity* by correlating the questionnaire's results with some behavior that would be expected to occur, given the results. For example, your questionnaire on political attitudes would be shown to have predictive validity if the questionnaire's results correctly predicted election outcomes.

The validity of a questionnaire may be affected by a variety of factors. For example, as noted earlier, how you define the behavior or attitude that you are measuring can affect validity. Validity also can be affected by the methods used to gather your data. In the "Sex in America" survey, some respondents were interviewed alone and others with someone else present. One cannot be sure that the responses given with another person present represent an accurate reflection of one's sexual behavior (Stevenson, 1995). Generally, methodological flaws, poor conceptualization, and unclear questions can all contribute to lowered levels of validity.

QUESTIONS TO PONDER

1. What is the validity of a questionnaire and why is it important?
2. What are the different types of validity you should consider?
3. What factors can affect the validity of your questionnaire?

ACQUIRING A SAMPLE FOR YOUR SURVEY

In Chapter 6, we distinguished between a population (all individuals in a well-defined group) and a sample (a smaller number of individuals selected from the population). Once you have designed and pretested your questionnaire, you then administer it to a group of participants. It is usually impractical to have everyone in the population (however that may be defined) complete your survey. Instead, you administer your questionnaire to a small sample of that population.

Proper sampling is a crucial aspect of sound survey research methodology. Without proper sampling, you can't generalize your results to your target population (e.g., accurately predict voter behavior in an election). Three sampling-related issues you must consider are representativeness, sampling technique, and sample size.

Representativeness

Regardless of the technique you use to acquire your sample, your sample should be representative of the population of interest. A **representative sample** closely matches the characteristics of the population. Imagine that you have a bag containing 300 golf balls: 100 are white, 100 are orange, and 100 are yellow. You then select a sample of