Asking People About Themselves: Survey Research

LEARNING OBJECTIVES

- Discuss reasons for conducting survey research.
- Identify factors to consider when writing questions for interviews and questionnaires, including defining research objectives and question wording.
- Describe different ways to construct questionnaire responses, including closed-ended questions, open-ended questions, and rating scales.
- Compare the two ways to administer surveys: written questionnaires and oral interviews.
- Define interviewer bias.
- Describe a panel study.
- Distinguish between probability and nonprobability sampling techniques.
- Describe simple random sampling, stratified random sampling, and cluster sampling.
- Describe haphazard sampling, purposive sampling, and quota sampling.
- Describe the ways that samples are evaluated for potential bias, including sampling frame and response rate.
Survey research employs questionnaires and interviews to ask people to provide information about themselves—their attitudes and beliefs, demographics (age, gender, income, marital status, and so on) and other facts, and past or intended future behaviors. In this chapter we will explore methods of designing and conducting surveys, including sampling techniques.

WHY CONDUCT SURVEYS?

A multitude of surveys are being conducted all the time. Just look at your daily newspaper. The Centers for Disease Control and Prevention is reporting results of a survey of new mothers asking about breast feeding. A college survey center is reporting the results of a telephone survey asking about political attitudes. If you look around your campus, you will find academic departments conducting surveys of seniors or recent graduates. If you make a major purchase, you will likely receive a request to complete a survey that asks about your satisfaction. I recently visited the American Psychological Association Web site and read a report called Stress in America that presented the results of an Internet survey of over 1,800 adults that was conducted in 2007. Surveys are clearly a common and important method of studying behavior.

Surveys provide us with a methodology for asking people to tell us about themselves. They have become extremely important as society demands data about issues rather than only intuition and anecdotes. My department needs data from graduates to help determine changes that should be made to the curriculum. Auto companies want data from buyers to assess and improve product quality and customer satisfaction. Without collecting such data, we are totally dependent upon stories we might hear or letters that a graduate or customer might write. Other surveys can be important for making public policy decisions by lawmakers and public agencies. In basic research, many important variables, including attitudes, current emotional states, and self-reports of behaviors, are most easily studied using questionnaires or interviews.

We often think of survey data providing a “snapshot” of how people think and behave at a given point in time. However, the survey method is also an important way for researchers to study relationships among variables and ways that attitudes and behaviors change over time. For example, Steinberg and Dornbusch (1991) examined the relationship between the number of hours that high school students work and variables such as grade point average, drug and alcohol use, and psychosomatic distress. The sample consisted of 3,989 students in grades 10-12 at nine high schools in California and Wisconsin. The researchers found that “long work hours during the school year are associated with lower investment and performance in school, greater psychological and somatic distress, drug and alcohol use, delinquency, and autonomy from parents” (Steinberg & Dornbusch, 1991, p. 304). Figure 7.1 shows a typical finding: There are some positive aspects of working fewer than 10 hours per week (as opposed
Why Conduct Surveys?

![Graph showing the relationship between work hours and grade point average.](image)

**FIGURE 7.1**
Relationship between hours of work and grade point average


to not being employed); however, increasingly negative effects are associated with longer work hours.

Survey research is also important as a complement to experimental research findings. Recall from Chapter 2 that Winograd-and Soloway (1986) conducted experiments on the conditions that lead to forgetting where we place something. To study this topic using survey methods, Brown and Rahhal (1994) asked both younger and older adults about their actual experiences when they hid something and later forgot its location. They reported that older adults take longer than younger adults to find the object and that older adults hide objects from potential thieves, whereas younger people hide things from friends and relatives. Interestingly, most lost objects are eventually found, usually by accident in a location that had been searched previously. This research illustrates a point made in previous chapters that multiple methods are needed to understand any behavior.

An assumption that underlies the use of questionnaires and interviews is that people are willing and able to provide truthful and accurate answers. Researchers have addressed this issue by studying possible biases in the way people respond. A response set is a tendency to respond to all questions from a particular perspective rather than to provide answers that are directly related to the questions. Thus, response sets can affect the usefulness of data obtained from self-reports. The most common response set is called social desirability, or "faking good." The social desirability response set leads the individual to answer in the most socially acceptable way—the way that "most people" are perceived to respond or the way that would reflect most favorably on the person. Social
desirability can be a problem in many research areas, but it is probably most acute when the question concerns a sensitive topic such as violent or aggressive behavior, substance abuse, or sexual practices. However, it should not be assumed that people consistently misrepresent themselves. Jourard (1969) suggested that people are most likely to lie when they don’t trust the researcher. If the researcher openly and honestly communicates the purposes and uses of the research, promises to provide feedback about the results, and assures confidentiality, then the participants can reasonably be expected to give honest responses.

We turn now to the major considerations in survey research: constructing the questions that are asked, choosing the methods for presenting the questions, and sampling the individuals taking part in the research.

CONSTRUCTING QUESTIONS TO ASK

A great deal of thought must be given to writing questions for questionnaires and interviews. This section describes some of the most important factors to consider when constructing questions.

Defining the Research Objectives

When constructing questions for a survey, the first thing the researcher must do is explicitly determine the research objectives: What is it that he or she wishes to know? The survey questions must be tied to the research questions that are being addressed. Too often, surveys get out of hand when researchers begin to ask any question that comes to mind about a topic without considering exactly what useful information will be gained by doing so. This process will usually require the researcher to decide on the type of questions to ask. As noted previously, there are three general types of survey questions (Judd, Smith, & Kidder, 1991).

Attitudes and Beliefs Questions about attitudes and beliefs focus on the ways that people evaluate and think about issues. Should more money be spent on mental health services? Are you satisfied with the way that police responded to your call? How do you evaluate this instructor?

Facts and Demographics Factual questions ask people to indicate things they know about themselves and their situation. In most studies, asking some demographic information is necessary to adequately describe your sample. Age and gender are typically asked. Depending on the topic of the study, questions on such information as ethnicity, income, marital status, employment status, and number of children might be included. Obviously, if you are interested
in making comparisons among groups, such as males and females, you must ask the relevant information about group membership. It is unwise to ask such questions if you have no real reason to use the information, however.

Other factual information you might ask will depend on the topic of your survey. Each year, Consumer Reports magazine asks me to tell them about the repairs that have been necessary on many of the products that I own, such as my car and dishwasher. Factual questions about illnesses and other medical information would be asked in a survey of health and quality of life.

**Behaviors** Other survey questions can focus on past behaviors or intended future behaviors. How many times last week did you exercise for 20 minutes or longer? How many children do you plan to have? Have you ever been so depressed that you called in sick to work?

**Question Wording**

A great deal of care is necessary to write the very best questions for a survey. Cognitive psychologists have identified a number of potential problems with question wording (see Graesser, Kennedy, Wiemer-Hastings, & Ottati, 1999). Many of the problems stem from a difficulty with understanding the question, including (a) unfamiliar technical terms, (b) vague or imprecise terms, (c) ungrammatical sentence structure, (d) phrasing that overloads working memory, and (e) embedding the question with misleading information. Here is a question that illustrates some of the problems identified by Graesser et al.:

> Did your mother, father, full-blooded sisters, full-blooded brothers, daughters, or sons ever have a heart attack or myocardial infarction?

There is memory overload because of the length of the question and the need to keep track of all those relatives while reading the question, and the respondent must worry about two different diagnoses with regard to each relative. Further, the term *myocardial infarction* may be unfamiliar to most people. How do you write questions to avoid such problems? The following items are important to consider when you are writing questions.

**Simplicity** The questions asked in a survey should be relatively simple. People should be able to easily understand and respond to the questions. Avoid jargon and technical terms that people won’t understand. Sometimes, however, you have to make the question a bit more complex to make it easier to understand. Usually this occurs when you need to define a term or describe an issue prior to asking the question. Thus, before asking whether someone approves of Proposition J, you will probably want to provide a brief description of the content of this ballot measure.
Double-Barreled Questions  Avoid “double-barreled” questions that ask two things at once. A question such as “Should senior citizens be given more money for recreation centers and food assistance programs?” is difficult to answer because it taps two potentially very different attitudes. If you are interested in both issues, ask two questions.

Loaded Questions  A loaded question is written to lead people to respond in one way. For example, the questions “Do you favor eliminating the wasteful excesses in the public school budget?” and “Do you favor reducing the public school budget?” will likely elicit different answers. Or consider that men are less likely to say they have “raped” someone than that they have “forced sex”; similarly, women are less likely to say they have been raped than forced to have unwanted sex (Koss, 1992). Questions that include emotionally charged words such as rape, waste, immoral, ungodly, or dangerous may influence the way that people respond and thus lead to biased conclusions.

Negative Wording  Avoid phrasing questions with negatives. This question is phrased negatively: “Do you feel that the city should not approve the proposed women’s shelter?” Agreement with this question means disagreement with the proposal. This phrasing can confuse people and result in inaccurate answers. A better format would be: “Do you believe that the city should approve the proposed women’s shelter?”

“Yea-Saying” and “Nay-Saying”  When you ask several questions about a topic, there is a possibility that a respondent will employ a response set to agree or disagree with all the questions. Such a tendency is referred to as “yea-saying” or “nay-saying.” The problem here is that the respondent may in fact be expressing true agreement, but alternatively may simply be agreeing with anything you say. One way to detect this response set is to word the questions so that consistent agreement is unlikely. For example, a study of family communication patterns might ask people how much they agree with the following statements: “The members of my family spend a lot of time together” and “I spend most of my weekends with friends.” Similarly, a measure of loneliness (e.g., Russell, Peplau, & Cutrona, 1980) will phrase some questions so that agreement means the respondent is lonely (“I feel isolated from others”) and others with the meaning reversed so that disagreement indicates loneliness (e.g., “I feel part of a group of friends”). Although it is possible that someone could legitimately agree with both items, consistently agreeing or disagreeing with a set of related questions phrased in both standard and reversed formats is an indicator that the individual is “yea-saying” or “nay-saying.”

Graesser and his colleagues have developed a computer program called QUAID (Question Understanding Aid) that analyzes question wording. Researchers can try out their questions online at the QUAID Web site (http://mnemosyne.csl.psyc.memphis.edu/quaid). You should also review the question wording examples in Table 7.1.

Graesser and his colleagues have developed a computer program called QUAID (Question Understanding Aid) that analyzes question wording. Researchers can try out their questions online at the QUAID Web site (http://mnemosyne.csl.psyc.memphis.edu/quaid). You should also review the question wording examples in Table 7.1.
TABLE 7.1 Question wording: What is the problem?

<table>
<thead>
<tr>
<th>Read each of the following questions and identify the problems for each.</th>
<th>Negative wording</th>
<th>Simplicity</th>
<th>Double-barreled</th>
<th>Loaded</th>
</tr>
</thead>
</table>
| Professors should not be required to take daily attendance.  
1 = (Strongly Disagree) and  
5 = (Strongly Agree) | | | | |
| I enjoy studying and spending time with friends on weekends. | | | | |
| Do you support the legislation that would unfairly tax hard-working farmers? | | | | |
| I would describe myself as attractive and intelligent. | | | | |
| Do you believe the relationship between cell phone behavior and consumption of fast food is orthogonal? | | | | |
| Restaurants should not have to be inspected each month. | | | | |
| Are you in favor of the boss's whim to cut lunchtime to 30 minutes? | | | | |

Answers are provided at the end of the chapter.

**RESPONSES TO QUESTIONS**

**Closed- Versus Open-Ended Questions**

Questions may be either closed- or open-ended. With closed-ended questions, a limited number of response alternatives are given; with open-ended questions, respondents are free to answer in any way they like. Thus, you could ask a person: “What is the most important thing children should learn to prepare them for life?” followed by a list of answers from which to choose (a closed-ended question) or you could leave this question open-ended for the person to provide the answer.

Using closed-ended questions is a more structured approach; they are easier to code and the response alternatives are the same for everyone. Open-ended questions require time to categorize and code the responses and are therefore more costly. Sometimes a respondent’s response cannot be categorized at all because the response doesn’t make sense or the person couldn’t think of an answer.
Still, an open-ended question can yield valuable insights into what people are thinking. Open-ended questions are most useful when the researcher needs to know what people are thinking and how they naturally view their world; closed-ended questions are more likely to be used when the dimensions of the variables are well defined.

Schwarz (1999) points out that the two approaches can sometimes lead to different conclusions. He cites the results of a survey question about preparing children for life. When “To think for themselves” was one alternative in a closed-ended list, 62% chose this option; however, only 5% gave this answer when the open-ended format was used. This finding points to the need to have a good understanding of the topic when asking closed-ended questions.

**Number of Response Alternatives**

With closed-ended questions, there are a fixed number of response alternatives. In public opinion surveys, a simple “yes or no” or “agree or disagree” dichotomy is often sufficient. In more basic research, it is often preferable to have a sufficient number of alternatives to allow people to express themselves—for example, a 5- or 7-point scale ranging from “strongly agree to strongly disagree” or “very positive to very negative.” Such a scale might appear as follows:

```
Strongly agree   _______ _______ _______ _______ _______ Strongly disagree
```

**Rating Scales**

Rating scales such as the one just shown are very common in many areas of research. Rating scales ask people to provide “how much” judgments on any number of dimensions—amount of agreement, liking, or confidence, for example. Rating scales can have many different formats. The format that is used depends on factors such as the topic being investigated. Perhaps the best way to gain an understanding of the variety of formats is simply to look at a few examples. The simplest and most direct scale presents people with five or seven response alternatives with the endpoints on the scale labeled to define the extremes. For example,

Students at the university should be required to pass a comprehensive examination to graduate.

```
Strongly agree   _______ _______ _______ _______ _______ Strongly disagree
```

How confident are you that the defendant is guilty of attempted murder?

```
Not at all confident ______ ______ ______ Very confident
```
**Graphic Rating Scale**  A graphic rating scale requires a mark along a continuous 100-millimeter line that is anchored with descriptions at each end.

How would you rate the movie you just saw?
Not very enjoyable __________________________ Very enjoyable

A ruler is then placed on the line to obtain the score on a scale that ranges from 0 to 100.

**Semantic Differential Scale**  The semantic differential scale is a measure of the meaning of concepts that was developed by Osgood and his associates (Osgood, Suci, & Tannenbaum, 1957). Respondents rate any concept—persons, objects, behaviors, ideas—on a series of bipolar adjectives using 7-point scales.

*Smoking cigarettes*

- Good
- Strong
- Active

|   |   |   |   |   |   |   |   | Bad
|---|---|---|---|---|---|---|---|
|   |   |   |   |   |   |   |   | Weak
|   |   |   |   |   |   |   |   | Passive

Research on the semantic differential shows that virtually anything can be measured using this technique. Ratings of specific things (marijuana), places (the student center), people (the governor, accountants), ideas (abortion, tax reduction), and behaviors (attending church, using public transit) can be obtained. A large body of research shows that the concepts are rated along three basic dimensions: the first and most important is *evaluation* (e.g., adjectives such as good–bad, wise–foolish, kind–cruel); the second is *activity* (active–passive, slow–fast, excitable–calm); and the third is *potency* (weak–strong, hard–soft, large–small).

**Nonverbal Scale for Children**  Young children may not understand the types of scales we’ve just described, but they are able to give ratings. For example, you could ask children to “Point to the face that shows how you feel about the toy.”

₀ ₁ ₂ ₃ ₄ ₅

**Labeling Response Alternatives**

The examples thus far have labeled only the endpoints on the rating scale. Respondents decide the meaning of the other response alternatives. This is a
reasonable approach, and people are usually able to use such scales without difficulty. Sometimes researchers need to provide labels to more clearly define the meaning of each alternative. Here is a fairly standard alternative to the agree-disagree scale shown previously:

<table>
<thead>
<tr>
<th>Strongly agree</th>
<th>Agree</th>
<th>Undecided</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
</table>

This type of scale assumes that the middle alternative is a "neutral" point halfway between the endpoints. Sometimes, however, a perfectly balanced scale may not be possible or desirable. Consider a scale asking a college professor to rate a student for a job or graduate program. This particular scale asks for comparative ratings of students:

In comparison with other graduates, how would you rate this student's potential for success?

<table>
<thead>
<tr>
<th>Lower</th>
<th>Upper</th>
<th>Upper</th>
<th>Upper</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>50%</td>
<td>50%</td>
<td>25%</td>
<td>10%</td>
<td>5%</td>
</tr>
</tbody>
</table>

Notice that most of the alternatives are asking people to make a rating in terms of the top 25% of students. This is done because students who apply for such programs tend to be very bright and motivated, and so professors rate them favorably. The wording of the alternatives attempts to force the raters to make finer distinctions among generally very good students.

Labeling alternatives is particularly interesting when asking about the frequency of a behavior. For example, you might ask, "How often do you exercise for at least 20 minutes?" What kind of scale should you use to let people answer this question? You could list (1) never, (2) rarely, (3) sometimes, (4) frequently. These terms convey your meaning but they are vague. Here is another set of alternatives, similar to ones described by Schwarz (1999):

- less than twice a week
- about twice a week
- about four times a week
- about six times a week
- at least once each day

A different scale might be:

- less than once per month
- about once a month
Schwarz (1999) calls the first scale a high-frequency scale because most alternatives indicate a high frequency of exercise. The other scale is referred to as low frequency. Schwarz points out that the labels should be chosen carefully because people may interpret the meaning of the scale differently, depending on the labels used. If you were actually asking the exercise question, you might decide on alternatives different from the ones described here. Moreover, your choice should be influenced by factors such as the population you are studying. If you are studying people who generally exercise a lot, you will be more likely to use a higher-frequency scale than you would if you were studying people who generally don’t exercise a great deal.

FINALIZING THE QUESTIONNAIRE

Formatting the Questionnaire

The printed questionnaire should appear attractive and professional. It should be neatly typed and free of spelling errors. Respondents should find it easy to identify the questions and the response alternatives to the questions. Leave enough space between questions so people don’t become confused when reading the questionnaire. If you have a particular scale format, such as a 5-point rating scale, use it consistently. Don’t change from 5- to 4- to 7-point scales, for example.

It is also a good idea to carefully consider the sequence in which you will ask your questions. In general, it is best to ask the most interesting and important questions first to capture the attention of your respondents and motivate them to complete the survey. Roberson and Sundstrom (1990) obtained the highest return rates in an employee attitude survey when important questions were presented first and demographic questions were asked last. In addition, it is a good idea to group questions together when they address a similar theme or topic. Doing so will make your survey appear more professional, and your respondents will be more likely to take it seriously.

Refining Questions

Before actually administering the survey, it is a good idea to give the questions to a small group of people and have them “think aloud” while answering them. The participants might be chosen from the population being studied, or they could be friends or colleagues who can give reasonable responses to the questions. For the “think aloud” procedure, you will need to ask the individuals to tell you how they interpret each question and how they respond to the response alternatives. This procedure can provide valuable information that you can use to improve the questions. (The importance of pilot studies such as this is discussed further in Chapter 9.)