

# CHAPTER 8

## Survey Research

### RESEARCH THAT MATTERS, QUESTIONS THAT COUNT

#### Survey Research in the Social Sciences

Attractions of Survey Research

*Versatility*

*Efficiency*

*Generalizability*

The Omnibus Survey

Errors in Survey Research

#### Writing Survey Questions

Avoid Confusing Phrasing

Minimize the Risk of Bias

Maximize the Utility of

Response Categories

Avoid Making Either Disagreement or Agreement

Disagreeable

Minimize Fence-Sitting and Floating

#### Combining Questions in Indexes

#### Designing Questionnaires

Build on Existing Instruments

Refine and Test Questions

Add Interpretive Questions

### CAREERS AND RESEARCH

Maintain Consistent Focus

### RESEARCH IN THE NEWS: WHAT CAN SURVEYS UNCOVER?

Order the Questions

Make the Questionnaire Attractive

Consider Translation

#### Organizing Surveys

Mailed, Self-Administered Surveys

Group-Administered Surveys

Telephone Surveys

*Reaching Sample Units*

*Maximizing Response to Phone Surveys*

In-Person Interviews

*Balancing Rapport and Control*

*Maximizing Response to Interviews*

Electronic Surveys

Mixed-Mode Surveys

A Comparison of Survey Designs

#### Ethical Issues in Survey Research

#### Conclusions

## Research That Matters, Questions That Count

Functional limitations caused by aging are associated with more psychological distress, whereas being married is associated with less psychological distress. But does marriage reduce the adverse effects of functional limitations? Alex Bierman at the University of Calgary sought to answer this question with data collected in the Aging, Status, and Sense of Control survey (ASOC) by the Survey Research Laboratory at the University of Illinois at Urbana-Champaign. ASOC used a longitudinal panel design that attempted to survey the same respondents in 1995, 1998, and 2001. There were 966 respondents aged 60 and older in the first wave and 907 who participated in all three waves.

The ASOC measured psychological distress with responses to seven questions that make up a version of the Center for Epidemiological Studies Depression Scale. These questions ask about the number of times in the previous week the respondent had trouble getting to sleep, felt that everything was an effort, felt they couldn't get going, etc. Functional limitations were measured with respondents' ratings of the difficulty they experience with such common tasks as climbing stairs, kneeling, and shopping. Bierman's analysis of the ASOC data indicated that married respondents tended to feel less psychological distress in relation to functional limitations they experienced than did unmarried respondents. This "protective effect" of marriage was stronger for men than for women.

1. Do the CES-D questions make sense to you as indicators of psychological distress? Everyone has feelings like this at some times, but do you think that the range of such feelings experienced could be used to indicate variation over a range from feeling good to feeling distressed? Explain your reasoning.
2. Bierman sought to determine whether marriage mitigates the effects of functional limitations on psychological distress. What would be the advantages and disadvantages of conducting a survey about these issues using in-person interviews, a self-administered paper questionnaire, a phone survey, or a survey on the web?
3. Thinking of the preceding chapter, can you imagine any way to study this effect of marriage using an experimental design?

In this chapter, you will learn how to write survey questions and how to design survey projects. You will also learn more about the ASOC. By the end of the chapter, you will know about the major challenges involved in survey research and how they can be reduced by adhering to guidelines for writing questions and using survey designs that match a survey's purpose. After you finish the chapter, test yourself by reading the 2012 *Society and Mental Health* article by Alex Bierman at the *Investigating the Social World* study site and completing the related interactive exercises for Chapter 8 at [edge.sagepub.com/schutt8e](http://edge.sagepub.com/schutt8e).

Bierman, Alex. 2012. "Functional Limitations and Psychological Distress: Marital Status as Moderator." *Society and Mental Health* 2(1):35–52.

“Education forms a unique dimension of social status, with qualities that make it especially important to health.” John Mirowsky and Catherine Ross (2003:1) make this claim at the start of *Education, Social Status, and Health* and then present evidence to support it throughout the book. Most of their evidence comes from two surveys. In this chapter, we will focus on one of them, the Aging, Status, and the Sense of Control (ASOC) survey that was funded by the National Institute on Aging.

I begin this chapter with a brief review of the reasons for using survey methods, but I will then focus attention on the Mirowsky and Ross ASOC survey and use it to illustrate some key features of survey research. Next, I will discuss guidelines for writing survey questions—a concern in every type of survey research. I will then explain the major steps in questionnaire design and discuss the features of five types of surveys, highlighting the unique problems attending each one and suggesting some possible solutions. I will give particular attention to the ways in which new means of communication such as cell phones and the Internet have been changing



Journal Link

Psychological Distress

survey research since the first ASOC survey in 1995 (there have been two more, in 1998 and 2001). I discuss ethics issues in the final section. By the chapter's end, you should be well on your way to becoming an informed consumer of survey reports and a knowledgeable developer of survey designs. As you read the chapter, I also hope that you will occasionally reflect on how education influences social status and health.

## ▣ Survey Research in the Social Sciences

**Survey research** involves the collection of information from a sample of individuals through their responses to questions. Mirowsky and Ross (2003) turned to survey research for their study of education, social status, and health because it is an efficient method for systematically collecting data from a broad spectrum of individuals and social settings. As you probably have observed, a great many social scientists—as well as newspaper editors, political pundits, government agencies, and marketing gurus—make the same methodological choice. In fact, surveys have become a multibillion-dollar industry in the United States that shapes what we read in the newspapers, see on TV, and find in government reports (Converse 1984; Tourangeau 2004:776).

**Survey research:** Research in which information is obtained from a sample of individuals through their responses to questions about themselves or others.

### Attractions of Survey Research

Survey research owes its popularity to three features: versatility, efficiency, and generalizability. Each of these features is changing as a result of new technologies.

#### Versatility

First, survey methods are versatile. Although a survey is not the ideal method for testing all hypotheses or learning about every social process, a well-designed survey can enhance our understanding of just about any social issue. Mirowsky and Ross's (2003) survey covered a range of topics about work and health, and there is hardly any other topic of interest to social scientists that has not been studied at some time with survey methods. Politicians campaigning for election use surveys, as do businesses marketing a product, governments assessing community needs, agencies monitoring program effectiveness, and lawyers seeking to buttress claims of discrimination or select favorable juries.

Computer technology has made surveys even more versatile. Computers can be programmed so that different types of respondents are asked different questions. Short videos or pictures can be presented to respondents on a computer screen. An interviewer may give respondents a laptop on which to record their answers to sensitive personal questions, such as about illegal activities, so that not even the interviewer will know what they said (Tourangeau 2004:788–794).

#### Efficiency

Surveys also are popular because data can be collected from many people at relatively low cost and, depending on the survey design, relatively quickly. John Mirowsky and Catherine Ross (2003:207) contracted with the Survey Research Laboratory (SRL) of the University of Illinois for their 25-minute 2003 telephone survey of 2,495 adult Americans. SRL estimated that the survey would incur direct costs of \$183,000—that's \$73.35 per respondent—and take as long as 1 year to complete. Both this cost and the length of time required were relatively high because SRL made special efforts to track down respondents from the first wave of interviews in 1995. One-shot telephone interviews can cost as little as \$30 per subject (Ross 1990). Large mailed surveys cost even less, about \$10 to \$15 per potential respondent, although the costs can increase greatly when intensive follow-up efforts are made. Surveys of the general population using personal interviews are much more expensive, with costs ranging from about \$100 per potential respondent, for studies in a limited geographical area, to



**Audio Link**  
U.S. Census Surveys

\$300 or more when lengthy travel or repeat visits are needed to connect with respondents (F. Fowler, personal communication, January 7, 1998; see also Dillman 1982; Groves & Kahn 1979). Surveys through the web have become the quickest way to gather survey data, but there are problems with this method that I will soon discuss.

Surveys are efficient because many variables can be measured without substantially increasing the time or cost. Mailed questionnaires can include as many as 10 pages of questions before respondents begin to balk. In-person interviews can be much longer. For example, the 2012 General Social Survey (GSS) had three versions in English and Spanish that measured 799 variables for the 1,794 cases that were newly interviewed in that year, and totaled 234 pages—although many sections applied only to particular respondents (National Opinion Research Center [NORC] 2014). The upper limit for phone surveys seems to be about 45 minutes.

Of course, these efficiencies can be attained only in a place with a reliable communications infrastructure (Labaw 1980:xiii–xiv). A reliable postal service, which is required for mail surveys, generally has been available in the United States—although residents of the Bronx, New York, have complained that delivery of local first-class mail often takes 2 weeks or more, almost ruling out mail surveys (Purdy 1994). The British postal service, the Royal Mail, has been accused of even worse performance: a “total shambles,” with mail abandoned in some cases and purposely misdelivered in other cases (Lyll 2004:A4). Phone surveys have been very effective in countries such as the United States, where 96% of households have phones (Tourangeau 2004:777). Also important to efficiency are the many survey research organizations—about 120 academic and nonprofit organizations in the United States—that provide trained staff and proper equipment (Survey Research Laboratory 2008).

Modern information technology has been a mixed blessing for survey efficiency. The Internet makes it easier to survey some populations, but it leaves out important segments. Caller ID and answering machines make it easy to screen out unwanted calls, but these tools also make it harder to reach people in phone surveys. In addition, as discussed in Chapter 5, a growing number of people use only cell phones. As a result, after a long decline to below 5% in 2001, the percentage of U.S. households without landline telephones climbed to 29% by 2011, and then to 40% by 2013 (Christian et al. 2010; McGeeney & Keeter 2014). U.S. Census Bureau 2013b) (see Exhibit 8.1). Survey researchers must spend more time and money to reach potential respondents (Tourangeau 2004:781–782).

### Generalizability

Survey methods lend themselves to probability sampling from large populations. Thus, survey research is very appealing when sample generalizability is a central research goal. In fact, survey research is often the only means available for developing a representative picture of the attitudes and characteristics of a large population.

Surveys are also the method of choice when cross-population generalizability is a key concern, because they allow a range of social contexts and subgroups to be sampled. The consistency of relationships can then be examined across the various subgroups. An ambitious Internet-based international survey sponsored by the National Geographic Society (2000) was completed by 80,012 individuals from 178 countries and territories.

Unfortunately (for survey researchers), the new technologies that are lowering the overall rate of response to phone surveys are also making it more difficult to obtain generalizable samples. Although in the United States, only 14% of households in 2013 didn't have access to the Internet at home or work, in these households persons tend to be elderly, poor, and have no more than a high school education compared to those who are “connected” (de Leeuw 2008:321; Pew Research Center 2014; Tourangeau 2004:792; U.S. Census Bureau 2013b).

Those who rely exclusively on cell phones tend to be younger and poorer than are those who also have landline phones. In addition, cell phone-only households are more likely in some states and regions than others and they are more likely to be Hispanic, compared with households with landlines (AAPOR 2014b).

Another challenge in survey research is the growing foreign-born population in the United States, 13% in 2012, requires foreign-language versions of survey forms; survey results cannot be generalized to the entire population (Grieco et al. 2012:2; Tourangeau 2004:783).

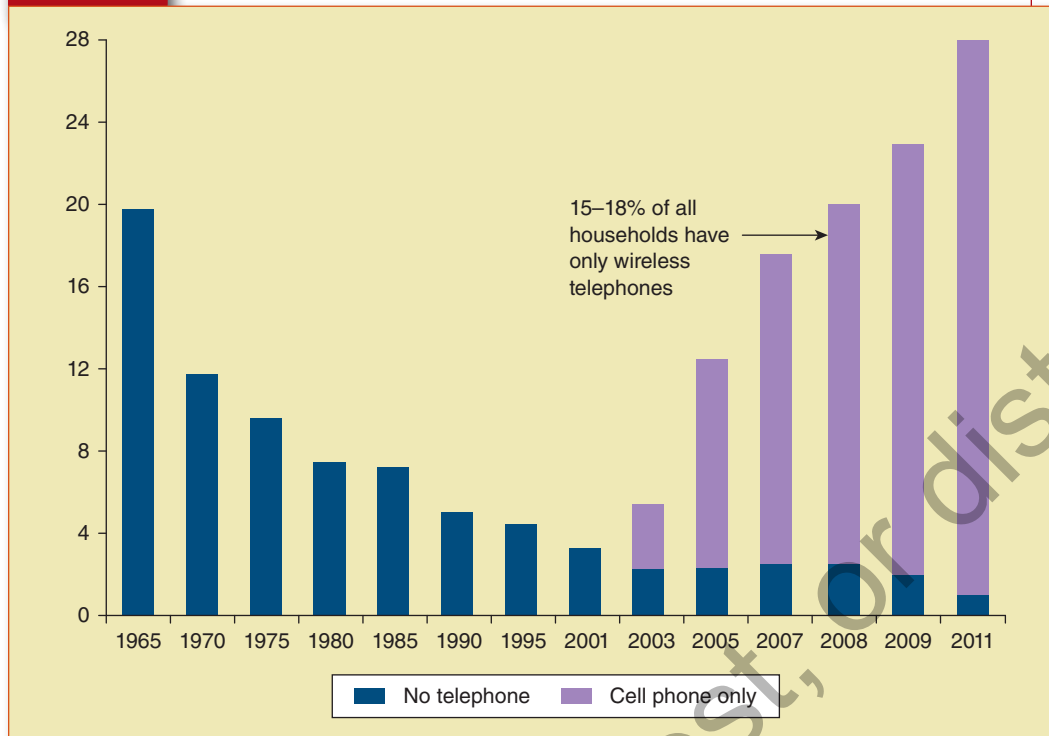


Research | Social Impact  
Link

Survey Methods and Results

## Exhibit 8.1

## Percentage of U.S. Households Without Telephones and With Cell Phones Only



*Sources:* Adapted from Link, Michael. 2008. "Solving the Problems Cell Phones Create for Survey Research." Presentation to the Harvard Program on Survey Research Spring Conference, New Technologies and Survey Research. Cambridge, MA: Institute of Quantitative Social Science, Harvard University, May 9; Christian, Leah, Scott Keeter, Kristen Purcell, and Aaron Smith. 2010. "Assessing the Cell Phone Challenge to Survey Research in 2010." Washington, DC: Pew Research Center for the People & the Press and Pew Internet & American Life Project.

## The Omnibus Survey

An omnibus survey shows just how versatile, efficient, and generalizable a survey can be. An **omnibus survey** covers a range of topics of interest to different social scientists, in contrast to the typical survey that is directed at a specific research question. The omnibus survey has multiple sponsors or is designed to generate data useful to a broad segment of the social science community rather than to answer a particular research question. It is usually directed to a sample of some general population, so the questions, about a range of different issues, are appropriate to at least some sample members.

One of sociology's most successful omnibus surveys is the GSS of the National Opinion Research Center at the University of Chicago. It is an extensive interview administered biennially to a probability sample of at least 3,000 Americans (4,820 in 2012), with a wide range of questions and topic areas chosen by a board of overseers. Some questions are asked of only a randomly selected subset of respondents. This **split-ballot design** allows more questions without increasing the survey's cost. It also facilitates experiments on the effect of question wording: Different forms of the same question are included in the split-ballot subsets. The GSS is widely available to universities, instructors, and

**Omnibus survey:** A survey that covers a range of topics of interest to different social scientists.

**Split-ballot design:** Unique questions or other modifications in a survey administered to randomly selected subsets of the total survey sample, so that more questions can be included in the entire survey or so that responses to different question versions can be compared.

students (Davis & Smith 1992; National Opinion Research Center 2011), as are many other survey data sets archived by the Inter-University Consortium for Political and Social Research (ICPSR) (more details about the ICPSR are in Chapter 13). Mirowsky and Ross contributed their survey data set to the ICPSR.

## Errors in Survey Research

It might be said that surveys are too easy to conduct. Organizations and individuals often decide that a survey will help solve some important problem because it seems so easy to write up some questions and distribute them. But without careful attention to sampling, measurement, and overall survey design, the effort is likely to be a flop. Such flops are too common for comfort, so the responsible survey researcher must take the time to design surveys properly and to convince sponsoring organizations that this time is worth the effort (Turner & Martin 1984:68).

For a survey to succeed, it must minimize four types of error (Groves 1989:vi, 10–12): (1) poor measurement, (2) nonresponse, (3) inadequate coverage of the population, and (4) sampling error.

**Poor measurement.** Measurement error was a key concern in Chapter 4, but there is much more to be learned about how to minimize these errors of observation in the survey process. The theory of *satisficing* can help us understand the problem. It takes effort to answer survey questions carefully: Respondents have to figure out what each question means, then recall relevant information, and finally decide which answer is most appropriate. Survey respondents *satisfice* when they reduce the effort required to answer a question by interpreting questions superficially and giving what they think will be an acceptable answer (Krosnick 1999:547–548). Presenting clear and interesting questions in a well-organized questionnaire will help reduce measurement error by encouraging respondents to answer questions carefully and to take seriously the request to participate in the survey. Tailoring questions to the specific population surveyed is also important. In particular, persons with less education are more likely to satisfice in response to more challenging questions (Holbrook, Green, & Krosnick 2003; Narayan & Krosnick 1996).

Errors in measurement also arise when respondents are unwilling to disclose their feelings and behaviors, unable to remember past events, and misunderstand survey questions. What people say they can do—such as ability to carry out various tasks—is not necessarily consistent with what they are able to do (Schutt 2011b:88). What people report that they have done is not necessarily what they have actually done (Brenner 2012). Careful assessment of survey question quality is thus an essential step in survey design. The next section focuses on how to write good survey questions.

**Nonresponse.** Nonresponse is a major and growing problem in survey research, although it is a problem that varies between particular survey designs. Social exchange theory can help us understand why nonresponse rates have been growing in the United States and Western Europe since the early 1950s (Dillman 2000:14–15; Groves & Couper 1998:155–189; Tourangeau 2004:782). According to social exchange theory, a well-designed survey effort will maximize the social rewards for survey participation and minimize its costs, as well as establish trust that the rewards will outweigh the costs (Blau 1964). The perceived benefits of survey participation have declined with decreasing levels of civic engagement and with longer work hours (Groves, Singer, & Corning 2000; Krosnick 1999:539–540). Perceived costs have increased with the widespread use of telemarketing and the ability of many people to screen out calls from unknown parties with answering machines and caller ID. In addition, recipients pay for time on cell phone calls, so the ratio of costs to benefits worsens for surveys attempting to reach persons using cell phones (Nagourney 2002). We will review more specifics about nonresponse in this chapter's sections on particular survey methods.



*Inadequate coverage of the population.* A poor sampling frame can invalidate the results of an otherwise well-designed survey. We considered the importance of a good sampling frame in Chapter 5; in this chapter, I will discuss special coverage problems related to each of the particular survey methods.

*Sampling error.* The process of random sampling can result in differences between the characteristics of the sample members and the population simply on the basis of chance. I introduced this as a topic in Chapter 5. You will learn how to calculate sampling error in Chapter 9.

It is most important to maintain a realistic perspective on the nature of surveys to avoid making unrealistic assumptions about the validity of survey results. Although surveys provide an efficient means for investigating a wide range of issues in large and diverse populations, the data they provide is necessarily influenced by these four sources of error. Survey researchers must make every effort to minimize each one. Only through learning more about different survey features and survey research alternatives can we prepare to weigh the advantages and disadvantages of survey research in particular circumstances and thus assess the value of a survey design in relation to a specific research question.

## Writing Survey Questions

Questions are the centerpiece of survey research. Because the way they are worded can have a great effect on the way they are answered, selecting good questions is the single most important concern for survey researchers. All hope for achieving measurement validity is lost unless the questions in a survey are clear and convey the intended meaning to respondents.

You may be thinking that you ask people questions all the time and have no trouble understanding the answers you receive, but can't you also think of times when you've been confused in casual conversation by misleading or misunderstood questions? Now, consider just a few of the differences between everyday conversations and standardized surveys that make writing survey questions much more difficult:

- Survey questions must be asked of many people, not just one.
- The same survey question must be used with each person, not tailored to the specifics of a given conversation.
- Survey questions must be understood in the same way by people who differ in many ways.
- You will not be able to rephrase a survey question if someone doesn't understand it because that would result in a different question for that person.
- Survey respondents don't know you and so can't be expected to share the nuances of expression that help you and your friends and family to communicate.

Writing questions for a particular survey might begin with a brainstorming session or a review of previous surveys. Then, whatever questions are being considered must be systematically evaluated and refined. Although most professionally prepared surveys contain previously used questions as well as some new ones, every question that is considered for inclusion must be reviewed carefully for its clarity and ability to convey the intended meaning. Questions that were clear and meaningful to one population may not be so to another. Nor can you simply assume that a question used in a previously published study was carefully evaluated.

Adherence to a few basic principles will go a long way toward ensuring clear and meaningful questions. Each of these principles summarizes a great deal of research, although none of them should be viewed as an inflexible mandate (Alwin & Krosnick 1991). As you will learn in the next section, every question must be considered relative to the other questions in a survey. Moreover, every survey has its own unique requirements and constraints; sometimes violating one principle is necessary to achieve others.



**Interactive Exercises**  
**Link**  
Survey Research

## Avoid Confusing Phrasing

What's a confusing question? Try this one that I received years ago from the Planetary Society in its National Priorities Survey for United States Space Program:

The Moon may be a place for an eventual scientific base, and even for engineering resources. Setting up a base or mining experiment will cost tens of billions of dollars in the next century. Should the United States pursue further manned and unmanned scientific research projects on the surface of the Moon?

Yes       No       No opinion

Does a “yes” response mean that you favor spending tens of billions of dollars for a base or mining experiment? Does “the next century” refer to the 21st century or to the 100 years after the survey (which was distributed in the 1980s)? Could you favor further research projects on the Moon but oppose funding a scientific base or engineering resources? Are engineering resources supposed to have something to do with a mining experiment? Does a mining experiment occur “on the surface of the Moon”? How do you answer if you favor unmanned scientific research projects on the Moon but not manned projects?

There are several ways to avoid such confusing phrasing. In most cases, a simple direct approach to asking a question minimizes confusion. Use shorter rather than longer words and sentences: “brave” rather than “courageous”; “job concerns” rather than “work-related employment issues” (Dillman 2000:52). Try to keep the total number of words to 20 or fewer and the number of commas to 3 or fewer (Peterson 2000:50). However, questions shouldn't be abbreviated in a way that results in confusion: To ask, “In what city or town do you live?” is to focus attention clearly on a specific geographic unit, a specific time, and a specific person (you); the simple format,

Residential location: \_\_\_\_\_,

does not do this.

Sometimes, when sensitive issues or past behaviors are the topic, longer questions can provide cues that make the respondent feel comfortable or aid memory (Peterson 2000:51).

Breaking up complex issues into simple parts also reduces confusion. In a survey about health services (such as the one by Mirowsky and Ross), you might be tempted to ask a complex question like this (Schaeffer & Presser 2003):

During the past 12 months since July 1st, 1987, how many times have you seen or talked with a doctor or a medical assistant about your health? Do not count any times you might have seen a doctor while you were a patient in a hospital, but count all the other times you actually saw or talked to a medical doctor of any kind about your health. (pp. 70–71)

This question can be simplified, thereby reducing confusion, by breaking it up into several shorter questions:

Have you been a patient in the hospital overnight in the past 12 months since July 1st, 1987?

(Not counting when you were in a hospital overnight) During the past 12 months since July 1st, 1987, how many times did you actually see any medical doctor about your own health?

During the past 12 months since July 1st, 1987, were there any times when you didn't actually see the doctor but saw a nurse or other medical assistant working for the doctor?

During the past 12 months since July 1st, 1987, did you get any medical advice, prescriptions, or results of tests over the telephone from a medical doctor, nurse, or medical assistant working for a doctor? (Cannell et al. 1989:Appendix A, p. 1)



A sure way to muddy the meaning of a question is to use **double negatives**: “Do you *disagree* that there should *not* be a tax increase?” Respondents have a hard time figuring out which response matches their sentiments. Such errors can easily be avoided with minor wording changes, but even experienced survey researchers can make this mistake unintentionally, perhaps while trying to avoid some other wording problem. For instance, in a survey commissioned by the American Jewish Committee, the Roper polling organization wrote a question about the Holocaust that was carefully worded to be neutral and value free: “Does it seem possible or does it seem impossible to you that the Nazi extermination of the Jews never happened?” Among a representative sample of adult Americans, 22% answered that it was possible the extermination never happened (Kifner 1994:A12). Many Jewish leaders and politicians were stunned, wondering how one in five Americans could be so misinformed. But a careful reading of the question reveals how confusing it is: Choosing “possible,” the seemingly positive response, means that you don’t believe the Holocaust happened. In fact, the Gallup organization then rephrased the question to avoid the double negative, giving a brief definition of the Holocaust and then asking, “Do you doubt that the Holocaust actually happened or not?” Only 9% responded that they doubted it happened. When a wider range of response choices was given, only 2.9% said that the Holocaust “definitely” or “probably” did not happen. To be safe, it’s best just to avoid using negative words such as “don’t” and “not” in questions.

**Double negative:** A question or statement that contains two negatives, which can muddy the meaning of the question.

So-called **double-barreled questions** are also guaranteed to produce uninterpretable results because they actually ask two questions but allow only one answer. For example, during the Watergate scandal, Gallup poll results indicated that when the question was “Do you think President Nixon should be impeached and compelled to leave the presidency, or not?” only about a third of Americans supported impeaching President Richard M. Nixon. But when the Gallup organization changed the question to ask respondents if they “think there is enough evidence of possible wrongdoing in the case of President Nixon to bring him to trial before the Senate, or not,” over half answered yes. Apparently, the first, double-barreled version of the question confused support for impeaching Nixon—putting him on trial before the Senate—with concluding that he was guilty before he had had a chance to defend himself (Kagay & Elder 1992:E5).

**Double-barreled question:** A single survey question that actually asks two questions but allows only one answer.

It is also important to identify clearly what kind of information each question is to obtain. Some questions focus on attitudes, or what people say they want or how they feel. Some questions focus on beliefs, or what people think is true. Some questions focus on behavior, or what people do. And some questions focus on attributes, or what people are like or have experienced (Dillman 1978:79–118; Gordon 1992). Rarely can a single question effectively address more than one of these dimensions at a time.

Whichever type of information a question is designed to obtain, be sure it is asked of only the respondents who may have that information. If you include a question about job satisfaction in a survey of the general population, first ask respondents whether they have a job. You will only annoy respondents if you ask a question that does not apply to them (Schaeffer & Presser 2003:74). These **filter questions** create **skip patterns**. For example, respondents who answer *no* to one question are directed to skip ahead to another question, but respondents who answer *yes* go on to the **contingent question**. Skip patterns should be indicated clearly with an arrow or other mark in the questionnaire as demonstrated in Exhibit 8.2.

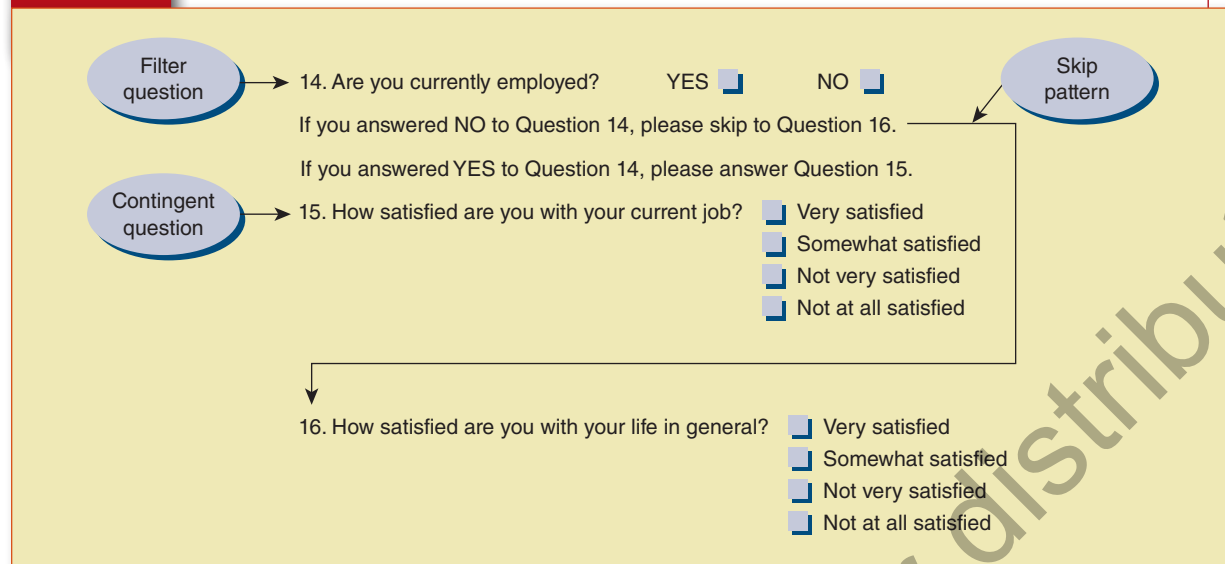
**Filter question:** A survey question used to identify a subset of respondents who then are asked other questions.

**Skip pattern:** The unique combination of questions created in a survey by filter questions and contingent questions.

**Contingent question:** A question that is asked of only a subset of survey respondents.

## Minimize the Risk of Bias

Specific words in survey questions should not trigger biases, unless that is the researcher’s conscious intent. Biased or loaded words and phrases tend to produce misleading answers. For example, a 1974 survey found that 18% of respondents supported sending U.S. troops “if a situation like Vietnam were to develop in another part of the world.” But when the question was reworded to mention sending troops to “stop a communist takeover”—“communist takeover” being a loaded phrase—favorable responses rose to 33% (Schuman & Presser 1981:285).

**Exhibit 8.2 Filter Questions and Skip Patterns**

Answers can also be biased by more subtle problems in phrasing that make certain responses more or less attractive to particular groups. To minimize biased responses, researchers have to test reactions to the phrasing of a question. For example, Mirowsky and Ross (personal e-mail, 2009) wanted to ask people, “Do you feel fit?” However, when the Survey Research Lab tried out this question with a small sample, people did not seem to understand that they were being asked about their level of energy and general feelings of fitness; they just focused on whether they had some type of health problem. It seemed that people had a biased concept of health as involving only problems rather than including the concept of positive health. As a result, Mirowsky and Ross rephrased the question to be more explicit: “Do you feel physically fit?”

Responses can also be biased when response alternatives do not reflect the full range of possible sentiment on an issue. When people pick a response choice, they seem to be influenced by where they are placing themselves relative to the other response choices. For example, the Detroit Area Study (Turner & Martin 1984:252) asked the following question: “People feel differently about making changes in the way our country is run. In order to keep America great, which of these statements do you think is best?” When the only response choices were “We should be very cautious of making changes” and “We should be free to make changes,” only 37% said that we should be free to make changes. However, when a response choice was added that suggested we should “constantly” make changes, 24% picked that response and another 32% chose the “free to make changes” response, for a total of 56% who seemed open to making changes in the way our country is run (Turner & Martin 1984:252). Including the more extreme positive alternative (“constantly” make changes) made the less extreme positive alternative more attractive.

If the response alternatives for a question fall on a continuum from positive to negative, the number of positive and negative categories should be balanced so that one end of the continuum doesn’t seem more attractive than the other (Dillman 2000:57–58). If you ask respondents, “How satisfied are you with the intramural sports program here?” and include “completely satisfied” as the most positive possible response, then “completely dissatisfied” should be included as the most negative possible response. This is called a *bipolar scale*.

Of course, the advice to minimize the risk of bias means nothing to those who conduct surveys to elicit bias. This is the goal of *push polling*, a technique that has been used in some political campaigns. In a push poll, the pollsters for a candidate call potential voters and ask them a series of questions that convey negative information about the opposing candidate. It’s really not a survey at all—just a propaganda effort—but it casts reputable survey research (and ethical political polling firms) in a bad light (Connolly & Manning 2001).

## Maximize the Utility of Response Categories

Response choices should be considered carefully because they help respondents to understand what the question is about and what types of responses are viewed as relevant (Clark & Schober 1994). Questions with fixed response choices must provide one and only one possible response for everyone who is asked the question—that is, the response choices must be exhaustive and mutually exclusive. Ranges of ages, incomes, years of schooling, and so forth should not overlap and should provide a response option for all respondents.

There are two exceptions to this principle: (1) Filter questions may tell some respondents to skip over a question (the response choices do not have to be exhaustive), and (2) respondents may be asked to “check all that apply” (the response choices are not mutually exclusive). Even these exceptions should be kept to a minimum. Respondents to a self-administered paper questionnaire should not have to do a lot of skipping around, or they may lose interest in answering carefully all the applicable questions. Some survey respondents react to a “check all that apply” request by just checking enough responses so that they feel they have “done enough” for that question and then ignoring the rest of the choices (Dillman 2000:63).

Vagueness in the response choices is also to be avoided. Questions about thoughts and feelings will be more reliable if they refer to specific times or events (Turner & Martin 1984:300). Usually a question like “On how many days did you read the newspaper in the last week?” produces more reliable answers than one like “How often do you read the newspaper? (frequently, sometimes, never).” In their survey, Mirowsky and Ross (2001:2) sensibly asked the question “Do you currently smoke 7 or more cigarettes a week?” rather than the vaguer question “Do you smoke?” Of course, being specific doesn’t help if you end up making unreasonable demands of your respondents’ memories. One survey asked, “During the past 12 months, about how many times did you see or talk to a medical doctor?” According to their written health records, respondents forgot 60% of their doctor visits (Goleman 1993b:C11). So unless your focus is on major events that are unlikely to have been forgotten, limit questions about specific past experiences to the past month.

Sometimes, problems with response choices can be corrected by adding questions. For example, if you ask, “How many years of schooling have you completed?” someone who dropped out of high school but completed the requirements for a General Equivalency Diploma (GED) might not be sure how to respond. By asking a second question, “What is the highest degree you have received?” you can provide the correct alternative for those with a GED as well as for those who graduated from high school.

Adding questions may also improve memory about specific past events. Imagine the problem you might have answering the question, “How often did you receive help from classmates while preparing for exams or completing assignments during the last month? (very often, somewhat often, occasionally, rarely, or never).” Now, imagine a series of questions that asks you to identify the exams and assignments you had in the past month and, for each one, inquires whether you received each of several types of help from classmates: study suggestions, study sessions, related examples, general encouragement, and so on. The more specific focus on particular exams and assignments should result in more complete recall (Dykema & Schaeffer 2000).

Response choices should be matched to the question they follow and reflect meaningful distinctions, as well as cover the range of possible responses—another way of saying that the response choices should be mutually exclusive and exhaustive. If the question is “How satisfied are you with your job?,” the response choices should focus on distinctions between levels of satisfaction and might range from “very satisfied” to “somewhat,” “not very,” and “not at all satisfied.” If one response choice is “somewhat satisfied,” “moderately satisfied” would not be a good additional response because it does not reflect a meaningful distinction. When measuring the importance of something to respondents, response choices should include “extremely” as well as “not at all important,” “slightly important,” “moderately important,” and “very important” because people tend to rank many issues as “very important.”

One common approach for measures of attitude intensity is to present a statement and then ask respondents to indicate their degree of agreement or disagreement. The last question in this section, about “my misfortunes,” is an example, using the form known as a **Likert item** (after social psychologist Rensis Likert, who popularized this approach). A Likert item phrases an attitude in terms of one end of

**Likert item:** A statement followed by response choices ranging from “strongly agree” to “strongly disagree.”

a continuum, so that the responses ranging from “strongly agree” to “strongly disagree” cover the full range of possible agreement. However, the risk of agreement bias should be considered carefully when interpreting responses to Likert-style items (see the next section).

**Bipolar response options:** Response choices to a survey question that include a middle category and parallel responses with positive and negative valence (can be labeled or unlabeled).

**Labeled unipolar response options:** Response choices for a survey question that use words to identify categories ranging from low to high (or vv).

**Unlabeled unipolar response options:** Response choices for a survey question that use numbers to identify categories ranging from low to high (or vv).

Other words used to distinguish points on an ordinal scale of attitude intensity are reflected in the response choices in Exhibit 8.3. One important decision is whether to use unipolar distinctions, such as “not at all” to “extremely,” or bipolar distinctions, such as “very comfortable” to “very uncomfortable.” The advantages of using **bipolar response options** are discussed in the next section. How many response categories are desirable? Five categories work well for unipolar ratings, and seven will capture most variation on bipolar ratings (Krosnick 2006; Schaeffer & Presser 2003:78–79). Responses are more reliable when these categories are labeled (**labeled unipolar response options**) rather than identified only by numbers (**unlabeled unipolar response options**) (Krosnick 1999:544; Schaeffer & Presser 2003:78). Exhibit 8.3 shows these alternatives, based on a question in the Mirowsky and Ross (2001) questionnaire and response alternatives. A special consideration for ratings is whether to include a middle, neutral response option. This issue is discussed later.

## Avoid Making Either Disagreement or Agreement Disagreeable

**Social desirability bias:** The tendency to “agree” with a statement just to avoid seeming disagreeable.

People often tend to “agree” with a statement just to avoid seeming disagreeable. This is termed *agreement bias*, **social desirability bias**, or an *acquiescence effect*. You can see the impact of this human tendency in a 1974 Michigan Survey Research Center survey that asked who was to blame for crime and lawlessness in the United States (Schuman & Presser 1981:208). When one question stated that individuals were

more to blame than social conditions, 60% of the respondents agreed. But when the question was rephrased and respondents were asked, in a balanced fashion, whether individuals or social conditions were more to blame, only 46% chose individuals. Numerous studies of agreement bias suggest that about 10% of respondents will “agree” just to be agreeable, without regard to what they really think (Krosnick 1999:553).

You can take several steps to reduce the likelihood of agreement bias. As a general rule, you should present both sides of attitude scales in the question itself (Dillman 2000:61–62): “In general, do you believe that *individuals* or *social conditions* are more to blame for crime and lawlessness in the United States?” The response choices themselves should be phrased to make each one seem as socially approved, as “agreeable,” as the others. You should also consider replacing a range of response alternatives that focus on the word *agree* with others. For example, “To what extent do you support or oppose the new health care plan?” (response choices range from “strongly support” to “strongly oppose”) is probably a better approach than the question “To what extent do you agree or disagree with the statement: ‘The new health care plan is worthy of support?’” (response choices range from “strongly agree” to “strongly disagree”). For the same reason, simple true–false and yes–no response choices should be avoided (Schaeffer & Presser 2003:80–81).

You may also gain a more realistic assessment of respondents’ sentiment by adding to a question a counterargument in favor of one side to balance an argument in favor of the other side. Thus, don’t just ask in an employee survey whether employees should be required to join the union; instead, ask whether employees should be required to join the union or be able to make their own decision about joining. In one survey, 10% more respondents said they favored mandatory union membership when the counterargument was left out than when it was included. It is reassuring to know, however, that this approach does not change the distribution of answers to questions about which people have very strong beliefs (Schuman & Presser 1981:186).

When an illegal or socially disapproved behavior or attitude is the focus, we have to be concerned that some respondents will be reluctant to agree that they have ever done or thought such a thing. In this situation, the goal is to write a question and response choices that make agreement seem more acceptable. For example, Dillman

**Exhibit 8.3** Labeled Unipolar, Unlabeled Unipolar, and Bipolar Response Options

## Original

>Q72a< How free do you feel to disagree with the person who supervises your work? Are you . . .

- <1> Not at all free,
- <2> Somewhat free,
- <3> Largely but not completely free, or
- <4> Completely free to disagree?
- <7> NO CODED RESPONSE APPLICABLE
- <8> DON'T KNOW
- <9> REFUSED

## Labeled unipolar version

How comfortable do you feel disagreeing with the person who supervises your work? (Please circle one number to indicate your response.)

1. Extremely comfortable
2. Very comfortable
3. Quite comfortable
4. Somewhat comfortable
5. Not at all comfortable

## Labeled bipolar version

Do you feel comfortable or uncomfortable disagreeing with the person who supervises your work? (Please circle one number to indicate your response.)

1. Very comfortable
2. Mostly comfortable
3. Slightly comfortable
4. Feel neither comfortable nor uncomfortable
5. Slightly uncomfortable
6. Mostly uncomfortable
7. Very uncomfortable

## Unlabeled unipolar version

Please circle a number from 1 to 10 to indicate how comfortable you feel disagreeing with the person who supervises your work. 1 means "not at all comfortable" and 10 means "extremely comfortable."

How comfortable do you feel disagreeing with the person who supervises your work?

Not at all

Extremely

1      2      3      4      5      6      7      8      9      10

Source: Based on Mirowsky and Ross (2001:9).

(2000:75) suggests that we ask, "Have you ever taken anything from a store without paying for it?" rather than "Have you ever shoplifted something from a store?" Asking about a variety of behaviors or attitudes that range from socially acceptable to socially unacceptable will also soften the impact of agreeing with those that are socially unacceptable.



## Minimize Fence-Sitting and Floating

Two related problems in writing survey questions also stem from people's desire to choose an acceptable answer. There is no uniformly correct solution to these problems; researchers have to weigh the alternatives in light of the concept to be measured and whatever they know about the respondents.

**Fence-sitters:** Survey respondents who see themselves as being neutral on an issue and choose a middle (neutral) response that is offered.

**Floaters:** Survey respondents who provide an opinion on a topic in response to a closed-ended question that does not include a "Don't know" option, but who will choose "Don't know" if it is available.

**Fence-sitters**, people who see themselves as being neutral, may skew the results if you force them to choose between opposites. In most cases, about 10% to 20% of such respondents—those who do not have strong feelings on an issue—will choose an explicit middle, neutral alternative (Schuman & Presser 1981:161–178). Having an explicit neutral response option is generally a good idea: It identifies fence-sitters and tends to increase measurement reliability (Schaeffer & Presser 2003:78).

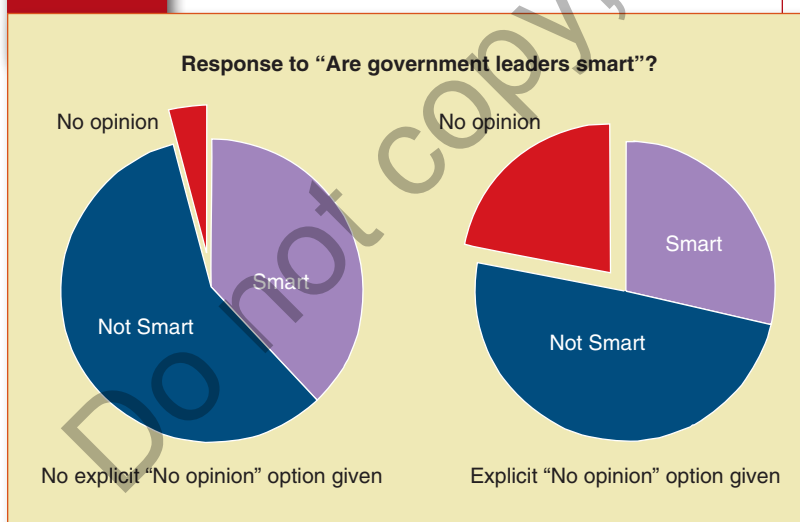
Even more people can be termed **floaters**: respondents who choose a substantive answer when they really don't know or have no opinion. A third of the public will provide an opinion on a proposed law that they know nothing about if they are asked for their opinion in a closed-ended survey question that does not include "Don't know" as an explicit response choice. However, 90% of these persons will select the "Don't know" response if they are explicitly given that option. On average, offering an explicit response option increases the "Don't know" responses by about a fifth (Schuman & Presser 1981:113–160).

Exhibit 8.4 depicts the results of one study that tested the effect of giving respondents an explicit "No opinion" option to the question "Are government leaders smart?" Notice how many more people chose "No opinion" when they were given that choice than when their only explicit options were "Smart" and "Not smart."

**Forced-choice questions:** Closed-ended survey questions that do not include "Don't know" as an explicit response choice.

Despite the prevalence of floating, people often have an opinion but are reluctant to express it. Actually, most political pollsters use **forced-choice questions** without a "Don't know" option. Just after President Clinton's victory, Frank Newport, editor in chief of the Gallup poll, defended pollsters' efforts to get all prospective voters to declare a preferred candidate:

**Exhibit 8.4** The Effect of Floaters on Public Opinion Polls



Source: Based on Schuman and Presser (1981:121).

It would not be very instructive for pollsters . . . to allow large numbers of voters to claim they are undecided all through the election season. We would miss the dynamics of change, we would be unable to tell how well candidates were doing in response to events, and publicly released polls would be out of synchronization with private, campaign polls. (Newport 1992:A28)

Because there are so many floaters in the typical survey sample, the decision to include an explicit "Don't know" option for a question is important. Unfortunately, the inclusion of an explicit "Don't know" response



choice leads some people who do have a preference to take the easy way out—to satiate—and choose “Don’t know.” This is particularly true in surveys of less-educated populations—except for questions that are really impossible to decipher, to which more educated persons are likely to say they “don’t know” (Schuman & Presser 1981:113–146). As a result, survey experts now recommend that questions not include “Don’t know” or “No opinion” options (Krosnick 1999:558; Schaeffer & Presser 2003:80). Adding an open-ended question in which respondents are asked to discuss their opinions can help identify respondents who are floaters (Smith 1984).

Researchers who use in-person or telephone interviews (rather than self-administered questionnaires) may get around the dilemma somewhat by reading the response choices without a middle or “Don’t know” alternative but recording a noncommittal response if it is offered. Mirowsky and Ross’s (2001) questionnaire for their phone survey about education and health included the following example (responses in ALL CAPS were not read):

My misfortunes are the result of mistakes I have made.  
(Do you . . . )

1. <1> Strongly agree,
2. <2> Agree,
3. <3> Disagree, or
4. <4> Strongly disagree?
5. <7> NO CODED RESPONSE APPLICABLE
6. <8> DON’T KNOW
7. <9> REFUSED

## ▣ Combining Questions in Indexes

Writing single questions that yield usable answers is always a challenge. Simple though they may seem, single questions are prone to error because of **idiosyncratic variation**, which occurs when individuals’ responses vary because of their reactions to particular words or ideas in the question. Differences in respondents’ backgrounds, knowledge, and beliefs almost guarantee that some will understand the same question differently.

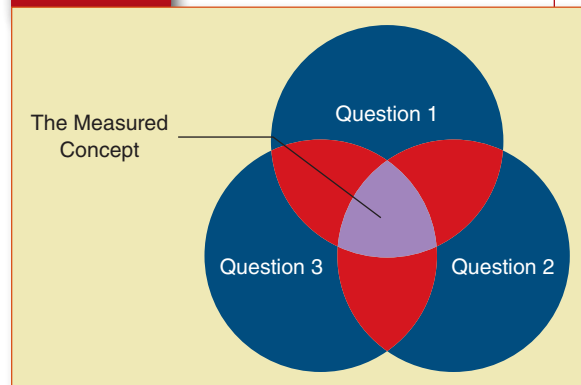
In some cases, the effect of idiosyncratic variation can be dramatic. For example, when people were asked in a survey whether they would “forbid” public speeches against democracy, 54% agreed. When the question was whether they would “not allow” public speeches against democracy, 75% agreed (Turner & Martin 1984:chap. 5). Respondents are less likely to respond affirmatively to the question, “Did you see a broken headlight?” than they are to the question, “Did you see *the* broken headlight?” (Turner & Martin 1984:chap. 9).

The guidelines in this chapter for writing clear questions should help reduce idiosyncratic variation caused by different interpretations of questions. But the best option is often to develop multiple questions about a concept and then to average the responses to those questions in a composite measure termed an *index* or *scale*.

The idea is that idiosyncratic variation in response to particular questions will average out, so that the main influence on the combined measure will be the concept upon which all the questions focus. The index can be considered a more complete measure of the concept than can any one of the component questions.

**Idiosyncratic variation:** Variation in responses to questions that is caused by individuals’ reactions to particular words or ideas in the question instead of by variation in the concept that the question is intended to measure.

### Exhibit 8.5 Overlapping Dimensions of a Concept



Creating an index is not just a matter of writing a few questions that seem to focus on a concept. Questions that seem to you to measure a common concept might seem to respondents to concern several different issues. The only way to know that a given set of questions does, in fact, form an index is to administer the questions to people like those you plan to study. If a common concept is being measured, people's responses to the different questions should display some consistency. In other words, responses to the different questions should be correlated. Exhibit 8.5 illustrates an index in which responses to the items are correlated; the substantial area of overlap indicates that the questions are measuring a common concept. Special statistics called *reliability measures* help researchers decide whether responses are consistent.

Because of the popularity of survey research, indexes already have been developed to measure many concepts, and some of these indexes have proved to be reliable in a range of studies. It usually is much better to use such an index to measure a concept than to try to devise questions to form a new index. Use of a preexisting index both simplifies the work involved in designing a study and facilitates comparison of findings to those obtained in other studies.

The questions in Exhibit 8.6 are a different form of the index to measure the concept of depression—the Center for Epidemiologic Studies Depression Index (CES-D)—that Bierman (2012) used in the “Research That Matters” article that began this chapter. Many researchers in different studies have found that these questions form a reliable index. Note that each question concerns a symptom of depression. People may have idiosyncratic reasons for having a particular symptom without being depressed; for example, persons

### Exhibit 8.6 Example of an Index: Short Form of the Center for Epidemiologic Studies Depression Index (CES-D)

At any time during the past week. . . (Circle one response on each line)	Never	Some of the Time	Most of the Time
a. Was your appetite so poor that you did not feel like eating?	1	2	3
b. Did you feel so tired and worn out that you could not enjoy anything?	1	2	3
c. Did you feel depressed?	1	2	3
d. Did you feel unhappy about the way your life is going?	1	2	3
e. Did you feel discouraged and worried about your future?	1	2	3
f. Did you feel lonely?	1	2	3

Source: Radloff (1977:387). Copyright 1977 by West Publishing Company/Applied Psychological Measurement, Inc.; reproduced by permission.

who have been suffering a physical ailment may report that they have a poor appetite. But by combining the answers to questions about several symptoms, the index score reduces the impact of this idiosyncratic variation.

Three cautions are in order:

1. *Our presupposition that each component question is indeed measuring the same concept may be mistaken.* Although we may include multiple questions in a survey to measure one concept, we may find that answers to the questions are not related to one another, and so the index cannot be created. Alternatively, we may find that answers to just a few of the questions are not related to the answers given to most of the others. We may therefore decide to discard these particular questions before computing the average that makes up the index.

2. *Combining responses to specific questions can obscure important differences in meaning among the questions.* My research on the impact of AIDS prevention education in shelters for the homeless provides an example. In this research, I asked a series of questions to ascertain respondents' knowledge about HIV risk factors and about methods of preventing exposure to those risk factors. I then combined these responses into an overall knowledge index. I was somewhat surprised to find that the knowledge index scores were no higher in a shelter with an AIDS education program than in a shelter without such a program. However, further analysis showed that respondents in the shelter with an AIDS education program were more knowledgeable than the other respondents about the specific ways of preventing AIDS, which were, in fact, the primary focus of the program. Combining responses to these questions with the others about general knowledge of HIV risk factors obscured an important finding (Schutt, Gunston, & O'Brien 1992).

3. *The questions in an index may cluster together in subsets.* All the questions may be measuring the intended concept, but we may conclude that this concept actually has several different aspects. A multidimensional index has been obtained. This conclusion can, in turn, help us refine our understanding of the original concept. For example, Carlo DiClemente and colleagues (1994) sought to determine how confident individuals in treatment for alcoholism were that they could abstain from drinking in different situations that presented typical drinking cues. The 20 situations they presented were of four different types: (1) negative affect, (2) social/positive, (3) physical and other concerns, or (4) withdrawal and urges. The questions used to measure these different dimensions are mixed together in the Alcohol Abstinence Self-Efficacy Scale so that individuals completing the index may not be aware of them (see Exhibit 8.7). However, the answers to questions representing the particular dimensions tend to be more similar to each other than to answers to questions representing other dimensions—they tend to cluster together. By creating subscales for each of these dimensions, researchers can identify not only the level of confidence in ability to resist drinking cues (abstinence self-efficacy) but also the types of drinking cues that are most difficult for individuals to resist.

An index score is usually calculated as the arithmetic average or sum of responses to the component questions, so that every question that goes into the index counts equally. Exhibit 8.7 shows how an index score is calculated from answers to the questions in the Alcohol Abstinence Self-Efficacy Scale (AASE).

Another approach to creating an index score is by giving different weights to the responses to different questions before summing or averaging the responses. Such a weighted index is also termed a *scale*. The scaling procedure might be as simple as arbitrarily counting responses to one question as worth two or three times as much as responses to another question, but most often, the weight applied to each question is determined through empirical testing. For example, based on Christopher Mooney and Mei Hsien Lee's (1995) research on abortion law reform, the scoring procedure for a scale of support for abortion might give a 1 to agreement that abortion should be allowed "when the pregnancy resulted from rape or incest" and a 4 to agreement with the statement that abortion should be allowed "whenever a woman decided she wanted one." In other words, agreeing that abortion is allowable in any circumstances is much stronger support for abortion rights than is agreeing that abortion should be allowed in the case of rape or incest.

**Exhibit 8.7 Alcohol Abstinence Self-Efficacy Scale (AASE)**

Listed below are a number of situations that lead some people to drink.

We would first like to know:

1. How tempted you may be to drink in each situation. Circle the number in each column that best describes the feelings of temptation in each situation at the present time according to the following scale:

- 1 = Not at all tempted
- 2 = Not very tempted
- 3 = Moderately tempted
- 4 = Very tempted
- 5 = Extremely tempted

Situation	Tempted				
	Not at All	Not Very	Moderately	Very	Extremely
1. When I am in agony because of stopping or withdrawing from alcohol use	1	2	3	4	5
2. When I have a headache	1	2	3	4	5
3. When I am feeling depressed	1	2	3	4	5
4. When I am on vacation and want to relax	1	2	3	4	5
5. When I am concerned about someone	1	2	3	4	5
6. When I am very worried	1	2	3	4	5
7. When I have the urge to try just one drink to see what happens	1	2	3	4	5
8. When I am being offered a drink in a social situation	1	2	3	4	5

Source: *Journal of Studies on Alcohol*, volume 55, pp 141–148, 1994. Center for Alcohol Studies, Rutgers.

## ▣ Designing Questionnaires

**Questionnaire:** The survey instrument containing the questions in a self-administered survey.

**Interview schedule:** The survey instrument containing the questions asked by the interviewer in an in-person or phone survey.

Survey questions are answered as part of a **questionnaire** (or **interview schedule**, as it's often called in interview-based studies), not in isolation from other questions. The context created by the questionnaire has a major impact on how individual questions are interpreted and whether they are even answered. As a result, survey researchers must give very careful attention to the design of the questionnaire as well as to the individual questions that it includes.

The way a questionnaire should be designed varies with the specific survey method used and with other particulars of a survey project. There can be no precise formula for identifying questionnaire features that reduce error. Nonetheless,

some key principles should guide the design of any questionnaire, and some systematic procedures should be considered for refining it. I will use Mirowsky and Ross's (1999) questionnaire for studying the psychological effects of changes in a household structure to illustrate some of these principles and procedures.

## Build on Existing Instruments

If another researcher already has designed a set of questions to measure a key concept, and evidence from previous surveys indicates that this measure is reliable and valid, then, by all means, use that instrument. Resources such as Delbert Miller and Neil J. Salkind's (2002) *Handbook of Research Design and Social Measurement* can give you many ideas about existing instruments; your literature review at the start of a research project should be an even better source. Mirowsky and Ross (2003:208–213) drew many of their measures from an extensive body of prior research (including their own).

But there is a trade-off here. Questions used previously may not concern quite the right concept or may not be appropriate in some ways to your population. For example, *sense of control* was a key concept in the Mirowsky and Ross ASOC survey (Mirowsky 1999:13), so they carefully reviewed prior research that had measured this concept. They found that people who were older and had lower incomes tended to “agree” more with statements to which they were asked to respond. As a result, Mirowsky and Ross decided to use a measure of sense of control that was not subject to agreement bias. A good rule of thumb is to use a previously designed instrument if it measures the concept of concern to you and if you have no clear reason for thinking that the instrument is not appropriate with your survey population. You should ask other researchers for their opinions about difficult concepts to measure before making a final decision.

## Refine and Test Questions

Adhering to the preceding question-writing guidelines will go a long way toward producing a useful questionnaire. However, simply asking what appear to you to be clear questions does not ensure that people have a consistent understanding of what you are asking. You need some external feedback—the more of it the better. This feedback is obtained from some type of pretest (Dillman 2000:140–147). Pretesting of some sort is an essential step in preparing any survey.

One important form of feedback results from simply discussing the questionnaire content with others. Persons who should be consulted include expert researchers, key figures in the locale or organization to be surveyed (e.g., elected representatives, company presidents, and community leaders), and some individuals from the population to be sampled. Run your list of variables and specific questions by such figures whenever you have a chance. Reviewing the relevant literature to find results obtained with similar surveys and comparable questions is also an important step to take, if you haven't already conducted such a review before writing your questions. Forming a panel of experts to review the questions can also help: Stanley Presser and Johnny Blair (1994) recommend a panel of a psychologist, a questionnaire design expert, and a general methodologist (cited in Peterson 2000:116).

Another increasingly popular form of feedback comes from guided discussions between potential respondents, called *focus groups*, to check for consistent understanding of terms and to identify the range of events or experiences about which people will be asked to report. By listening to and observing the focus group discussions, researchers can validate their assumptions about what level of vocabulary is appropriate and what people are going to be reporting (Fowler 1995). (See Chapter 10 for more about this technique.)

Professional survey researchers also use a technique for improving questions called the **cognitive interview** (Dillman 2000:66–67; Fowler 1995). Although the specifics vary, the basic approach is to ask people to describe what they are thinking when they answer questions. The researcher asks a test question, then probes with follow-up questions about how the respondent understood one or more words in

**Cognitive interview:** A technique for evaluating questions in which researchers ask people test questions and then probe with follow-up questions to learn how they understood the question and what their answers mean.



**Researcher Interview Link**  
Writing Survey Questions



the question, how confusing it was, and so forth (Schaeffer & Presser 2003:82). This method can identify many problems with proposed questions, particularly if the individuals interviewed reflect the population to be surveyed. Different particular approaches to cognitive interviewing can identify different problems with survey questions. However, there is as yet no single approach to cognitive interviewing that can be considered most effective (Presser et al. 2004:109–130).

**Survey pretest:** A method of evaluating survey questions and procedures by testing them on a small sample of individuals like those to be included in the actual survey and then reviewing responses to the questions and reactions to the survey procedures.

In a traditional **survey pretest**, interviewers administer the questionnaire to a small set of respondents (perhaps 15–25) who are similar to those who will be sampled in the planned survey. After the interviews are completed, the interviewers discuss the experience with the researcher and, through this discussion, try to identify questions that caused problems. The SRL pretested the Mirowsky and Ross ASOC survey in the summer of 1994, before their first wave of interviews (Mirowsky 1999:34). Try it yourself if you develop a questionnaire. Prepare for the pretest by completing the questionnaire yourself and then revising it. Next, try it out on some colleagues or other friends, and revise it again. For the actual pretest, draw a small sample of individuals from the population you are studying, or one very similar to it, and try out the survey procedures with them, including as many mailings as you plan if you will mail your questionnaire, and actual interviews if you plan to conduct in-person interviews. In the pretest version of a written questionnaire, you may include some space for individuals to comment on each key question or, with in-person interviews, audiotape the test interviews for later review.

Conclude the pretest by reviewing the responses to each question and listening to the audiotapes and reading the comments. Revise any questions that respondents do not seem to interpret as you had intended or that are not working well for other reasons. If the response rate is relatively low, consider whether it can be improved by some modifications in procedures.

**Behavior coding:** Observation in which the researcher categorizes, according to strict rules, the number of times certain behaviors occur.

The value of a pretest can be enhanced with **behavior coding**: A researcher observes the interviews or listens to taped interviews and codes, according to strict rules, the number of times that difficulties occur with questions (Krosnick 1999:541). Such difficulties include respondents asking for clarification and interviewers rephrasing questions rather than reading them verbatim (Presser & Blair 1994:74–75). This information is then used to improve question wording and instructions for interviewers about reading the questions (Schaeffer & Presser 2003:82).

Which method of improving questions is best? Each has unique advantages and disadvantages. Behavior coding, with its clearly specified rules, is the most reliable method across interviewers and repetitions, whereas simple pretesting is the least reliable. However, focus groups or cognitive interviews are better for understanding the bases of problems with particular questions. Review of questions by an expert panel is the least expensive method and identifies the greatest number of problems with questions (Presser & Blair 1994).

**Interpretive questions:** Questions included in a questionnaire or interview schedule to help explain answers to other important questions.

## Add Interpretive Questions

A survey researcher can also include **interpretive questions** in the survey itself to help the researcher understand what the respondent meant by his or her responses to particular questions. An example from a study of people with motor vehicle driving violations illustrates the importance of interpretive questions:

When asked whether their emotional state affected their driving at all, respondents would reply that their emotions had very little effect on their habits. Then, when asked to describe the circumstances surrounding their last traffic violation, respondents typically replied, “I was mad at my girlfriend,” or “I had a quarrel with my wife,” or “We had a family quarrel,” or “I was angry with my boss.” (Labaw 1980:71)





## CAREERS AND RESEARCH

### Floyd J. (“Jack”) Fowler Jr., PhD, Founder and Director of the Center for Survey Research

Jack Fowler “wrote the book” on survey research—two books, actually, with SAGE Publications: *Improving Survey Questions* and *Survey Research Methods*. This career focus crept up on Fowler while he was in school. As an undergraduate major in English at Wellesley College, Fowler found himself fascinated by social science and went on to earn his PhD in social psychology at the University of Michigan. In graduate school, he got hooked on survey research.

Fowler was asked to serve as a research assistant in a series of studies designed to identify the sources of error in the National Health Interview Survey, a major source of health data in the United States. This was an opportunity to relate research methodology to real-life problems and improve the way the world works, and Fowler seized it. He went on to found the Center for Survey Research at the University of Massachusetts Boston and to serve as its director for more than two decades. Fowler describes his professional life as “essentially a series of projects” that have made a difference by helping address important problems in areas ranging from health care, crime, and housing to medical decision making and views of local government.

His advice for students interested in a similar career:

Methods, methods, methods. Make sure you are firmly grounded in the methods of collecting and analyzing data. The research priorities will change, society and the nature of the problems change so fast. However, if you know how to collect and analyze data, you will always be relevant. . . . To enjoy work most days and to feel confident it is making a positive difference is about as good as one can ask for.

Were these respondents lying in response to the first question? Probably not. More likely, they simply didn’t interpret their own behavior in terms of general concepts such as *emotional state*. But their responses to the first question were likely to be misinterpreted without the further detail provided by answers to the second.

Consider five issues when developing interpretive questions—or when you review survey results and need to consider what the answers tell you:

1. *What do the respondents know?* Answers to many questions about current events and government policies are almost uninterpretable without also learning what the respondents know. In studies such as Mirowsky and Ross’s (2003), however, which focused on personal experiences and feelings, questions assessing knowledge are not as necessary.

2. *What relevant experiences do the respondents have?* Such experiences undoubtedly color the responses. For example, the meaning of opinions about crime and punishment may differ greatly between those who have been crime victims themselves and those who have not. Mirowsky and Ross (2001:7) preceded a series of questions concerning feelings about work with a question about their current employment status. If respondents were not working, they were not asked the other questions about work.

3. *How consistent are the respondents’ attitudes, and do they express some larger perspective or ideology?* An employee who seeks more wages because he or she believes that all employer profits result from exploitation is expressing a different sentiment from one who seeks more wages because he or she really wants a more expensive car with which to impress his or her neighbors.

4. *Are respondents' actions consistent with their expressed attitudes?* We probably should interpret differently the meaning of expressed support for gender equality from married men who help with household chores and those who do not. Questions about behavior may also provide a better way to assess orientations than will questions about attitudes. Patricia Labaw (1980:100) points out that “the respondent’s actual purchase of life insurance is a more accurate representation of what he believes about his life insurance needs than anything he might say in response to a direct question” about whether it is important to carry life insurance.

5. *How strongly are the attitudes held?* The attitudes of those with stronger beliefs are more likely to be translated into action than are attitudes that are held less strongly. Just knowing the level of popular support for, say, abortion rights or gun control thus fails to capture the likelihood of people to march or petition their representatives on behalf of the cause; we also need to know what proportion of supporters feel strongly (Schuman & Presser 1981:chap. 9). Thus, rather than just asking unmarried respondents if they wish to remarry, Mirowsky and Ross (2001:1) used the following question and response choices to measure strength of desire to remarry in their telephone survey:

How much would you like to get remarried someday?  
Would you say . . .

1. <1> Not at all,
2. <2> Somewhat, or
3. <3> Very much?

The qualitative insights produced by open-ended questions (see Chapter 4) can be essential for interpreting the meaning of fixed responses. For example, Renee Anspach (1991) asked administrators, case managers, clients, and family members in four community mental health systems whether their programs were effective. They usually rated their programs as effective when given fixed-choice responses. However, their responses to a series of open-ended questions pointed to many program failings. Anspach concluded that the respondents’ positive answers to her initial question reflect their desire to make the program appear effective, for several reasons: Administrators wanted to maintain funding and employee morale, and case managers wanted to encourage cooperation by clients and their families, as well as to deflect blame for problems to clients, families, or system constraints.

### Maintain Consistent Focus

A survey (with the exception of an omnibus survey) should be guided by a clear conception of the research problem under investigation and the population to be sampled. Does the study seek to describe some phenomenon in detail, to explain some behavior, or to explore some type of social relationship? Until the research objective is formulated clearly, survey design cannot begin. Throughout the process of questionnaire design, this objective should be the primary basis for making decisions about what to include and exclude and what to emphasize or treat in a cursory fashion. Moreover, the questionnaire should be viewed as an integrated whole, in which each section and every question serves a clear purpose related to the study’s objective and complements other sections or questions.



Audio Link  
Questionnaires



## In the News

### Research in the News

#### WHAT CAN SURVEYS UNCOVER?

A survey of 2,000 retired New York City police officers uncovered some problems in crime reporting. The survey focused on manipulation of reported crimes, such as downgrading a crime to a less serious offense or discouraging individuals from filing reports. A subsample of recently retired officers (N = 871) showed that more than half had “personal knowledge” of such manipulation. Criminologist Eli Silverman is using this survey to shed light on the systemic culture of improper reporting.

#### For Further Thought



1. What features of a survey do you think would make honest reporting of misbehavior like this more likely? Consider asking about the behaviors of acquaintances, the auspices of the survey, and the method of survey administration.
2. Based on your own experience, do you think surveys of college students can provide valid estimates of student grades? . . . of instructor performance? . . . of the use of alcohol or marijuana?

*News source:* Ruderman, Wendy. 2012. “Crime Report Manipulation Is Common Among New York Police, Study Finds.” *The New York Times*, June 29:A17.

Surveys often include too many irrelevant questions and fail to include questions that, the researchers realize later, are crucial. One way to ensure that possibly relevant questions are asked is to use questions suggested by prior research, theory, experience, or experts (including participants) who are knowledgeable about the setting under investigation. Of course, not even the best researcher can anticipate the relevance of every question. Researchers tend to try to avoid “missing something” by erring on the side of extraneous questions (Labaw 1980:40).

### Order the Questions

The order in which questions are presented will influence how respondents react to the questionnaire as a whole and how they may answer some questions (Schwarz 2010:47). As a first step, the individual questions should be sorted into broad thematic categories, which then become separate sections in the questionnaire. For example, Mirowsky and Ross’s (1999) questionnaire contained the following sections: health, employment experience, sociodemographics, social–psychological attitudes, and social support. Both the sections and the questions within the sections must then be organized in a logical order that would make sense in a conversation. Throughout the design process, the grouping of questions in sections and the ordering of questions within sections should be adjusted to maximize the questionnaire’s overall coherence.

The first question deserves special attention, particularly if the questionnaire is to be self-administered. This question signals to the respondent what the survey is about, whether it will be interesting, and how easy it will be to complete. For these reasons, the first question should connect to the primary

purpose of the survey, be interesting, easy, and will apply to everyone in the sample (Dillman 2000:92–94). Mirowsky and Ross (1999) began their survey about health and related issues with a question about the respondent's overall health:

First, I'd like to ask you about your health. In general, would you say your health is . . .

- |                      |                   |
|----------------------|-------------------|
| 1. <1> Very good,    | 4. <4> Poor, or   |
| 2. <2> Good,         | 5. <5> Very poor? |
| 3. <3> Satisfactory, |                   |

One or more filter or screening questions may also appear early in the survey to identify respondents for whom the questionnaire is not intended or perhaps to determine which sections of a multipart questionnaire a respondent is to skip (Peterson 2000:106–107).

Question order can lead to **context effects** when one or more questions influence how subsequent questions

**Context effects:** Occur in a survey when one or more questions influence how subsequent questions are interpreted.

**Part-whole question effects:** These occur when responses to a general or summary question about a topic are influenced by responses to an earlier, more specific question about that topic.

are interpreted (Schober 1999:88–89). For example, when a sample of the general public was asked, “Do you think it should be possible for a pregnant woman to obtain a legal abortion if she is married and does not want any more children?” Fifty-eight percent said yes. However, when this question was preceded by a less permissive question that asked whether the respondent would allow abortion of a defective fetus, only 40% said yes. Asking the question about a defective fetus altered respondents' frame of reference, perhaps by making abortion simply to avoid having more children seem frivolous by comparison (Turner & Martin 1984:135). Context effects have also been identified in the measurement of general happiness, in what is termed a **part-whole question effect** (Peterson 2000:113). Married people tend to report that they

are happier “in general” if the general happiness question is preceded by a question about their happiness with their marriage (Schuman & Presser 1981:23–77).

Prior questions can influence how questions are comprehended, what beliefs shape responses, and whether comparative judgments are made (Tourangeau 1999). The potential for context effects is greatest when two or more questions concern the same issue or closely related issues, as in the example of the two questions about abortion. The impact of question order also tends to be greater for general, summary-type questions, as with the example about general happiness.

Context effects can be identified empirically if the question order is reversed on a subset of the questionnaires (the so-called split-ballot design) and the results compared. However, knowing that a context effect occurs does not tell us which order is best. Reviewing the overall survey goals and any other surveys with which comparisons should be made can help us decide on question order. What is most important is to be aware of the potential for problems resulting from question order and to evaluate carefully the likelihood of context effects in any particular questionnaire. Those who report survey results should mention, at least in a footnote, the order in which key questions were asked when more than one question about a topic was used (Labaw 1980). An alternative approach is to randomize the order in which key questions are presented, so that any effects of question order cancel each other out.

**Matrix questions:** a series of questions that concern a common theme and that have the same response choices.

Some questions may be presented in a *matrix* format. **Matrix questions** are a series of questions that concern a common theme and that have the same response choices. The questions are written so that a common initial phrase applies to each

**Exhibit 8.8** A Page From Ross's Interview Schedule

45. In the past 12 months about how many times have you gone on a diet to lose weight?				v94
Never .....	0			
Once .....	1			
Twice .....	2			
Three times or more .....	3			
Always on a diet .....	4			
46. What is your height without shoes on?				v95
_____ ft.                      _____ in.				
47. What is your weight without clothing?				v96
_____ lbs.				
48a. Do you currently smoke 7 or more cigarettes a week?				v97
Yes .....	1 --> (SKIP TO Q.49)			
No .....	2			
48b. Have you ever smoked 7 or more cigarettes a week?				v98
Yes .....	1			
No .....	2			
49. How much difficulty do you have ...				
	<b>No</b>	<b>Some</b>	<b>A great</b>	
	<b>difficulty,</b>	<b>difficulty, or</b>	<b>deal of</b>	
	<b>difficulty?</b>	<b>difficulty?</b>	<b>difficulty?</b>	
a. Going up and down stairs? Would you say .....	1	2	3	v99
b. Kneeling or stooping? .....	1	2	3	v100
c. Lifting or carrying objects less than 10 pounds, like a bag of groceries? .....	1	2	3	v101
d. Using your hands or fingers? .....	1	2	3	v102
e. Seeing, even with glasses? .....	1	2	3	v103
f. Hearing? .....	1	2	3	v104
g. Walking? .....	1	2	3	v105

Source: From Catherine E. Ross, *Work, Family, and the Sense of Control* (1990). Reprinted with permission of the author.

one (see Question 49 in Exhibit 8.8). This format shortens the questionnaire by reducing the number of words that must be used for each question. It also emphasizes the common theme among the questions and so invites answering each question in relation to other questions in the matrix. It is very important to provide an explicit instruction to "Check one response on each line" in a matrix question, because some respondents will think that they have completed the entire matrix after they have responded to just a few of the specific questions.

## Make the Questionnaire Attractive

An attractive questionnaire is more likely to be completed and less likely to confuse either the respondent or, in an interview, the interviewer. An attractive questionnaire also should increase the likelihood that different respondents interpret the same questions in the same way.

Printing a multipage questionnaire in booklet form usually results in the most attractive and simple-to-use questionnaire. Printing on both sides of folded-over legal-size paper (8½" by 14") is a good approach, although pages can be printed on one side only and stapled in the corner if finances are very tight (Dillman 2000:80–86). An attractive questionnaire does not look cramped; plenty of white space—more between questions than within question components—makes the questionnaire appear easy to complete. Response choices are distinguished clearly and consistently, perhaps by formatting them with light print (while questions are formatted with dark print) and keeping them in the middle of the pages. Response choices are listed vertically rather than horizontally across the page.

The proper path through the questionnaire for each respondent is identified with arrows or other graphics and judicious use of spacing and other aspects of layout. Respondents should not be confused about where to go next after they are told to skip a question. Instructions should help route respondents through skip patterns, and such skip patterns should be used infrequently. Instructions should also explain how each type of question is to be answered (e.g., by circling a number or writing a response) in a neutral way that isn't likely to influence responses. Some distinctive formatting should be used to identify instructions.

The visual design of a questionnaire has more subtle effects on how respondents answer questions. Seemingly minor differences, such as whether responses are grouped under headings or just listed, whether separate response choices are provided or just the instruction to write in a response from a list of choices, and how much space there is between response choices can all affect the distribution of responses to a question (Dillman & Christian 2005:43–48).

Exhibit 8.8 contains portions of the questionnaire Ross (1990) used in a previous phone survey about aging and health. This page illustrates three of the features that I have just reviewed: (1) numeric designation of response choices, (2) clear instructions, and (3) an attractive, open layout. Because this questionnaire was read over the phone, rather than being self-administered, there was no need for more explicit instructions about the matrix question (Question 49) or for a more distinctive format for the response choices (Questions 45 and 48). A questionnaire designed to be self-administered also should include these additional features.

## Consider Translation

Should the survey be translated into one or more languages? In the 21st century, no survey plan in the United States or many other countries can be considered complete until this issue has been considered. In the United States in 2006, 15.3% of persons aged 18 years and older were foreign born (Pew Hispanic Center 2008:Table 1) and more than half of these adults said that they did not speak English very well (Pew Hispanic Center 2008:Table 20). Depending on the specific region or group that is surveyed, these proportions can be much higher and can include persons fluent in various languages (with Spanish being the most common). Although English becomes the primary language spoken by almost all children of immigrants, many first-generation immigrants are not fluent in English (Hakimzadeh & Cohn 2007:i; Pew Hispanic Center 2008:Table 21). As a result, they can only be included in a survey if it is translated into their native language.

When immigrants are a sizable portion of a population, omitting them from a survey can result in a misleading description of the population. Foreign-born persons in the United States tend to be younger than native-born persons and their average income is lower (Pew Hispanic Center 2008:Tables 8a, 29). They also are more likely to be married, to be in a household with five or more family members, and to have less than a high school education (Pew Hispanic Center 2008:Tables 13, 18, 22). However, none of these differences are true for all immigrant groups. In particular, persons from South and East Asia and the Middle East tend to have more education and higher incomes than do persons born in the United States (Pew Hispanic Center 2008:Tables 22, 29).



So, survey researchers find increasingly that they must translate their questionnaires into one or more languages to represent the population of interest. This does not simply mean picking up a bilingual dictionary, clicking “translate” in a web browser, or hiring a translator to translate the questions and response choices word for word. Such a literal translation may not result in statements that are interpreted in the same way to non-English speakers. The U.S. Census Bureau’s (2006) guidelines for translation designate the literal translation as only one step in the process. What is needed is to achieve some equivalence of the concepts in different cultures (Church 2010:154–159). The U.S. Census Bureau and the World Health Organization (n.d.) recommend that questionnaires be translated by a team that includes trained translators, persons who are specialists in the subject matter of the survey, persons with expertise in questionnaire design, and experts with several of these skills who can review the translation and supervise a pretest (Pan & de la Puente 2005).

A properly translated questionnaire will be

- *Reliable*: conveys the intended meaning of the original text
- *Fluent*: reads well and makes sense in the target language
- *Appropriate*: the style, tone, and function are appropriately transferred

Needless to say, this translation process adds cost and complexity to survey design.

## 📌 Organizing Surveys

There are five basic social science survey designs: (1) mailed, (2) group-administered, (3) phone, (4) in-person, and (5) electronic. Survey researchers can also combine elements of two or more of these basic designs in mixed-mode surveys. Exhibit 8.9 summarizes the typical features of the five basic survey designs.

**Manner of administration.** The five survey designs differ in the manner in which the questionnaire is administered (see Exhibit 8.9). Mailed, group, and **electronic surveys** are completed by the respondents themselves. During phone and in-person interviews, the researcher or a staff person asks the questions and records the respondent’s answers. However, new *mixed-mode* surveys break down these distinctions. For example, in *audio computer-assisted self-interviewing* (or *audio-CASI*), the interviewer gives the respondent a laptop and a headset (Tourangeau 2004:790–791). The respondent reads the questions on the computer screen, hears the questions in the headset, and responds by choosing answers on the computer screen.

**Electronic survey:** A survey that is sent and answered by computer, either through e-mail or on the web.

**Exhibit 8.9** Typical Features of the Five Survey Designs

Design	Manner of Administration	Setting	Questionnaire Structure	Cost
Mailed survey	Self	Individual	Mostly structured	Low to moderate
Group survey	Self	Group	Mostly structured	Very low
Phone survey	Professional	Individual	Structured	Moderate
In-person interview	Professional	Individual	Structured or unstructured	High
Electronic survey	Self	Individual	Mostly structured	Low

**Questionnaire structure.** Survey designs also differ in the extent to which the researcher structures the content and order of questions in advance. Most mailed, group, phone, and electronic surveys are highly structured, fixing in advance the content and order of questions and response choices. Some of these types of surveys, particularly mailed surveys, may include some open-ended questions (respondents write in their answers rather than checking off one of several response choices). In-person interviews are often highly structured, but they may include many questions without fixed response choices. Moreover, some interviews may proceed from an interview guide rather than a fixed set of questions. In these relatively unstructured interviews, the interviewer covers the same topics with respondents but varies questions according to the respondent's answers to previous questions. Extra questions are added as needed to clarify or explore answers to the most important questions. Computers make it easy for researchers to use complex branching patterns in questionnaires administered in person, on the phone, or on the web because the computer can present different questions based on responses to prior questions (Tourangeau 2004:789).



**Journal Link**  
Administering Surveys

**Setting.** Most surveys are conducted in settings where only one respondent completes the survey at a time; most mail and electronic questionnaires and phone interviews are intended for completion by only one respondent. The same is usually true of in-person interviews, although sometimes researchers interview several family members at once. A variant of the standard survey is a questionnaire distributed simultaneously to a group of respondents, who complete the survey while the researcher (or assistant) waits. Students in classrooms are typically the group involved, although this type of group distribution also occurs in surveys of employees and members of voluntary groups.

**Cost.** As mentioned earlier, in-person interviews are the most expensive type of survey. Phone interviews are much less expensive, although costs are rising because of the need to make more calls to reach potential respondents. Surveying by mail is cheaper yet. Electronic surveys can be the least expensive method because there are no interviewer costs, no mailing costs, and, for many designs, almost no costs for data entry. However, extra staff time and programming expertise are required to prepare an electronic questionnaire (Tourangeau, Conrad, & Couper 2012).

Because of their different features, the five designs vary in the types of error to which they are most prone and the situations in which they are most appropriate. The different designs can also be improved in different ways by adding some features of the other designs. This section focuses on the various designs' unique advantages and disadvantages and identifies techniques for reducing error within each design and by combining designs.

**Mailed survey:** A survey involving a mailed questionnaire to be completed by the respondent.

## Mailed, Self-Administered Surveys

A **mailed survey** is conducted by mailing a questionnaire to respondents, who then administer the survey themselves. The central concern in a mailed survey is maximizing the response rate. Even an attractive questionnaire full of clear questions will probably be returned by no more than 30% of a sample unless extra steps are taken to increase the rate of response. It's just too much bother for most potential recipients; in the language of social exchange theory, the costs of responding are perceived to be much higher than any anticipated rewards for doing so. Of course, a response rate of 30% is a disaster; even a response rate of 60% represents so much nonresponse error that it is hard to justify using the resulting data. Fortunately, the conscientious use of a systematic survey design method can be expected to lead to an acceptable 70% or higher rate of response to most mailed surveys (Dillman 2000).

Sending follow-up mailings to nonrespondents is the single most important requirement for obtaining an adequate response rate to a mailed survey. The follow-up mailings explicitly encourage initial

nonrespondents to return a completed questionnaire; implicitly, they convey the importance of the effort. Dillman (2000:155–158, 177–188) has demonstrated the effectiveness of a standard procedure for the mailing process:

- A few days before the questionnaire is to be mailed, send a brief letter to respondents that notifies them of the importance of the survey they are to receive.
- Send the questionnaire with a well-designed, personalized cover letter (see the following description), a self-addressed, stamped return envelope, and, if possible, a token monetary reward. The materials should be inserted in the mail-out envelope so that they will all be pulled out together when the envelope is opened (Dillman 2000:174–175). There should be no chance that the respondent will miss something.
- Send a reminder postcard, thanking respondents and reminding nonrespondents, to all sample members 2 weeks after the initial mailing. The postcard should be friendly in tone and must include a phone number for those people who may not have received the questionnaire. It is important that this postcard be sent before most nonrespondents will have discarded their questionnaire, even though this means the postcard will arrive before all those who might have responded to the first mailing have done so.
- Send a replacement questionnaire with a new cover letter only to nonrespondents, 2 to 4 weeks after the initial questionnaire mailing. This cover letter should be a bit shorter and more insistent than the original cover letter. It should note that the recipient has not yet responded, and it should stress the survey's importance. Of course, a self-addressed, stamped return envelope must be included.
- The final step is taken 6 to 8 weeks after the initial survey mailing. This step uses a different mode of delivery (either priority or special delivery) or a different survey design—usually an attempt to administer the questionnaire over the phone. These special procedures emphasize the importance of the survey and encourage people to respond.

The **cover letter** for a mailed questionnaire is critical to the success of a mailed survey. This statement to respondents sets the tone for the questionnaire. A carefully prepared cover letter should increase the response rate and result in more honest and complete answers to the survey questions; a poorly prepared cover letter can have the reverse effects.

The cover letter or introductory statement must be

- *Credible*: The letter should establish that the research is being conducted by a researcher or organization that the respondent is likely to accept as a credible, unbiased authority. According to one investigation, a sponsor known to respondents may increase the rate of response by as much as 17%. Government sponsors tend to elicit high rates of response. Research conducted by well-known universities and recognized research organizations (e.g., Gallup or RAND) is also usually credible in this sense. The next most credible sponsors are state headquarters of an organization and then other people in a similar field. Publishing firms, students (sorry!), and private associations elicit the lowest response rates.
- *Personalized*: The cover letter should include a personalized salutation (using the respondent's name, e.g., not just "Dear Student"), close with the researcher's signature (blue ballpoint pen is best because that makes it clear that the researcher has personally signed), and refer to the respondent in the second person ("Your participation . . .").

**Cover letter:** The letter sent with a mailed questionnaire that explains the survey's purpose and auspices and encourages the respondent to participate.

- *Interesting:* The statement should interest the respondent in the contents of the questionnaire. Never make the mistake of assuming that what is of interest to you will also interest your respondents. Try to put yourself in their shoes before composing the statement, and then test your appeal with a variety of potential respondents.
- *Responsible:* Reassure the respondent that the information you obtain will be treated confidentially, and include a phone number to call if the respondent has any questions or would like a summary of the final report. Point out that the respondent's participation is completely voluntary (Dillman 1978:165–172).

Exhibit 8.10 is an example of a cover letter for a questionnaire.

Other steps are necessary to maximize the response rate (Fowler 1988:99–106; Mangione 1995:79–82; Miller 1991:144):

- It is particularly important, in self-administered surveys, that the individual questions are clear and understandable to all the respondents because no interviewers will be on hand to clarify the meaning of the questions or to probe for additional details.

### Exhibit 8.10 Sample Questionnaire Cover Letter

#### University of Massachusetts Boston Department of Sociology

Jane Doe  
AIDS Coordinator  
Shattuck Shelter

Dear Jane:

AIDS is an increasing concern for homeless people and for homeless shelters. The enclosed survey is about the AIDS problem and related issues confronting shelters. It is sponsored by the Life Lines AIDS Prevention Project for the Homeless—a program of the Massachusetts Department of Public Health.

As an AIDS coordinator/shelter director, you have learned about homeless persons' problems and about implementing programs in response to those problems. The Life Lines Project needs to learn from your experience. Your answers to the questions in the enclosed survey will improve substantially the base of information for improving AIDS prevention programs.

Questions in the survey focus on AIDS prevention activities and on related aspects of shelter operations. It should take about 30 minutes to answer all the questions.

Every shelter AIDS coordinator (or shelter director) in Massachusetts is being asked to complete the survey. And every response is vital to the success of the survey: The survey report must represent the full range of experiences.

You may be assured of complete confidentiality. No one outside of the university will have access to the questionnaire you return. (The ID number on the survey will permit us to check with nonrespondents to see if they need a replacement survey or other information.) All information presented in the report to Life Lines will be in aggregate form, with the exception of a list of the number, gender, and family status of each shelter's guests.

Please mail the survey back to us by Monday, June 4, and feel free to call if you have any questions.

Thank you for your assistance.

Yours sincerely,

*Russell K. Schutt*

Russell K. Schutt, PhD  
Project Director

*Stephanie Howard*

Stephanie Howard  
Project Assistant

- Use no more than a few open-ended questions because respondents are likely to be put off by the idea of having to write out answers.
- Write an identifying number on the questionnaire so you can determine whom the nonrespondents are. This is essential for follow-up efforts. Of course, the identification must be explained in the cover letter.
- Enclose a token incentive with the survey. A \$2 or \$5 bill seems to be the best incentive. It is both a reward for the respondent and an indication of your trust that the respondent will carry out his or her end of the “bargain.” The response rate to mailed surveys increases by 19 percentage points, on average, in response to such an incentive (Church 1993). Offering a large monetary reward or some type of lottery ticket only for those who return their questionnaire is actually less effective, apparently because it does not indicate trust in the respondent (Dillman 2000:167–170).
- Include a stamped, self-addressed return envelope with each copy of the questionnaire. This reduces the cost for responding. The stamp helps personalize the exchange and is another indication of trust in the respondent (who could use the stamp for something else). Using a stamp rather than metered postage on the mail-out envelope does not seem to influence the response rate, but it is very important to use first class rather than bulk rate postage (Dillman 2000:171–174).
- Consider presurvey publicity efforts. A vigorous advertising campaign increased considerably the response to the 2000 Census mailed questionnaire; the results were particularly successful among minority groups, who had been targeted because of low response rates in the 1990 Census (Holmes 2000).

If Dillman’s procedures are followed, and the guidelines for cover letters and questionnaire design also are adhered to, the response rate is almost certain to approach 70%. One review of studies using Dillman’s method to survey the general population indicates that the average response to a first mailing will be about 24%; the response rate will rise to 42% after the postcard follow-up, to 50% after the first replacement questionnaire, and to 72% after a second replacement questionnaire is sent by certified mail (Dillman et al. 1974).

The response rate may be higher with particular populations surveyed on topics of interest to them, and it may be lower with surveys of populations that do not have much interest in the topic. When a survey has many nonrespondents, getting some ideas about their characteristics, by comparing late respondents with early respondents, can help determine the likelihood of bias resulting from the low rate of response. If those who returned their questionnaires at an early stage are more educated or more interested in the topic of the questionnaire, the sample may be biased; if the respondents are not more educated or more interested than nonrespondents, the sample will be more credible.

If resources did not permit phone calls to all nonrespondents, a random sample of nonrespondents can be selected and contacted by phone or interviewed in person. It should be possible to secure responses from a substantial majority of these nonrespondents in this way. With appropriate weighting, these new respondents can then be added to the sample of respondents to the initial mailed questionnaire, resulting in a more representative total sample (for more details, see Levy & Lemeshow 1999:398–402).

Related to the threat of nonresponse in mailed surveys is the hazard of incomplete response. Some respondents may skip some questions or just stop answering questions at some point in the questionnaire. Fortunately, this problem does not occur often with well-designed questionnaires. Potential respondents who have decided to participate in the survey usually complete it. But there are many exceptions to this observation because questions that are poorly written, too complex, or about sensitive personal issues simply turn off some respondents. The revision or elimination of such questions during the design phase should minimize the problem. When it does not, it may make sense to impute values for the missing data (in effect, estimate the values of missing data). One imputation procedure would be to substitute the mean (arithmetic average) value of a variable for those cases that have a missing value on the variable (Levy & Lemeshow 1999:404–416).



**Group-administered survey:** A survey that is completed by individual respondents who are assembled in a group.

## Group-Administered Surveys

A **group-administered survey** is completed by individual respondents assembled in a group. The response rate is not usually a major concern in surveys that are distributed and collected in a group setting because most group members will participate. The real difficulty with this method is that it is seldom feasible because it requires what might be called a captive audience. With the exception of students, employees, members of the armed forces, and some institutionalized populations, most populations cannot be sampled in such a setting.

Whoever is responsible for administering the survey to the group must be careful to minimize comments that might bias answers or that could vary between different groups in the same survey (Dillman 2000:253–256). A standard introductory statement should be read to the group that expresses appreciation for their participation, describes the steps of the survey, and emphasizes (in classroom surveys) that the survey is not the same as a test. A cover letter like the one used in mailed surveys also should be distributed with the questionnaires. To emphasize confidentiality, respondents should be given an envelope in which to seal their questionnaires after they are completed.

Another issue of special concern with group-administered surveys is the possibility that respondents will feel coerced to participate and, as a result, will be less likely to answer questions honestly. Also, because administering a survey in this way requires approval of the powers that be—and this sponsorship is made quite obvious by the fact that the survey is conducted on the organization's premises—respondents may infer that the researcher is not at all independent of the sponsor. No complete solution to this problem exists, but it helps to make an introductory statement emphasizing the researcher's independence and giving participants a chance to ask questions about the survey. The sponsor should also understand the need to keep a low profile and to allow the researcher both control over the data and autonomy in report writing. Participation in group-administered surveys of grade school and high school students can be reduced because of the requirement of parental permission, but here the group context can be used to the researcher's advantage. Jane Onoye, Deborah Goebert, and Stephanie Nishimura (2012) at the University of Hawai'i at Manoa found that offering a class a reward such as a pizza if a high rate of participation was achieved led to more parental consent forms being returned than when students were offered a \$5 gift card for participating.

**Phone survey:** A survey in which interviewers question respondents over the phone and then record their answers.

## Telephone Surveys

In a **phone survey**, interviewers question respondents over the phone and then record respondents' answers. Phone interviewing became a very popular method of conducting surveys in the United States because almost all families had phones by the latter part of the 20th century. But two matters may undermine the validity of a phone survey: not reaching the proper sampling units and not getting enough complete responses to make the results generalizable.

### Reaching Sample Units

There are three different ways of obtaining a sampling frame of telephone exchanges or numbers: (1) Phone directories provide a useful frame for local studies; (2) a nationwide list of area code or exchange numbers can be obtained from a commercial firm (random digit dialing is used to fill in the last four digits); and (3) commercial firms can provide files based on local directories from around the nation. There are coverage errors with each of these frames: 10% to 15% of directory listings will turn out not to still be valid residential numbers; more than 35% of U.S. households with phones have numbers that are unlisted in directories, and the percentage is as high as 60% in some communities; and less than 25% of the area codes and exchanges in the one national comprehensive list (available from Bell Core Research, Inc.) refer to residential units (Levy & Lemeshow 1999:455–460). In planning a survey, researchers must consider the advantages and disadvantages of these methods for a particular study and develop means to compensate for the weaknesses of the specific method chosen.

Research Social Impact  
Link  
Sample Units



Most telephone surveys use random digit dialing at some point in the sampling process (Lavrakas 1987). A machine calls random phone numbers within the designated exchanges, whether or not the numbers are published. When the machine reaches an inappropriate household (such as a business in a survey that is directed to the general population), the phone number is simply replaced with another. Most survey research organizations use special methods (some version of the Mitofsky–Waksberg method) to identify sets of phone numbers that are likely to include working numbers and so make the random digit dialing more efficient (Tourangeau 2004:778–780).

The University of Illinois SRL used this approach to draw the original sample for Mirowsky and Ross's study of education, social status, and health (Mirowsky 1999). Because the research had a particular focus on health problems related to aging, the researchers used a stratified sampling procedure and oversampled older Americans:

The survey of Aging, Status, and the Sense of Control (ASOC) is a national telephone probability sample of United States households. A first wave of interviews was completed at the beginning of 1995. Respondents were selected using a prescreened random-digit dialing method that increases the hit rate and decreases standard errors compared with the standard Mitofsky–Waksberg method while producing a sample with the same demographic profile (Lund & Wright 1994; Waksberg 1978). The ASOC survey has two subsamples, designed to produce an 80 percent oversample of persons aged 60 or older. The general sample draws from all households; the oversample draws only from households with one or more seniors. In the general sample the adult (18 or older) with the most recent birthday was selected as respondent. In the oversample the senior (60 or older) with the most recent birthday was selected. For practical reasons the survey was limited to English-speaking adults. Up to 10 callbacks were made to select and contact a respondent, and up to 10 to complete the interview once contact was made. (p. 34)

For the third wave of interviews in 2000–2001, SRL planned an intensive effort to contact the original members of the Wave I sample (Mirowsky 1999):

Attempts will be made to contact and interview all wave 1 respondents, whether or not they were in wave 2, except for individuals who made it clear they did not want to be contacted in the future. A number of new strategies for maximizing follow-up will be tried (Smith 1995; Lyberg & Dean 1992):

(1) Using tested optimal time-of-day/day-of-week callback sequences, lengthening the period of time over which calls are made, and trying a final sequence of five calls three months after an initial sequence of calls fails to make contact; (2) Giving interviewers additional training on establishing rapport and interacting flexibly; (3) Sending advance letters on letterhead to all baseline respondents that include the survey laboratory phone number that will appear on caller ID, an 800 number to call for additional information about the study, several lines of tailored motivational text, and the location of a web page with information about the study, including the e-mail address and phone number of the project coordinator; (4) Sending a letter after first refusal, signed by the investigator, explaining the study and the importance of participation, and giving an 800 number to call if they decide to participate; (5) Attempting to find respondents not at the original phone number by using directory assistance, the Equifax database, and six web database programs and search engines; (6) Interviewing persons other than the respondent who answer the phone or persons previously identified by the respondent as likely to know their whereabouts, to locate the respondent or identify a likely reason for noncontact (e.g., passed away, moved to a nursing home, too sick to participate, retired and moved away). (pp. 34–35)

However households are contacted, the interviewers must ask a series of questions at the start of the survey to ensure that they are speaking to the appropriate member of the household. Exhibit 8.11 displays a portion of the instructions that the SRL used to select the appropriate respondent for Mirowsky and Ross's phone survey about education and health. This example shows how appropriate and inappropriate households can be distinguished in a phone survey, so that the interviewer is guided to the correct respondent.



**Journal Link**  
Telephone Surveys

**Exhibit 8.11 Phone Interview Procedure for Respondent Designation****868 Aging, Status, and the Sense of Control  
Informant Questionnaire****Introduction and Selection of Respondent or Informant**

[Not shown to interviewer, this is a check item]

[if contact attempts less than 15, the interviewer will go to >h101<.

Interviewers will only make a total of 20 attempts]

[if contact attempts greater than 15, the interviewer will go to >h102<]

YOU ARE CALLING [RNAME].

May I speak with [RNAME]?

- <1> YES, CONNECTED TO RESPONDENT
- <2> NOT AVAILABLE
- <3> NEVER ABLE TO INTERVIEW—TOO HARD OF HEARING, PERMANENTLY ILL, OR FOR SOME OTHER REASON
- <4> NO ONE THERE BY THAT NAME, OR NO LONGER LIVE THERE
- <5> LANGUAGE PROBLEM
- <6> DECEASED
- <7> OTHER
- <8> CHILD, NO ADULTS AVAILABLE
- <9> REFUSED

[if <1> go to expl] (Respondents go to main study questionnaire)

[if <2-9> go to wh02]

IF RESPONDENT IS NOT AVAILABLE AFTER SEVERAL ATTEMPTS OR YOU ARE UNABLE TO COMPLETE THE INTERVIEW WITH RESPONDENT, ATTEMPT TO UPDATE RESPONDENT'S ADDRESS AND TELEPHONE INFORMATION WITH AN INFORMANT.

(Introduction for a respondent)

My name is [Interviewer name] and I am calling from the University of Illinois. Approximately 2 years ago you participated in a telephone interview regarding health and different experiences pertaining to sense of control. We are calling to complete a follow-up survey that will take about 30 minutes.

- <1> YES, RESPONDENT IS AVAILABLE
- <3> RESPONDENT PREFERS CALLBACK—SET UP APPOINTMENT
- <4> NEVER ABLE TO INTERVIEW—TOO HARD OF HEARING, PERMANENTLY ILL, OR FOR SOME OTHER REASON
- <5> DUPLICATE
- <6> LANGUAGE PROBLEM
- <7> OTHER
- <9> REFUSED

(Introduction for an informant)

My name is [Interviewer name] and I am calling from the University of Illinois. Approximately 2 years ago [RNAME] participated in a telephone interview regarding health and different experiences pertaining to sense of control. We are calling to complete a follow-up survey with [RNAME]. Since [RNAME] is not available we would like to update (his or her) telephone and address information.

- <1> INFORMANT IS AVAILABLE
- <3> INFORMANT PREFERS CALLBACK—SET UP APPOINTMENT

- <4> NEVER ABLE TO INTERVIEW—TOO HARD OF HEARING, PERMANENTLY ILL, OR FOR SOME OTHER REASON
- <5> DUPLICATE
- <6> LANGUAGE PROBLEM
- <7> OTHER
- <9> REFUSED

...

YOU WILL VERIFY THE SPELLING OF THE RESPONDENT'S NAME AND ADDRESS.

Because this study is about how people may change during their lives, we may want to call [RNAME] again in a few years. I'd like to verify the information we have about [RNAME]. First, I would like to ask you about the spelling of [RNAME]'s name.

- <1> PROCEED [go to U1b]
  - <9> REFUSED [go to U4]
- Is [RNAME]'s first name spelled [R First Name]?
- <1> Yes
  - <2> No [go to U2]
  - <7> NO CODED RESPONSE APPLICABLE
  - <8> DON'T KNOW
  - <9> REFUSED

...

Can [fill RNAME] still be reached at this phone number?

- <1> Yes
  - <2> No
- >U22<

What is the (correct) phone number to reach [fill RNAME]?

RECORD WHO YOU COMPLETED THE INTERVIEW WITH.

DID YOU SPEAK TO THE RESPONDENT OR INFORMANT? (DO NOT ASK.)

- <1> RESPONDENT
- <2> INFORMANT
- <8> DON'T KNOW

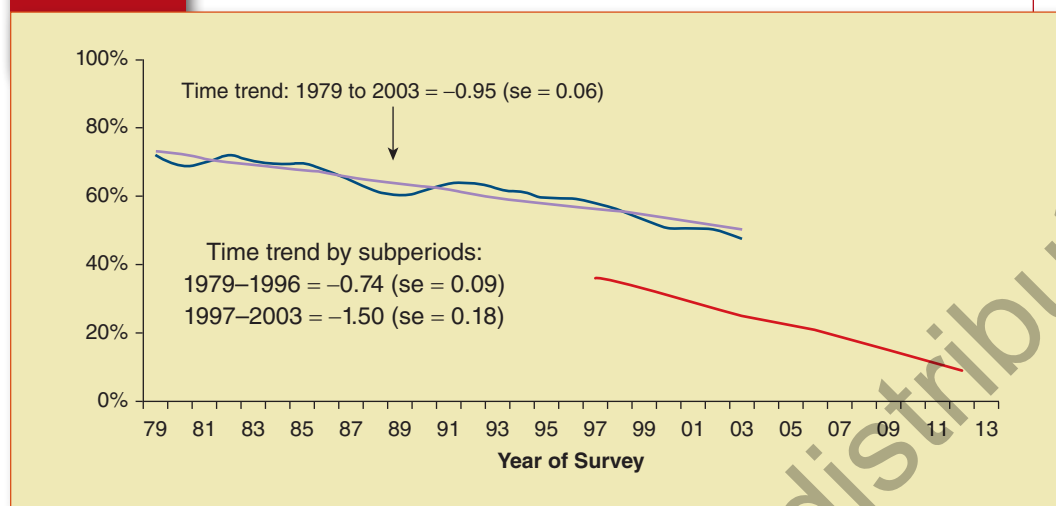
Source: Ross (1990:7).

### Maximizing Response to Phone Surveys

Four issues require special attention in phone surveys. First, because people often are not home, multiple callbacks will be needed for many sample members. Those with more money and education are more likely to be away from home; such persons are more likely to vote Republican, so the results of political polls can be seriously biased if few callback attempts are made (Kohut 1988).

This problem has been compounded in recent years by social changes that are lowering the response rate in phone surveys (Tourangeau 2004:781–783) (see Exhibit 8.12). The Pew Research Center reports a decline in the response rate based on all those sampled from 36% in 1997 to only 9% in 2012 (Kohut et al. 2012).

The number of callbacks needed to reach respondents by telephone has increased greatly in the past 20 years, with increasing numbers of single-person households, dual-earner families, and out-of-home activities. Survey research organizations have increased the usual number of phone contact attempts from between 4 to 8 to 20. The growth of telemarketing has created another problem for telephone survey researchers: Individuals have become

**Exhibit 8.12** Phone Survey Response Rates by Year, 1979–2012

*Source:* Adapted from Curtin, Richard, Stanley Presser, and Eleanor Singer. “Changes in Telephone Survey Nonresponse Over the Past Quarter Century.” *Public Opinion Quarterly*, 69:87–98. Copyright © 2005, Oxford University Press, on behalf of the American Association for Public Opinion Research. Reprinted with permission.

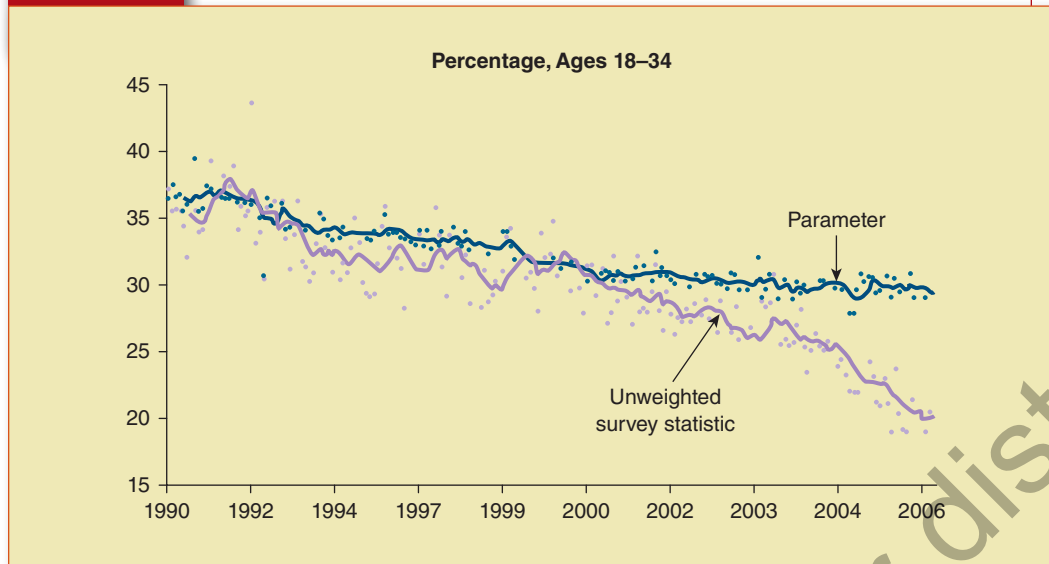
more accustomed to “just say no” to calls from unknown individuals and organizations or to simply use their answering machines to screen out unwanted calls (Dillman 2000:8, 28). Cell phone users are also harder (and more costly) to contact in phone surveys because their numbers are not in published directories. Households with a cell phone but no landline tend to be younger, so the rate of phone survey participation is declining among those 18 to 34 years of age (Keeter 2008) (Exhibit 8.13).

The second issue researchers using phone surveys must cope with are difficulties because of the impersonal nature of phone contact. Visual aids cannot be used, so the interviewer must be able to convey verbally all information about response choices and skip patterns. Instructions to the interviewer must clarify how to ask each question, and response choices must be short. The SRL developed the instructions shown in Exhibit 8.14 to clarify procedures for asking and coding a series of questions that Ross (1990) used in another survey to measure symptoms of stress within households.

Third, interviewers must be prepared for distractions because the respondent likely will be interrupted by other household members. Sprinkling interesting questions throughout the questionnaire may help maintain respondent interest. In general, rapport between the interviewer and the respondent is likely to be lower with phone surveys than with in-person interviews, and so respondents may tire and refuse to answer all the questions (Miller 1991:166). Distractions are a special problem when respondents are called on a cell phone because they could be driving, in a restaurant or other crowded area, at work, or otherwise involved in activities that make responding difficult and that would not occur in a survey using a landline in the home (AAPOR 2014).

The fourth special consideration for phone surveys is that careful interviewer training is essential. This is how one survey research organization describes its training:

In preparation for data collection, survey interviewers are required to attend a two-part training session. The first part covers general interviewing procedures and techniques as related to the proposed survey. The second entails in-depth training and practice for the survey. This training includes instructions on relevant subject matter, a question-by-question review of the survey instrument and various forms of role-playing and practice interviewing with supervisors and other interviewers. (J. E. Blair, personal communication to C. E. Ross, April 10, 1989)

**Exhibit 8.13** Trend in Percentage, Ages 18 to 34, Responding to Phone Surveys

Source: Pew Research Studies, Keeter (2008).

Procedures can be standardized more effectively, quality control maintained, and processing speed maximized when phone interviewers use **computer-assisted telephone interviews (CATI)**:

The interviewing will be conducted using “CATI” (Computer-Assisted Telephone Interviewing). . . . The questionnaire is “programmed” into the computer, along with relevant skip patterns throughout the instrument. Only legal entries are allowed. The system incorporates the tasks of interviewing, data entry, and some data cleaning. (J. E. Blair, personal communication to C. E. Ross, April 10, 1989)

Computerized **interactive voice response (IVR)** survey technology allows even greater control over interviewer–respondent interaction. In an IVR survey, respondents receive automated calls and answer questions by pressing numbers on their touch-tone phones or speaking numbers that are interpreted by computerized voice recognition software. These surveys can also record verbal responses to open-ended questions for later transcription. Although they present some difficulties when many answer choices must be used or skip patterns must be followed, IVR surveys have been used successfully with short questionnaires and when respondents are highly motivated to participate (Dillman 2000:402–411). When these conditions are not met, potential respondents may be put off by the impersonality of this computer-driven approach.

Phone surveying had for decades been the method of choice for relatively short surveys of the general population. Response rates in phone surveys traditionally tended to be very high—often above 80%—because few individuals would hang up on a polite caller or suddenly stop answering questions (at least within the first 30 minutes or so). Mirowsky and Ross (2003:207) achieved a response rate of 71.6% for people who could be contacted in their Wave I survey in 1995. However, phone surveying is not a panacea and it should no longer be considered the best method to use for general purpose surveys. You have already learned of the dramatic decline in phone survey response rates, although this can be somewhat mitigated by extra effort. In a recent

**Computer-assisted telephone interview (CATI):**

A telephone interview in which a questionnaire is programmed into a computer, along with relevant skip patterns, and only valid entries are allowed; incorporates the tasks of interviewing, data entry, and some data cleaning.

**Interactive voice response (IVR):**

A survey in which respondents receive automated calls and answer questions by pressing numbers on their touch-tone phones or speaking numbers that are interpreted by computerized voice recognition software.

**Exhibit 8.14** Sample Interviewer Instructions*Question:*

41. On how many of the past 7 days have you . . .  
 Number of days
- a. Worried a lot about little things? \_\_\_\_\_
- b. Felt tense or anxious? \_\_\_\_\_

*Instructions for interviewers:*

Q41 For the series of “On how many of the past 7 days,” make sure the respondent gives the numerical answer. If he/she responds with a vague answer like “not too often” or “just a few times,” ask again “On how many of the past 7 days would you say?” Do NOT lead the respondent with a number (e.g., “Would that be 2 or 3?”). If R says “all of them,” verify that the answer is “7.”

*Question:*

45. In the past 12 months about how many times have you gone on a diet to lose weight?
- |                     |   |
|---------------------|---|
| Never               | 0 |
| Once                | 1 |
| Twice               | 2 |
| Three times or more | 3 |
| Always on a diet    | 4 |

*Instructions for interviewers:*

Q45 Notice that this question ends with a question mark. That means that you are not to read the answer categories. Rather, wait for R to respond and circle the appropriate number.

*Source:* Ross (1990).

phone survey of low-income women in a public health program (Schutt & Fawcett 2005), the University of Massachusetts Center for Survey Research (CSR) achieved a 55.1% response rate from all eligible sampled clients after a protocol that included as many as 30 contact attempts, although the response rate rose to 72.9 when it was calculated as a percentage of clients whom CSR was able to locate (Roman 2005:7). Response rates can be much lower in populations that are young, less educated, and poor. Those who do respond are more likely to be engaged in civic issues than those who do not respond, so estimates of related attitudes and behaviors in phone surveys can be quite biased (Kohut et al. 2012).

## In-Person Interviews

**In-person interview:** A survey in which an interviewer questions respondents face-to-face and records their answers.

What is unique to the **in-person interview**, compared with the other survey designs, is the face-to-face social interaction between interviewer and respondent. If money is no object, in-person interviewing is often the best survey design.

In-person interviewing has several advantages: Response rates are higher than with any other survey design; questionnaires can be much longer than with mailed or phone surveys; the questionnaire can be complex, with both open-ended and closed-ended questions and frequent branching patterns; the order in which questions are read and answered can be controlled by the interviewer; the physical and social circumstances of the interview can be monitored; and respondents' interpretations of questions can be probed and clarified.

But researchers must be alert to some special hazards resulting from the presence of an interviewer. Respondents should experience the interview process as a personalized interaction with an interviewer who



is very interested in the respondent's experiences and opinions. At the same time, however, every respondent should have the same interview experience—asked the same questions in the same way by the same type of person, who reacts similarly to the answers (de Leeuw 2008:318). Therein lies the researcher's challenge—to plan an interview process that will be personal and engaging and yet consistent and nonreactive (and to hire interviewers who can carry out this plan). Careful training and supervision are essential because small differences in intonation or emphasis on particular words can alter respondents' interpretations of questions' meaning (Groves 1989:404–406; Peterson 2000:24). Without a personalized approach, the rate of response will be lower and answers will be less thoughtful—and potentially less valid. Without a consistent approach, information obtained from different respondents will not be comparable—less reliable and less valid.

### Balancing Rapport and Control

Adherence to some basic guidelines for interacting with respondents can help interviewers maintain an appropriate balance between personalization and standardization:

- Project a professional image in the interview: that of someone who is sympathetic to the respondent but nonetheless has a job to do.
- Establish rapport at the outset by explaining what the interview is about and how it will work and by reading the consent form. Ask the respondent if he or she has any questions or concerns, and respond to these honestly and fully. Emphasize that everything the respondent says is confidential.
- During the interview, ask questions from a distance that is close but not intimate. Stay focused on the respondent and make sure that your posture conveys interest. Maintain eye contact, respond with appropriate facial expressions, and speak in a conversational tone of voice.
- Be sure to maintain a consistent approach; deliver each question as written and in the same tone of voice. Listen empathetically, but avoid self-expression or loaded reactions.
- Repeat questions if the respondent is confused. Use nondirective probes—such as “Can you tell me more about that?”—for open-ended questions.

As with phone interviewing, computers can be used to increase control of the in-person interview. In a **computer-assisted personal interviewing (CAPI)** project, interviewers carry a laptop computer that is programmed to display the interview questions and to process the responses that the interviewer types in, as well as to check that these responses fall within allowed ranges (Tourangeau 2004:790–791). Interviewers seem to like CAPI, and the data obtained are comparable in quality to data obtained in a noncomputerized interview (Shepherd et al. 1996). A CAPI approach also makes it easier for the researcher to develop skip patterns and experiment with different types of questions for different respondents without increasing the risk of interviewer mistakes (Couper et al. 1998).

**Computer-assisted personal interview (CAPI):** A personal interview in which the laptop computer is used to display interview questions and to process responses that the interviewer types in, as well as to check that these responses fall within allowed ranges.

The presence of an interviewer may make it more difficult for respondents to give honest answers to questions about socially undesirable behaviors such as drug use, sexual activity, and not voting (Schaeffer & Presser 2003:75). CAPI is valued for this reason because respondents can enter their answers directly in the laptop without the interviewer knowing what their response is. Alternatively, interviewers can simply hand respondents a separate self-administered questionnaire containing the more sensitive questions. After answering these questions, the respondent seals the separate questionnaire in an envelope so that the interviewer does not know the answers. When this approach was used for the GSS questions about sexual activity, about 21% of men and 13% of women who were married or had been married admitted to having cheated on a spouse (“Survey on Adultery” 1993:A20). The degree of rapport becomes a special challenge when survey questions concern issues

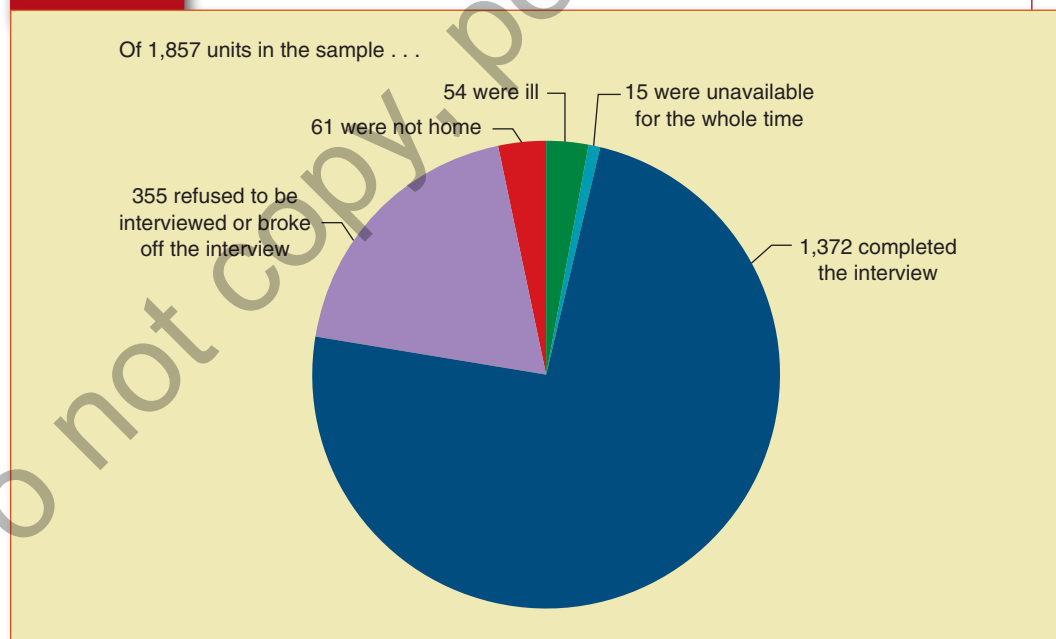
related to such demographic characteristics as race or gender (Groves 1989). If the interviewer and respondent are similar on the characteristics at issue, the responses to these questions may differ from those that would be given if the interviewer and respondent differ on these characteristics. For example, a white respondent may not disclose feelings of racial prejudice to a black interviewer that he would admit to a white interviewer.

Although in-person interview procedures are typically designed with the expectation that the interview will involve only the interviewer and the respondent, one or more other household members are often within earshot. In a mental health survey in Los Angeles, for example, almost half the interviews were conducted in the presence of another person (Pollner & Adams 1994). It is reasonable to worry that this third-party presence will influence responses about sensitive subjects—even more so because the likelihood of a third party being present may correspond with other subject characteristics. For example, in the Los Angeles survey, another person was present in 36% of the interviews with Anglos, in 47% of the interviews with African Americans, and in 59% of the interviews with Hispanics. However, there is no consistent evidence that respondents change their answers because of the presence of another person. Analysis of this problem with the Los Angeles study found very little difference in reports of mental illness symptoms between respondents who were alone and those who were in the presence of others.

### Maximizing Response to Interviews

Even if the right balance has been struck between maintaining control over interviews and achieving good rapport with respondents, in-person interviews still can be problematic. Because of the difficulty of finding all the members of a sample, response rates may suffer. Exhibit 8.15 displays the breakdown of nonrespondents to the 1990 GSS. Of the total original sample of 2,165, only 86% (1,857) were determined to be valid selections of dwelling units with potentially eligible respondents. Among these potentially eligible respondents, the response

**Exhibit 8.15** Reasons for Nonresponse in Personal Interviews (1990 General Social Survey)



Source: Data from Davis and Smith (1992):54XX.

rate was 74%. The GSS is a well-designed survey using carefully trained and supervised interviewers, so this response rate indicates the difficulty of securing respondents from a sample of the general population even when everything is done “by the book.”

Several factors affect the response rate in interview studies. Contact rates tend to be lower in central cities partly because of difficulties in finding people at home and gaining access to high-rise apartments and partly because of interviewer reluctance to visit some areas at night, when people are more likely to be home (Fowler 1988:45–60). Single-person households also are more difficult to reach, whereas households with young children or elderly adults tend to be easier to contact (Groves & Couper 1998:119–154).

Refusal rates vary with some respondents' characteristics. People with less education participate somewhat less in surveys of political issues (perhaps because they are less aware of current political issues). Less education is also associated with higher rates of “Don't know” responses (Groves 1989). High-income persons tend to participate less in surveys about income and economic behavior (perhaps because they are suspicious about why others want to know about their situation). Unusual strains and disillusionment in a society can also undermine the general credibility of research efforts and the ability of interviewers to achieve an acceptable response rate. These problems can be lessened with an advance letter introducing the survey project and by multiple contact attempts throughout the day and evening, but they cannot entirely be avoided (Fowler 1988:52–53; Groves & Couper 1998). Encouraging interviewers to tailor their response when potential respondents express reservations about participating during the initial conversation can also lead to lower rates of refusal: Making small talk to increase rapport and delaying asking a potential respondent to participate may reduce the likelihood of a refusal after someone first expresses uncertainty about participating (Maynard, Freese, & Schaeffer 2010:810).

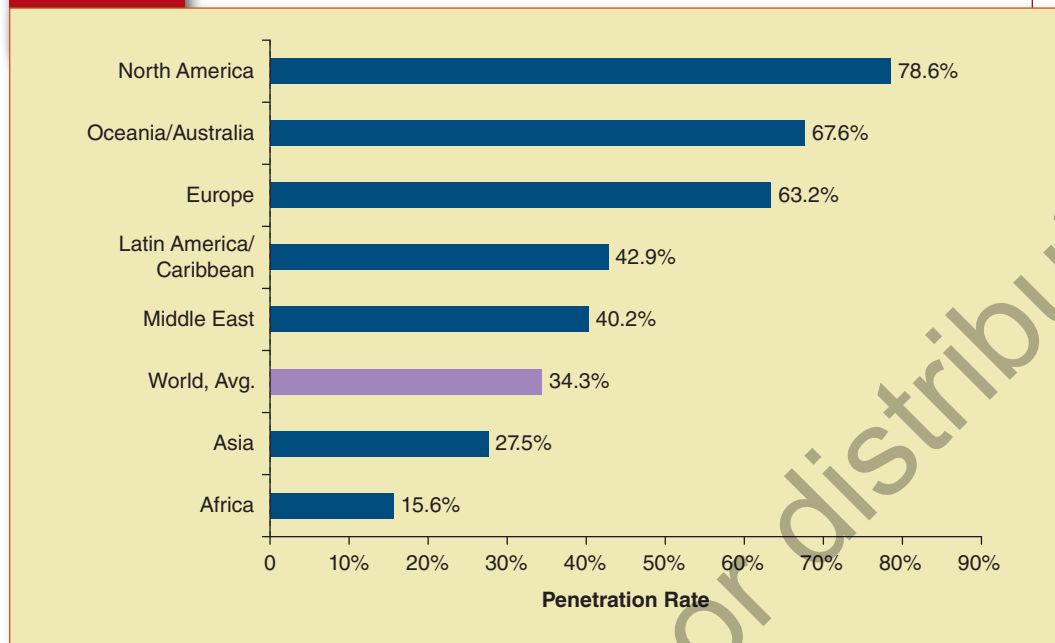
## Electronic Surveys

Widespread use of e-mail and the Internet, increasingly with high-speed connections and often through smartphones, creates new opportunities for survey researchers. Surveys can be e-mailed to respondents and returned in the same way; they can be posted on a website, and they can even be designed for completion on a smartphone. I will focus in this section on the currently most popular electronic survey approach, web surveys, to illustrate both the advantages and the limitations of these approaches.

**Web surveys** have become an increasingly useful survey method for two reasons: growth in the fraction of the population using the Internet and technological advances that make web survey design relatively easy. Many specific populations have very high rates of Internet use, so a web survey can be a good option for groups such as professionals, middle-class communities, members of organizations, and, of course, college students. Because of the Internet's global reach, web surveys also make it possible to conduct large, international surveys. However, coverage remains a major problem with many populations (Tourangeau et al. 2012). About one quarter of U.S. households are not connected to the Internet (File 2013b), so it is not yet possible to survey directly a representative sample of the U.S. population on the web—and given a plateau in the rate of Internet connections, this coverage problem may persist for the near future (Couper & Miller 2008:832). Rates of Internet usage are much lower in other parts of the world, with a worldwide average of 34.3% and rates as low as 15.6% in Africa and 27.5% averaged across all of Asia (see Exhibit 8.16; Internet World Statistics 2012). Households without Internet access also tend to be older, poorer, and less educated than do those that are connected, so web surveys of the general population can result in seriously biased estimates (File 2013b; Pew Research Center 2013). Coverage problems can be compounded in web surveys because of much lower rates of survey completion: It is just too easy to stop working on a web survey—much easier than it is to break off interaction with an interviewer (Tourangeau et al. 2012).

The extent to which the population of interest is connected to the web is the most important consideration when deciding whether to conduct a survey through the web. Other considerations that may increase the attractiveness of a web survey include the need for a large sample, for rapid turnaround, for collecting sensitive

**Web survey:** A survey that is accessed and responded to on the World Wide Web.

**Exhibit 8.16** Worldwide Internet Penetration Rates by Region, 2012

Source: Internet World Stats, [www.internetworldstats.com/stats.htm](http://www.internetworldstats.com/stats.htm). Copyright © 2012, Miniwatts Marketing Group. Reprinted with permission.

information that might be embarrassing to acknowledge in person, the availability of an e-mail list of the population, and the extent to which the interactive and multimedia features will enhance interest in the survey (Sue & Ritter 2012:10–11). Jennie Connor, Andrew Gray, and Kypros Kypri (2010) achieved a 63% response rate with a web survey about substance use that began with an initial e-mail invitation to a representative sample of undergraduate students at six New Zealand campuses.

There are several different approaches to engaging people in web surveys, each with unique advantages and disadvantages and somewhat different effects on the coverage problem. Many web surveys begin with an e-mail message to potential respondents that contains a direct “hotlink” to the survey website (Gaiser & Schreiner 2009:70). It is important that such e-mail invitations include a catchy phrase in the subject line as well as attractive and clear text in the message itself (Sue & Ritter 2012:110–114). This approach is particularly useful when a defined population with known e-mail addresses is to be surveyed. The researcher can then send e-mail invitations to a representative sample without difficulty. To ensure that the appropriate people respond to a web survey, researchers may require that respondents enter a personal identification number (PIN) to gain access to the web survey (Dillman 2000:378; Sue & Ritter 2012:103–104). Connor, Gray, and Kypri (2010:488) used this approach in their survey of New Zealand undergraduates:

Selected students received a letter which invited them to participate in an internet-based survey as part of the Tertiary Student Health Project, and provided a web address for the survey form. Details of the recruitment and data collection methods have been described in detail previously. Data were collected via a confidential online computerised survey that was completed at a time and place of the respondent's choice.

However, lists of unique e-mail addresses for the members of defined populations generally do not exist outside of organizational settings. Many people have more than one e-mail address, and often there is no apparent link between an e-mail address and the name or location of the person to whom it is assigned. As a result, there is no available method for drawing a random sample of e-mail addresses for people from any general population, even if the focus is only on those with Internet access (Dillman 2007:449).

Web surveys that use volunteer samples may instead be linked to a website that is used by the intended population and everyone who visits that site is invited to complete the survey. This was the approach used in the international web survey sponsored by the National Geographic Society in 2000 (Witte, Amoroso, & Howard 2000). However, although this approach can generate a very large number of respondents (50,000 persons completed Survey 2000), the resulting sample will necessarily reflect the type of people who visit that website (middle class, young North Americans, in Survey 2000) and thus be a biased representation of the larger population (Couper 2000:486–487; Dillman 2000:355). Some control over the resulting sample can be maintained by requiring participants to meet certain inclusion criteria (Selm & Jankowski 2006:440).

Coverage bias can also be a problem with web surveys that are designed for a population with high levels of Internet use. If the topic of the survey leads some people to be more likely to respond on the web, the resulting sample can be very unrepresentative. William Wells, Michael Cavanaugh, Jeffrey Bouffard, and Matt Nobles (2012:461) identified this problem in a comparison of attitudes of students responding to a web survey about gun violence with students at the same university who responded to the same survey administered in classes. Here is their e-mail survey introduction to potential respondents:

Recently, in response to shootings on university campuses like Virginia Tech and Northern Illinois University, several state legislatures (South Dakota, Texas, Washington) have begun debating whether to change rules banning students and employees from carrying concealed weapons on campus. This is an important public safety issue and the faculty in . . . are interested in knowing how people on this campus feel about it.

Students who responded to the web survey were much more likely to support the right to carry concealed weapons on campus than were those who responded in the classroom survey. In general, having a more extreme attitude motivated people to participate.

Some web surveys are designed to reduce coverage bias by providing computers and Internet connections to those who do not have them. This design-based recruitment method begins by contacting people by phone and providing those who agree to participate with whatever equipment they lack. This approach considerably increases the cost of the survey, so it is normally used as part of creating the panel of respondents who agree to be contacted for multiple surveys over time. The start-up costs can then be spread across many surveys. GfK Knowledge Networks is a company that received funding from the U.S. National Science Foundation to create such a web survey panel. CentER Data in the Netherlands also uses this panel approach (Couper & Miller 2008:832–833). Another approach to reducing coverage bias in web surveys is to recruit a volunteer panel of Internet users and then weight the resulting sample to make it comparable to the general population in such demographics as gender, race, age, and education. This is the method adopted by many market research organizations (Couper & Miller 2008:832–833); although response rates to volunteer samples are very low and the participants are often unlike the general population, it appears that weighting can reduce coverage bias by 30% to 60% (Tourangeau et al. 2012).

Of course, coverage bias is not as important when a convenience sample will suffice for an exploratory survey about some topic. Audrey Freshman (2012:41) used a web survey of a convenience sample to study symptoms of posttraumatic stress disorder (PTSD) among victims of the Bernie Madoff financial scandal.

This convenience, nonprobability sample was solicited via direct link to the study placed in online Madoff survivor support groups and comment sections of newspapers and blogs dealing with the event. The study announcement encouraged victims to forward the link to other former investors who might be interested in responding to the survey, thereby creating a snowball effect. The link led directly to a study description and enabled respondents to give informed consent prior to study participation. Participants were assured of anonymity of their responses and were instructed how to proceed in the event of increased feelings of distress as a result of study material. The survey was presumed to take approximately five to 10 minutes to complete. (p. 41)

Although a majority of respondents met clinical criteria for a diagnosis of PTSD, there is no way to know if this sample represents the larger population of Madoff's victims.

In contrast to problems of coverage, web surveys have some unique advantages for increasing measurement validity (Selm & Jankowski 2006; Tourangeau et al. 2012). Questionnaires completed on the web can elicit more



honest reports about socially undesirable behavior or experiences, including illicit behavior and victimization in the general population and failing course grades among college students, when compared with results with phone interviews (Kreuter, Presser, & Tourangeau 2008; Parks, Pardi, & Bradizza 2006). Onoye and colleagues (2012) found that conducting a survey on the web increased self-reports of substance use compared with a paper-and-pencil survey. Web surveys are relatively easy to complete because respondents simply click on response boxes and the survey can be programmed to move respondents easily through sets of questions, not presenting questions that do not apply to the respondent, thus leading to higher rates of item completion (Kreuter et al. 2008).

Use of the visual, interactive web medium can also help. Pictures, sounds, and animation can be used as a focus of particular questions and graphic and typographic variation can be used to enhance visual survey appeal (see Exhibit 8.17). Definitions of terms can also “pop up” when respondents scroll over them (Dillman 2007:458–459). In these ways, a skilled web programmer can generate a survey layout with many attractive features that make it more likely that respondents will give their answers—and have a clear understanding of the question (Smyth et al. 2004:4–5). Responses can quickly be checked to make sure they fall within the allowable range. Because answers are recorded directly in the researcher’s database, data entry errors are almost eliminated and results can be reported quickly. By taking advantage of these features, Titus Schleyer and Jane Forrest (2000:420) achieved a 74% response rate in a survey of dental professionals who were already Internet users.

Despite some clear advantages of some types of web surveys, researchers who use this method must be aware of some important disadvantages. Coverage bias is the single biggest problem with web surveys of the general population and of segments of the population without a high level of Internet access, and none of the different web survey methods fully overcome this problem. Weighting web survey panels of Internet users by demographic and other characteristics does not result in similar responses on many questions with those that are obtained from a mailed survey to a sample of the larger population (Rookey, Hanway, & Dillman 2008). Although providing Internet access to all who agree to participate in a web survey panel reduces coverage bias, many potential respondents do not agree to participate in such surveys: The rate of agreement to participate was 57% in one Knowledge Networks survey and just 41.5% in a survey of students at the University of Michigan (Couper 2000:485–489). Only about one third of Internet users contacted in phone surveys agree to provide an e-mail address for a web survey and then only one third of those actually complete the survey (Couper 2000:488). Web surveys that take more than 15 minutes are too long for most respondents (de Leeuw 2008:322). Surveys by phone continue to elicit higher rates of response (Kreuter et al. 2008). Some researchers have found that when people are sent a mailed survey that also provides a link to a web survey alternative, they overwhelmingly choose the paper survey (Couper 2000:488).

Despite their advantages for measurement, visual and other highlights that are possible in web surveys should be used with caution to avoid unintended effects on interpretation of questions and response choices (Tourangeau et al. 2012). For example, respondents tend to believe that a response in the middle is the typical response, that responses near each other are related, and that things that look alike are similar. Even minor visual cues can make a difference in responses. In one survey, 5% of respondents shifted their response when one response was given more space relative to others.

Surveys are also now being conducted through social media such as Facebook, on smartphones, and via text messages (Sue & Ritter 2012:119–122). Research continues into the ways that the design of web surveys can influence rates of initial response, the likelihood of completing the survey, and the validity of the responses (Couper, Traugott, & Lamias 2001; Kreuter et al. 2008; Porter & Whitcomb 2003; Tourangeau et al. 2012). At this point, there is reason enough to consider the option of a web survey for many investigations, but proceed with

caution and consider carefully their strengths and weaknesses when designing a web survey of any type and when analyzing findings from it.

**Mixed-mode survey:** A survey that is conducted by more than one method, allowing the strengths of one survey design to compensate for the weaknesses of another and maximizing the likelihood of securing data from different types of respondents; for example, non-respondents in a mailed survey may be interviewed in person or over the phone.

## Mixed-Mode Surveys

Survey researchers increasingly are combining different survey designs to improve the overall participation rate and to take advantage of the unique strengths of different methods. **Mixed-mode surveys** allow the strengths of one survey design to compensate for the weaknesses of another, and they can maximize the likelihood of securing



### Exhibit 8.17 Survey Monkey Web Survey Example

**Demographic**

Please provide us with some background information to help us design the website around your needs.

1. What kind of connection do you have to the Internet?

28.8 Kbps modem  
 56 Kbps modem  
 ISDN  
 Cable Modem  
 DSL  
 Do not know

2. Where are you located?

United States  
 Other Country

Please Specify

Next

---

**Future Directions**

3. The following features may be added to this site in the near future. Please rate the importance of the following features.

	Very Important	Important	Somewhat Important	Not Important
Increased Multi-lingual Support	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Multiple Users per Account	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Graphical Charts of Response Data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Increased Export Functionality	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Prev Next

Source: Survey Monkey.

data from different types of respondents (Dillman 2007:451–453; Selm & Jankowski 2006). For example, a survey may be sent electronically to sample members who have e-mail addresses and mailed to those who don't. Phone reminders may be used to encourage responses to web or paper surveys, or a letter of introduction may be sent in advance of calls in a phone survey (Guterbock 2008). Alternatively, nonrespondents in a mailed survey may be interviewed in person or over the phone. In one comparative study, the response rate to a telephone survey rose from 43% to 80% when it was followed by a mailed questionnaire (Dillman 2007:456). Kristen Olson, Jolene Smyth, and Heather Wood (2012) at the University of Nebraska–Lincoln found that providing a survey in the mode that potential respondents preferred—either phone, mailed, or web—increased the overall rate of participation by a small amount. As noted previously, an interviewer may also mix modes by using a self-administered questionnaire to present sensitive questions to a respondent in an in-person interview.

The mixed-mode approach is not a perfect solution. Rebecca Medway and Jenna Fulton (2012) reviewed surveys that gave the option of responding to either a mailed questionnaire or a web questionnaire and found that this reduced the response rate compared with using only a mailed questionnaire. Perhaps the need to choose between the modes or the delay in deciding to start the web survey led some potential respondents not to bother.

Respondents to the same question may give different answers because of the survey mode, rather than because they actually have different opinions. For example, when equivalent samples were asked by phone or mail, “Is the gasoline shortage real or artificial?” many more phone respondents than mail respondents answered that it was “very real” (Peterson 2000:24). Respondents to phone survey questions tend to endorse more extreme responses to scalar questions (which range from more to less) than do respondents to mail or web surveys (Dillman 2007:456–457). Responses may also differ between questions—one third of the questions in one survey—when asked in web and phone survey modes, even with comparable samples (Rookey et al. 2008:974). When responses differ by survey mode, there is often no way to know which responses are more accurate, although it appears that web surveys are likely to result in more admissions of socially undesirable experiences (Kreuter et al. 2008; Peterson 2000:24). Use of the same question structures, response choices, and skip instructions across modes substantially reduces the likelihood of mode effects, as does using a small number of response choices for each question (Dillman 2000:232–240; Dillman & Christian 2005), but web survey researchers are only beginning to identify the effect of visual appearance on the response to questions (Dillman 2007:472–487).

## A Comparison of Survey Designs



Which survey design should be used when? Group-administered surveys are similar, in most respects, to mailed surveys, except that they require the unusual circumstance of having access to the sample in a group setting. We therefore don't need to consider this survey design by itself; what applies to mailed surveys applies to group-administered survey designs, with the exception of sampling issues. The features of mixed-mode surveys depend on the survey types that are being combined. Thus, we can focus our comparison on the four survey designs that involve the use of a questionnaire with individuals sampled from a larger population: (1) mailed surveys, (2) phone surveys, (3) in-person surveys, and (4) electronic surveys. Exhibit 8.18 summarizes their strong and weak points.

The most important consideration in comparing the advantages and disadvantages of the four methods is the likely response rate they will generate. Mailed surveys must be considered the least preferred survey design from a sampling standpoint, although declining rates of response to phone surveys are changing this comparison.

Contracting with an established survey research organization for a phone survey is often the best alternative to a mailed survey. The persistent follow-up attempts that are necessary to secure an adequate response rate are much easier over the phone than in person. But, as explained earlier, the process requires an increasing number of callbacks to many households and rates of response have been declining. Current federal law prohibits automated dialing of cell phone numbers, so it is very costly to include the growing number of cell phone-only individuals in a phone survey.

In-person surveys are preferable in the possible length and complexity of the questionnaire itself, as well as with respect to the researcher's ability to monitor conditions while the questionnaire is completed. Mailed surveys often are preferable for asking sensitive questions, although this problem can be lessened in an interview by giving respondents a separate sheet to fill out or a laptop in which to enter their answers. Although interviewers may themselves distort results, either by changing the wording of questions or by failing to record answers properly, survey research organizations can reduce this risk through careful interviewer training and monitoring. Some survey supervisors will have interviews tape recorded so that they can review the dialogue between interviewer and respondents and provide feedback to the interviewers to help improve their performance. Some survey organizations have also switched to having in-person interviews completed entirely by the respondents on a laptop as they listen to prerecorded questions.

A phone survey limits the length and complexity of the questionnaire but offers the possibility of very carefully monitoring interviewers (Dillman 1978; Fowler 1988:61–73):

Supervisors in [one organization's] Telephone Centers work closely with the interviewers, monitor their work, and maintain records of their performance in relation to the time schedule, the quality of their work, and help detect and correct any mistakes in completed interviews prior to data reduction and processing. (J. E. Blair, personal communication to C. E. Ross, April 10, 1989)



**Exhibit 8.18 Advantages and Disadvantages of the Four Survey Designs**

Characteristics of Design	Mail Survey	Phone Survey	In-Person Survey	Web Survey
Representative sample				
Opportunity for inclusion is known				
For completely listed populations	High	High	High	Medium
For incompletely listed populations	Medium	Medium	High	Low
Selection within sampling units is controlled (e.g., specific family members must respond)	Medium	High	High	Low
Respondents are likely to be located				
If samples are heterogeneous	Medium	Medium	High	Low
If samples are homogeneous and specialized	High	High	High	High
Questionnaire construction and question design				
Allowable length of questionnaire	Medium	Medium	High	Medium
Ability to include				
Complex questions	Medium	Low	High	High
Open questions	Low	High	High	Medium
Screening questions	Low	High	High	High
Tedious, boring questions	Low	High	High	Low
Ability to control question sequence	Low	High	High	High
Ability to ensure questionnaire completion	Medium	High	High	Low
Distortion of answers				
Odds of avoiding social desirability bias	High	Medium	Low	High
Odds of avoiding interviewer distortion	High	Medium	Low	High
Odds of avoiding contamination by others	Medium	High	Medium	Medium
Administrative goals				
Odds of meeting personnel requirements	High	High	Low	Medium
Odds of implementing quickly	Low	High	Low	High
Odds of keeping costs low	High	Medium	Low	High

Source: Adapted from Dillman (1978):74–75. *Mail and Telephone Surveys: The Total Design Method*. Reprinted by permission of John Wiley & Sons, Inc.

People interviewed by phone tend to be less interested in the survey than do those interviewed in person, so they tend to satisfice more—apparently in a desire to complete the survey more quickly—and they tend to be less trusting of the survey motives (Holbrook et al. 2003).

The advantages and disadvantages of electronic surveys must be weighed in light of the population that is to be surveyed and capabilities at the time that the survey is to be conducted. At this time, too many people lack Internet connections for survey researchers to use the Internet to survey the general population.

These various points about the different survey designs lead to two general conclusions. First, in-person interviews are the strongest design and generally preferable when sufficient resources and a trained interview staff are available; telephone surveys have many of the advantages of in-person interviews at much less cost, but response rates are an increasing problem. Second, the “best” survey design for any particular study will be determined by the study’s unique features and goals rather than by any absolute standard of what the best survey design is.

## ▣ Ethical Issues in Survey Research

Survey research usually poses fewer ethical dilemmas than do experimental or field research designs. Potential respondents to a survey can easily decline to participate, and a cover letter or introductory statement that identifies the sponsors of, and motivations for, the survey gives them the information required to make this decision. The methods of data collection are quite obvious in a survey, so little is concealed from the respondents. Only in group-administered surveys might the respondents be, in effect, a captive audience (probably of students or employees), and so these designs require special attention to ensure that participation is truly voluntary. (Those who do not wish to participate may be told they can just hand in a blank form.)

Current federal regulations to protect human subjects allow survey research to be exempted from formal review unless respondents can be identified and disclosure of their responses could place them at risk. Specifically, the Code of Federal Regulations (2009) limits the survey exemption this way:

(2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless:

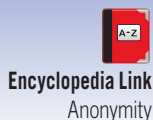
(i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects’ responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects’ financial standing, employability, or reputation. (p. 46.101(b)2)

**Confidentiality:** Provided by research in which identifying information that could be used to link respondents to their responses is available only to designated research personnel for specific research needs.

**Confidentiality** is most often the primary focus of ethical concern in survey research. Many surveys include some essential questions that might, in some way, prove damaging to the subjects if their answers were disclosed. To prevent any possibility of harm to subjects because of the disclosure of such information, the researcher must preserve subject confidentiality. Nobody but research personnel should have access to information that could be used to link respondents to their responses, and even that access should be limited to what is necessary for specific research purposes.

Only numbers should be used to identify respondents on their questionnaires, and the researcher should keep the names that correspond to these numbers in a safe, private location, unavailable to staff and others who might otherwise come across them. Follow-up mailings or contact attempts that require linking the ID numbers with names and addresses should be carried out by trustworthy assistants under close supervision. For electronic surveys, encryption technology should be used to make information provided over the Internet secure from unauthorized persons.

Mirowsky and Ross (1999) focused special attention on the maintenance of respondent confidentiality because of the longitudinal nature of the ASOC survey. To recontact people for interviews after the first wave in 1995, they had to keep files with identifying information. Here is how Mirowsky (1999) described their procedures in their application for funding of the third wave (the 2001 survey):



In order to follow respondents, their phone numbers and first names were recorded. They also were asked to give the phone number and first name of someone outside the household who would know how to contact them if they moved. Multiple attempts were made to contact each respondent at 18 months after the initial interview, as described in the section on tracking in the progress report. Those contacted were asked again about someone who will know how to contact them if they move. If a respondent is not reached at the old number, the outside contact is called and asked to help locate the person.

The privacy of respondents is maintained by having separate groups of people collect data and analyze it. The staff of the Survey Research Laboratory of the University of Illinois at Chicago collects the data. Computer-aided interviewing software gave each respondent a sequence number in two separate data sets. One data set consists solely of the first names, phone numbers, and contacts that could identify a respondent. That data set is secured by the Survey Research Laboratory. It is only available to laboratory staff when it is needed for follow-up. It is not available at any time to the principal investigator, his associates, or anyone else not on the survey laboratory interviewing staff. The second data set contains disposition codes and responses to questions. It does not contain any information that could identify individual respondents. The second data set is available to the principal investigator and his associates for analysis. (p. 41)

Not many surveys can provide true **anonymity**, so that no identifying information is ever recorded to link respondents with their responses. The main problem with anonymous surveys is that they preclude follow-up attempts to encourage participation by initial nonrespondents, and they prevent panel designs, which measure change through repeated surveys of the same individuals. In-person surveys rarely can be anonymous because an interviewer must, in almost all cases, know the name and address of the interviewee. However, phone surveys that are meant only to sample opinion at one point in time, as in political polls, can safely be completely anonymous. When no future follow-up is desired, group-administered surveys also can be anonymous. To provide anonymity in a mail survey, the researcher should omit identifying codes from the questionnaire but could include a self-addressed, stamped postcard so the respondent can notify the researcher that the questionnaire has been returned without creating any linkage to the questionnaire itself (Mangione 1995:69).

**Anonymity:** Provided by research in which no identifying information is recorded that could be used to link respondents to their responses.



**Video Link**  
Using Survey Data

## ▣ Conclusions

Survey research is an exceptionally efficient and productive method for investigating a wide array of social research questions. Mirowsky and Ross (2003) and Mirowsky (1999) were able to survey representative samples of Americans and older Americans and follow them for 6 years. These data allowed Mirowsky and Ross to investigate the relationships among education, social status, and health and how these relationships are changing.

In addition to the potential benefits for social science, considerations of time and expense frequently make a survey the preferred data collection method. One or more of the six survey designs reviewed in this chapter (including mixed mode) can be applied to almost any research question. It is no wonder that surveys have become the most popular research method in sociology and that they frequently inform discussion and planning about important social and political questions. As use of the Internet increases, survey research should become even more efficient and popular.

The relative ease of conducting at least some types of survey research leads many people to imagine that no particular training or systematic procedures are required. Nothing could be further from the truth. But as a result of this widespread misconception, you will encounter a great many nearly worthless survey results. You must be prepared to examine carefully the procedures used in any survey before accepting its findings as credible. And if you decide to conduct a survey, you must be prepared to invest the time and effort that proper procedures require.



## Key Terms

Anonymity	299	Electronic survey	277	Mailed survey	278
Behavior coding	270	Fence-sitters	264	Mixed-mode survey	294
Bipolar response options	262	Filter question	259	Omnibus survey	255
Cognitive interview	269	Floater	264	Part-whole question effects	274
Computer-assisted personal interview (CAPI)	289	Forced-choice questions	264	Phone survey	282
Computer-assisted telephone interview (CATI)	287	Group-administered survey	282	Questionnaire	268
Confidentiality	298	Idiosyncratic variation	265	Skip pattern	259
Context effects	274	In-person interview	288	Social desirability bias	262
Contingent question	259	Interactive voice response (IVR)	287	Split-ballot design	255
Cover letter	279	Interpretive questions	270	Survey pretest	270
Double-barreled question	259	Interview schedule	268	Survey research	253
Double negative	259	Labeled unipolar response options	262	Unlabeled unipolar response options	262
		Likert item	261	Web survey	291

## Highlights

- Surveys are the most popular form of social research because of their versatility, efficiency, and generalizability. Many survey data sets, such as the GSS, are available for social scientists to use in teaching and research.
- Omnibus surveys cover a range of topics of interest and generate data useful to multiple sponsors.
- Survey designs must minimize the risk of errors of observation (measurement error) and errors of nonobservation (errors resulting from inadequate coverage, sampling error, and nonresponse). The likelihood of both types of error varies with the survey goals. For example, political polling can produce inconsistent results because of rapid changes in popular sentiment.
- Social exchange theory asserts that behavior is motivated by the return expected to the individual for the behavior. Survey designs must maximize the social rewards, minimize the costs of participating, and establish trust that the rewards will outweigh the costs.
- Questions must be worded carefully to avoid confusing respondents, encouraging a less-than-honest response, or triggering biases. Inclusion of “Don’t know” choices and neutral responses may help, but the presence of such options also affects the distribution of answers. Open-ended questions can be used to determine the meaning that respondents attach to their answers. Answers to any survey questions may be affected by the questions that precede them in a questionnaire or interview schedule.
- Sets of questions that comprise an index can reduce idiosyncratic variation in measurement of a concept. Indexes may be unidimensional or multidimensional. Responses to the questions in an index should be tested after data are collected to ensure that they can be combined as measures of a single concept, or of several related concepts, as intended.
- Questions can be tested and improved through review by experts, focus group discussions, cognitive interviews, behavior coding, and pilot testing. Every questionnaire and interview schedule should be pretested on a small sample that is like the sample to be surveyed.
- Interpretive questions should be used in questionnaires to help clarify the meaning of responses to critical questions.
- A survey questionnaire or interview schedule should be designed as an integrated whole, with each question and section serving some clear purpose and complementing the others.
- The cover letter for a mailed questionnaire should be credible, personalized, interesting, and responsible.
- Response rates in mailed surveys are typically well below 70% unless multiple mailings are made to nonrespondents and the questionnaire and cover letter are attractive, interesting, and

carefully planned. Response rates for group-administered surveys are usually much higher.

- Phone interviews using random digit dialing allow fast turn-around and efficient sampling. Multiple callbacks are often required, and the rate of nonresponse to phone interviews is rising. Phone interviews should be no more than 30 to 45 minutes. Response rates to phone surveys have declined dramatically due to cell phones and caller ID.
- In-person interviews have several advantages over other types of surveys: They allow longer and more complex interview schedules, monitoring of the conditions when the questions are answered, probing for respondents' understanding of the questions, and high response rates. However, the interviewer must balance the need to establish rapport with the respondent with the importance of maintaining control over the delivery of the interview questions.
- Electronic surveys may be e-mailed or posted on the web. Interactive voice response systems using the telephone are another option. At this time, use of the Internet is not sufficiently widespread to allow web surveys of the general population, but these approaches can be fast and efficient for populations with high rates of computer use.
- Mixed-mode surveys allow the strengths of one survey design to compensate for the weaknesses of another. However, questions and procedures must be designed carefully to reduce the possibility that responses to the same question will vary as a result of the mode of delivery.
- In deciding which survey design to use, researchers must consider the unique features and goals of the study. In general, in-person interviews are the strongest, but most expensive, survey design.
- Most survey research poses few ethical problems because respondents are able to decline to participate—an option that should be stated clearly in the cover letter or introductory statement. Special care must be taken when questionnaires are administered in group settings (to “captive audiences”) and when sensitive personal questions are to be asked; subject confidentiality should always be preserved.

### STUDENT STUDY SITE



Sharpen your skills with SAGE edge at [edge.sagepub.com/schutt8e](http://edge.sagepub.com/schutt8e). SAGE edge for students provides a personalized approach to help you accomplish your coursework goals in an easy-to-use learning environment.

## Discussion Questions

1. Response rates to phone surveys are declining, even as phone usage increases. Part of the problem is that lists of cell phone numbers are not available and wireless service providers may not allow outside access to their networks. Cell phone users may also have to pay for incoming calls. Do you think regulations should be passed to increase the ability of survey researchers to include cell phones in their random digit dialing surveys? How would you feel about receiving survey calls on your cell phone? What problems might result from “improving” phone survey capabilities in this way?
2. In-person interviews have for many years been the “gold standard” in survey research because the presence of an interviewer increases the response rate, allows better rapport with the interviewee, facilitates clarification of questions and instructions, and provides feedback about the interviewee’s situation. However, researchers who design in-person interviewing projects are now making increasing use of technology to ensure consistent questioning of respondents and to provide greater privacy for respondents answering questions. But having a respondent answer questions on a laptop while the interviewer waits is a very different social process than actually asking the questions verbally. Which approach would you favor in survey research? What trade-offs can you suggest there might be in quality of information collected, rapport building, and interviewee satisfaction?

3. Each of the following questions was used in a survey that I received at some time in the past. Evaluate each question and its response choices using the guidelines for writing survey questions presented in this chapter. What errors do you find? Try to rewrite each question to avoid such errors and improve question wording.

a. From an *Info World* (computer publication) product evaluation survey:

How interested are you in PostScript Level 2 printers?

\_\_\_\_Very \_\_\_\_Somewhat \_\_\_\_Not at all

b. From the Greenpeace National Marine Mammal Survey:

Do you support Greenpeace's nonviolent, direct action to intercept whaling ships, tuna fleets, and other commercial fishermen to stop their wanton destruction of thousands of magnificent marine mammals?

\_\_\_\_Yes \_\_\_\_No \_\_\_\_Undecided

c. From a U.S. Department of Education survey of college faculty:

How satisfied or dissatisfied are you with each of the following aspects of your instructional duties at this institution?

	Very Dissat.	Somewhat Dissat.	Somewhat Satisf.	Very Satisf.
a. The authority I have to make decisions about what courses I teach	1	2	3	4
b. Time available for working with students as advisor, mentor	1	2	3	4

d. From a survey about affordable housing in a Massachusetts community:

Higher than single-family density is acceptable to make housing affordable.

Strongly Agree	Undecided	Disagree	Strongly Agree	Disagree
1	2	3	4	5

e. From a survey of faculty experience with ethical problems in research:

Are you reasonably familiar with the codes of ethics of any of the following professional associations?

	Very Familiar	Familiar	Not Too Familiar
American Sociological Association	1	2	0
Society for the Study of Social Problems	1	2	0
American Society of Criminology	1	2	0

If you are familiar with any of the above codes of ethics, to what extent do you agree with them?

Strongly Agree Agree No opinion Disagree Strongly Disagree

Some researchers have avoided using a *professional code of ethics* as a guide for the following reason. Which responses, if any, best describe your reasons for not using all or any of parts of the codes?

	Yes	No
1. Vagueness	1	0
2. Political pressures	1	0
3. Codes protect only individuals, not groups	1	0

f. From a survey of faculty perceptions:

Of the students you have observed while teaching college courses, please indicate the percentage who significantly improve their performance in the following areas.

Reading \_\_\_\_%

Organization \_\_\_\_%

Abstraction \_\_\_\_%

g. From a University of Massachusetts Boston student survey:

A person has a responsibility to stop a friend or relative from driving when drunk.

Strongly Agree \_\_\_\_ Agree \_\_\_\_ Disagree \_\_\_\_ Strongly Disagree \_\_\_\_

Even if I wanted to, I would probably not be able to stop most people from driving drunk.

Strongly Agree \_\_\_\_ Agree \_\_\_\_ Disagree \_\_\_\_ Strongly Disagree \_\_\_\_

## Practice Exercises

1. Consider how you could design a split-ballot experiment to determine the effect of phrasing a question or its response choices in different ways. Check recent issues of the local newspaper for a question used in a survey of attitudes about some social policy or political position. Propose some hypothesis about how the wording of the question or its response choices might have influenced the answers people gave, and devise an alternative that differs only in this respect. Distribute these questionnaires to a large class (after your instructor makes the necessary arrangements) to test your hypothesis.
2. I received in my university mailbox some years ago a two-page questionnaire that began with the following cover letter at the top of the first page:
3. Test your understanding of survey research terminology by completing one set of Interactive Exercises from the study site on survey design. Be sure to review the text on the pages indicated in relation to any answers you missed.
4. Review this chapter's Research That Matters article on the book's study site, [edge.sagepub.com/schutt8e](http://edge.sagepub.com/schutt8e). Describe the sampling and measurement methods used and identify both strong and weak points of the survey design. Would a different type of survey design (in-person, phone, mailed, web) have had any advantages? Explain your answer.

### Faculty Questionnaire

This survey seeks information on faculty perception of the learning process and student performance in their undergraduate careers. Surveys have been distributed in universities in the Northeast, through random deposit in mailboxes of selected departments. This survey is being conducted by graduate students affiliated with the School of Education and the Sociology Department. We greatly appreciate your time and effort in helping us with our study.

Critique the "Faculty Questionnaire" cover letter, and then draft a more persuasive one.

## Ethics Questions

1. Group-administered surveys are easier to conduct than other types of surveys, but they always raise an ethical dilemma. If a teacher allows a social research survey to be distributed in his or her class, or if an employer allows employees to complete a survey on company time, is the survey truly voluntary? Is it sufficient to read a statement to the group members stating that their participation is entirely up to them? How would you react to a survey in your class? What general guidelines should be followed in such situations?
2. Patricia Tjaden and Nancy Thoennes (2000) sampled adults with random digit dialing to study violent victimization from a nationally representative sample of adults. What ethical dilemmas do you see in reporting victimizations that are identified in a survey? What about when the survey respondents are under the age of 18? What about children under the age of 12?

## Web Exercises

1. Who does survey research and how do they do it? These questions can be answered through careful inspection of ongoing surveys and the organizations that administer them at [www.ciser.cornell.edu/info/polls.shtml](http://www.ciser.cornell.edu/info/polls.shtml). Spend some time reading

about the different survey research organizations, and write a brief summary of the types of research they conduct, the projects in which they are involved, and the resources they offer on their websites. What are the distinctive features of different survey research organizations?

2. Go to the Research Triangle Institute site at [www.rti.org](http://www.rti.org). Click on “Survey Research & Services” then “Innovations.” Read about their methods for computer-assisted interviewing and their cognitive laboratory methods for refining

questions. What does this add to my treatment of these topics in this chapter?

3. Go to the UK Data Service at <http://discover.ukdataservice.ac.uk/variables>. In the search box, enter topics of interest such as “health” or “inequality.” Review five questions for two topic areas and critique them in terms of the principles for question writing that you have learned. Do you find any question features that might be attributed to the use of British English?

## Video Interview Questions

Listen to the researcher interview for Chapter 8 at [edge.sagepub.com/schutt8e](http://edge.sagepub.com/schutt8e).

1. What two issues should survey researchers consider when designing questions?
2. Why is cognitive testing of questions important?

## SPSS Exercises

What can we learn from the General Social Survey (GSS) data about the orientations of people who support capital punishment? Is it related to religion? Reflective of attitudes toward race? What about political views? Is it a guy thing? Do attitudes and behavior concerning guns have some relation to support for capital punishment?

1. To answer these questions, we will use some version of each of the following variables in our analysis: PARTYID3, GUNLAW, HELPBLK, RACDIF1, FUND, OWNGUN, and CAPPUN. Check the wording of each of these questions at the University of Michigan’s GSS website (click on “Browse GSS Variables” and use the mnemonic listing of variables to find those in the list above): [www.norc.org/GSS+Website](http://www.norc.org/GSS+Website)

How well does each of these questions meet the guidelines for writing survey questions? What improvements would you suggest?

2. Now generate cross-tabulations to show the relationship between each of these variables, treated as independent variables, and support for capital punishment. A cross-tabulation can be used to display the distribution of responses on the dependent variable for each category of the independent variable. For this purpose, you should substitute several slightly different versions of the variables you just reviewed. From the menu, select Analyze/Descriptive Statistics/Crosstabs:

*Rows:* CAPPUN

*Columns:* SEX, PARTYID3, GUNLAW, HELPBLK, RACDIF1, FUND, OWNGUN

*Cells:* column percentages

(If you have had a statistics course, you will also want to request the chi-square statistic for each of these tables.)

Describe the relationship you have found in the tables, noting the difference in the distribution of the dependent (row) variable—support for capital punishment—between the categories of each of the independent (column) variables.

3. Summarize your findings. What attitudes and characteristics are associated strongly with support for the death penalty?
4. What other hypotheses would you like to test? What else do you think needs to be considered to help you understand the relationships you have identified? For example, should you consider the race of the respondents? Why or why not?



5. Let's take a minute to learn about recoding variables. If you generate the frequencies for POLVIEWS and for POLVIEWS3, you'll see how I recoded POLVIEWS3. Why? Because I wanted to use a simple categorization by political party views in the cross-tabulation. You can try to replicate my recoding in SPSS. From the menu, click Transform/

Recode/Into different variables. Identify the old variable name and type in the new one. Type in the appropriate sets of old values and the corresponding new values. You may need to check the numerical codes corresponding to the old values with the variable list pull-down menu (the ladder icon with a question mark).

## Developing a Research Proposal

These steps focus again on the "Research Design" decisions, but this time assuming that you will use a survey design (Exhibit 3.10, #13 to #17).

1. Write 10 questions for a one-page questionnaire that concerns your proposed research question. Your questions should operationalize at least three of the variables on which you have focused, including at least one independent and one dependent variable (you may have multiple questions to measure some variables). Make all but one of your questions closed-ended. If you completed the "Developing a Research Proposal" exercises in Chapter 4, you can select your questions from the ones you developed for those exercises.
2. Conduct a preliminary pretest of the questionnaire by conducting cognitive interviews with two students or other persons like those to whom the survey is directed. Follow up the closed-ended questions with open-ended probes that ask the students what they meant by each response or what came to mind when they were asked each question. Account for the feedback you receive when you revise your questions.
3. Polish the organization and layout of the questionnaire, following the guidelines in this chapter. Prepare a rationale for the order of questions in your questionnaire. Write a cover letter directed to the appropriate population that contains appropriate statements about research ethics (human subjects' issues).