

Figure 2.2 The survey research triangle: Survey error, constraints, and effects

dominant
appr.

There is, of course, never a perfect paradigm for understanding everything in a field, and so there are other perspectives that can be used in approaching survey research as a field of study. However, the total survey error approach is certainly the dominant approach today, and part of its advantage is that some alternative approaches can be treated as part of its perspective. For example, one recent aspect of the scientific approach to surveys has been the application of increased understanding of the human thought process (the so-called cognitive revolution in psychology) to understanding the survey research process. Although this could be considered an alternative perspective, it can also be subsumed under the total survey error approach, as part of increasing our understanding of the errors in the measurement process associated with the respondent.

Finally, the total survey error approach provides a useful way to structure this book. It reminds us of the importance of identifying, discussing, measuring, and seeking ways to minimize all the factors that lead to errors in surveys. It emphasizes the need to balance error minimization against survey costs, time, and ethical considerations. Additionally, it recognizes that there are a number of survey-related effects that will always affect survey results. Figure 2.2 summarizes the main elements of these three factors, while implying the difficulty of balancing between them. Chapters 4 through 14 will focus on the various sources of survey error, survey constraints, and some survey effects.

3

Survey Modes

RESPONSES TO EMERGING TECHNOLOGIES

There are several different modes through which surveys can be administered: face-to-face, telephone, mail, Internet, self-administered, and so on. Before the sources of survey error are examined in detail, it is useful to give preliminary consideration to these different modes of survey administration. Two basic distinctions between survey modes will be developed in this chapter, after which some related research approaches are discussed. Finally, computer-assisted survey information collection and Internet surveys will be described in more detail.

Survey Modes

A variety of different survey modes have been used over the years, with much of the change reflecting changes in available technology. As will be seen, the different modes vary in their costs and time requirements, factors that must be considered when one is choosing an interviewing mode.

Two Important Distinctions

There are two important distinctions to make about survey modes (see table 3.1). One is how personal they are in terms of the extent of interviewer involvement: whether the survey is *self-administered* or *interviewer-administered*. Face-to-face and telephone surveys are generally interviewer administered. Face-to-face interviews are usually in the respondents' natural surroundings, as when interviewers go to people's homes or offices. Skillful interviewers can usually get a high response rate and can conduct

Table 3.1 Survey Modes by Administration Approaches

	Self-Administered	Interviewer-Administered
Not computerized (static)	SAQ: Self-administered questionnaires	Telephone surveys
	Mail surveys	Face-to-face surveys
	Static e-mail and Internet surveys	
Computerized (interactive)	Interactive Internet surveys	CATI: computer-assisted telephone interviewing
	CSAQ: computerized self- administered questionnaires	
	IVR: interactive voice response	CAPI: computer-assisted personal interviewing

very high-quality interviews. It is harder to exercise quality control because these interviews are usually not monitored, though even monitoring is possible with modern recording technology. The response rate and interview quality are generally lower in phone interviews. Telephone surveys can be conducted out of a central telephone facility (a "calling center"), which improves quality control since a supervisor can oversee the process. The data can be collected much faster than with mail surveys or face-to-face interviews, which is important to clients who want daily updates.

Interviewers provide several advantages in surveys. They can probe answers that are unclear or inadequate. Interviewers are essential if respondents' skills in reading, writing, and/or computer usage are too low to permit them to fill out self-administered surveys. Also, interviewers are especially useful in establishment surveys of business firms, since they can try to interview the desired respondent in a business rather than letting someone in the organization decide who will fill out a questionnaire. However, interviewers also add a source of error to surveys, as will be detailed in chapter 4, and these errors are eliminated when no interviewer is involved.

Survey modes that do not involve interviewers include self-administered questionnaires, mail surveys, Internet surveys, and interactive voice response surveys. These surveys afford respondents greater privacy, as no one can overhear answers and the respondents will not be embarrassed by their answers to an interviewer. Self-administered questionnaires, such as those given in classes, are often given to volunteer subjects and/or captive groups.¹ Unfortunately, that means that the respondents are not necessarily representative of the larger group of interest, except

when questionnaires are given to people as they undergo an experience (such as having people fill out a questionnaire as they complete a training program). Still, self-administered questionnaires are particularly useful when respondents need to check their records in order to answer questions, as when surveys of businesses have to answer detailed personnel and financial questions. Mail surveys can be sent out to a representative sample of the population of interest. Large numbers of mail questionnaires are usually sent out, in the hope that people will fill them out and mail them back. However, because of the impersonality of mail surveys, the response rate can be very low, and people who return them may not be representative of the larger population. There has been a considerable increase in self-administered surveys (Dillman 2000, 7–8). Many federal government surveys are now self-administered, for reasons such as lower cost and the ability to conduct them in-house without hiring a professional firm to conduct interviewer-administered surveys.

The other important distinction involves whether or not the survey is *computerized in an interactive sense*. Computerizing permits control of the flow of the questionnaire. For one thing, researchers often want the answer to one question to determine which question is asked next (for example, questions about the education of one's children can be skipped for people who do not have children). Computer-based surveys can show people the next question that is appropriate given their answer to the last question, whereas directions in a written questionnaire as to how the next question depends on answers to the last question can be confusing. As described below, computerized surveys also permit "survey experiments" in which random subsamples are asked different versions of the same question or the same questions in different orders to test the effects associated with question wording and question order.

As shown in table 3.1, computerized surveys can be interviewer administered or self-administered. Interviewer-administered computerized surveys can be conducted either on the telephone (computer-assisted telephone interviewing, CATI) or face to face with interviewers using personal or handheld computers (computer-assisted personal interviewing, CAPI). Computer-assisted surveys combine the advantages of an interviewer and of computerized control of the question flow.

Internet surveys have become the best-known form of self-administered computerized surveys. "Virtual interviewing" resembles self-administered questionnaires in that there is no interviewer. However, some of the less expensive Internet surveys involve only a static questionnaire posted on the Web, not taking advantage of the potential for controlling

the question flow in an interactive manner. It is also possible to conduct self-administered computerized surveys without using the Internet, either through automated phone surveys (now often using interactive voice response technology) or by having respondents fill out a questionnaire at a computer without going onto the Internet (known as CSAQ technology, for computerized self-administered questionnaires).

A third possible distinction between different survey modes involves the “channel of communication”—whether visual or aural communication is invoked. Self-administered questionnaires use only the visual channel, whereas telephone surveys just use the aural channel. Internet surveys are primarily visual, though some use voice as well. Face-to-face interviews are primarily aural but can include visual stimuli (as well as non-verbal communication from the interviewer). Surveys that are presented only visually are problematic for respondents who have difficulty reading, while surveys that are presented only aurally can challenge the respondents’ listening abilities. Some computerized surveys have people put on earphones to hear the survey while they read it on the computer screen, as a means of providing respondents both types of input.

The choice of survey mode can affect survey results. These effects are discussed in chapter 12, after survey procedures are discussed in detail in intervening chapters.

Measuring Change

In addition to the selection of a survey mode, an important design aspect of surveys involves whether and how to measure change over time. A *cross-section survey* takes a single one-time sample. Many surveys are simply single cross-sectional surveys with no intended time perspective, but survey projects are increasingly interested in assessing the amount of change that occurs in facts or attitudes.

There are several types of *longitudinal surveys* (Rose 2000a; Menard 2002), with the more complex designs permitting more complete assessment of change. Repeated cross-section surveys involve taking a series of separate surveys from the same sampling frame without reinterviewing the same people; this strategy allows an estimate of net change at the aggregate level but cannot show which people change. By contrast, a *longitudinal panel survey* involves reinterviews with the same people each time, a design that excels in looking at individual-level change but cannot take changes in the composition of the population into account.² The U.S. National Election Studies began a panel survey in 1990, with the first interview (known as the first “wave”) after the midterm election that year.

Reinterviews were conducted in 1991 after the first Gulf War and again before and after the 1992 presidential election.

These two approaches can be mixed in various ways. *Overlapping panels* bring new people into the study periodically, whereas *rotating panels* rotate people out of the survey after a certain number of interviews. A *revolving panel* drops subsamples for one period and then includes them again in a later period. Other options include following people when they move and maintaining records at the individual level across panel waves. A *split panel* combines a panel study with either a repeated or a rotating panel survey, though the combination limits the resources that could otherwise be used to buy a higher sample size for one of its components (Kalton and Citro 2000). The most complete form is the *longitudinal survey with rotation*, bringing in new cases periodically, following people when they move, and maintaining records for individuals across waves (Bailar 1989). Minimally, augmenting panel surveys with fresh cross-sections is useful for obtaining estimates of biases resulting from panel surveys (Bartels 2000). This strategy was followed in the 1992–97 NES panel study, in which some respondents were interviewed as many as seven times, but the panel was “refreshed” in 1994 and 1996 with the addition of new cross-sections to make up for the inevitable loss of respondents over repeated interviewing. The sources of bias in panel studies will be discussed in later chapters of this book.

The choice between cross-sectional and longitudinal designs has implications for the choice of survey mode. Long-term panel surveys need to track people when they move, and such tracking is facilitated when at least the initial contact with the respondent is through an interviewer. As will be seen later in this chapter, there are also some interesting experiments in maintaining ongoing short-term panel surveys over the Internet. In any case, when selecting a survey mode, researchers need to think about how best to obtain change measures if they are required.

Related Research Approaches

Several other research modes share some features of surveys. Indeed, participants in some of these modes are often chosen on the basis of a short phone survey. *Focus groups* select a small group of people to come together to answer questions posed by a discussion facilitator. The group discussion is meant to resemble natural decision processes more than surveys, though a person with a strong view on the topic being discussed can dominate the conversation and discourage others from expressing contrary opinions. The results are looked at qualitatively, not quantitatively as with

surveys, such as listening to see which ways to frame an advertising pitch for a product or political issue work best. *Deliberative polls* get together a large group of people to be given information about a policy problem and discuss possible solutions with each other. Deliberative polls are rare, but they are interesting because they allow people to discuss ideas before finalizing their opinions as people do in real life. *Audience reaction research* entails having people record their reactions to a presentation—for example, turning a dial to record how positive or negative their reactions are to each moment of a political advertisement.

Diaries have respondents keep track themselves of their behavior over a period of time. Prior to the advent of cable television, this was one of the main methods of measuring television viewing: people were asked to write down what they watched. Diaries have also been used to have people keep track of their food intake, their time use, and their household purchases. Diaries would be highly accurate if people filled them out as they performed these acts, but cooperation usually declines over time, with people recording fewer activities. Further, some people change their behavior because of the data collection (such as watching the TV shows they feel they should watch instead of those they would normally watch).

The sampling methods used in surveys are often employed in these research approaches, and surveys are often conducted in conjunction with them. For example, the 1996 National Issues Convention deliberative poll was based on interviews taken nationally by NORC with 915 people, with a completion rate of 72%. In the end, 50% of those respondents accepted paid trips to Austin, Texas, for three and a half days to discuss the public issues of that election year, and they filled out questionnaires at the beginning and end of those sessions so that opinion change could be measured (Merkle 1996).

Two additional research modes employ special forms of contacting survey respondents. *Exit polls* ask people leaving election polling places how they voted (Mitofsky 1991), while *intercept surveys* ask questions of people in malls or other settings. Exit polls and intercept surveys use specific sampling approaches in deciding which people to interview (see the discussion of systematic sampling in chapter 10).

Some survey modes are not considered scientifically valid. As an example, the *phone-in poll* has people phone one number if they want to vote yes on a topic and a different phone number if they want to vote no. Similarly, the *click-in poll* is a poll on an Internet site that a person happens to visit, such as the CNN Web site's QuickVote poll on its topic of the day. These polls are not considered valid, because only interested people participate

and because affected parties can distort the results by participating multiple times, though their advocates argue that these problems are not necessarily more serious than the problems of other survey modes.

Some pseudopolls are not ethically valid. The American Association of Public Opinion Research (AAPOR) has made strong policy statements against companies that claim to be conducting a survey as a means of gaining people's attention when the company has no interest in the person's answers. In particular, AAPOR has attacked the practices of using purported surveys as entrees for selling products (known as "sugging," for "selling under the guise" of a survey) or soliciting money (known as "frugging," for "fund raising under the guise" of a survey). Finally, a *push poll* is propaganda that looks like a poll. Some political candidates have hired firms to call people just before an election, claim they are conducting a poll, and then ask voters if they would vote for the candidate's opponent if they were told that he or she had done some terrible deed or supported some unpopular policy. This may resemble a poll, but the real objective is to change opinions rather than to measure them (Traugott and Kang 2000).

Computer-Assisted Survey Information Collection

It is useful to discuss computer-assisted survey modes in more detail before proceeding. Early survey research techniques were primarily paper-and-pencil (P&P) based, but computer-assisted survey information collection (CASIC)—a term coined by researchers for the federal government—is now common. Particularly important in this shift was the devising of procedures for computer-assisted telephone interviewing (CATI). CATI began in 1971 with Clifton Research Services (Couper and Nicholls 1998) and moved into the academic arena through the efforts of Merrill Shanks at the University of California, Berkeley. The federal government's statistical agencies began to adopt CATI by the early 1980s. The new personal computer technology permitted small companies to run their own CATI operations, leading to rapid expansion of the number of private survey research firms through the 1980s and 1990s.

CATI has several advantages over conventional paper-and-pencil techniques. It permits the interviewer to enter responses directly into a computer, eliminating data entry as a separate stage of the survey process. It controls the order in which questions are asked, facilitating the use of complex branching strategies in which the answer to one question determines the next question to be asked. Two additional innovations have

provided further advantages for CATI. One is the development of computer programs to control the assignment to interviewers of phone numbers to call, including keeping track of appointments when respondents ask to be phoned back for the interview at a particular time. Another is the ability to do random experiments, using a computer to randomize question order, the order of answer categories, and/or which versions of questions are asked of particular respondents.

Computerization has the further advantage of moving the data editing process (see chapter 11) to the interview itself. The computer can be programmed to catch implausible combinations of answers and have the interviewer correct one of the entries or ask the respondent to resolve the apparent inconsistency. Additionally, since the data are entered directly into a computer, CATI permits fast analysis of surveys, as epitomized by the reporting in the morning news of CATI surveys taken the previous night. Thus CASIC has increased the speed and efficiency with which surveys are conducted, along with the completeness and consistency of the data that are collected. Computer use does, however, add some complexity to the interviewer's task, requiring him or her to pay attention to the machine and not just the respondent.

Computerized surveys are particularly useful for panel surveys in which researchers want to ask respondents about changes from their previous interview. Relevant answers from previous waves can be preloaded into the computer, so the interviewer can tell when changes have occurred and can prompt the respondent's memory. Rather than asking what jobs the respondent has had since the last interview in September two years earlier, the interviewer can say, "You were working at the XYZ plant when we talked two years ago. Are you still there, or have you had other jobs since?"

The development of notebook computers permitted the development of computer-assisted personal interviewing (CAPI). CAPI was first implemented by government statistical agencies that were committed to face-to-face interviews by local field interviewers (Couper and Nicholls 1998). An extra advantage of CAPI is the ability to send survey data from a notebook through a modem to a centralized facility, as opposed to having to mail completed paper-and-pencil questionnaires back to headquarters.

The computerization of surveys has also affected noninterviewer surveys. Computerized self-administered questionnaires (CSAQ) have respondents go through a questionnaire themselves on a computer. This is done in e-mail surveys and Internet surveys, which require respondents to have access to both computers and the Internet. Knowledge

Networks, an Internet survey organization, even supplies Web TV devices and Internet connections to members of their sample in exchange for participation in their surveys. CSAQ can also be achieved in CAPI: the interviewer can hand a notebook computer to the respondents and have them go through some questions by themselves. This can be particularly useful for sensitive questions (such as about drug use) that the respondent may not want to answer in front of an interviewer. The invention of small personal digital assistants (PDAs) leads to further possibilities, with respondents being asked to click through the survey on a PDA. This technology has been suggested as a way to automate election exit polling, so that responses would be recorded automatically and could be uploaded quickly and accurately to a central computer that would do the tabulation and reporting.

Further innovations include interactive voice response (IVR), which includes both voice recognition entry and touch-tone telephone data entry. In the ideal version, people are first contacted by a human interviewer or through the mail. They are asked to phone a special number (or are transferred to that number if they are contacted by phone), and they enter a PIN so that access to the survey is controlled. Respondents are asked questions by an automated unit, and, depending on the system, they can respond either verbally or by pressing numbers on the phone keypad. Voice recognition software can analyze people's oral responses so that question flow can be regulated interactively according to the answers given to previous questions; however, it cannot yet handle complex answers. As an example of the importance of this system, the Gallup Organization conducted over one million IVR interviews in 1999 (Tourangeau, Steiger, and Wilson 2002). Dillman (2000, 402-12) is positive toward this technology, to the point of his predicting that it will mostly replace CAPI (353).

It should be emphasized that CASIC has disadvantages as well as advantages. In particular, good programming of complicated surveys is expensive and time consuming. It is easy for programmers to make logical errors in computerizing the questionnaire, leading to questions' being skipped or the wrong question's being asked. In addition to the expense of computers, there is the cost of a computer network and the personnel costs for maintaining the computer system and ensuring that it is always available for interviewing. It is essential that interviews be saved, even if the system crashes. All in all, the fixed cost of computerizing a survey is high, even if the variable cost of increasing the number of respondents is relatively low once it is computerized (Groves and Tortora 1998). As a

result, especially when few interviews are needed, researchers often do not employ CASIC capabilities, instead having interviewers read questionnaires to respondents.

Interviewer-assisted computerized surveys still differ from computer surveys that are not interviewer assisted. CATI and CAPI were intended originally to improve data quality, and they clearly have been successful for interviewing the mass public. On the other hand, CSAQ and IVR are generally intended to reduce costs and speed data collection, but they have not yet been shown to be as successful in surveys of the mass public (Nicholls, Baker, and Martin 1997).

Internet Polling

Internet polling is still a fairly new mode of computer-assisted data collection. Like many other aspects of the Internet, its potential is exciting, but it is too early to tell how successful it will be. Some approaches are effective, but others have not worked well. Putting a poll up on the Internet can be inexpensive, so many groups put up polls without paying enough attention to quality. However, quality research is never cheap. In the long run, quality standards will be established for Internet surveys. Indeed, there are already several interesting attempts to create high-quality Internet surveys. This section will briefly mention some poor approaches for Internet polls, after which some of the better approaches will be described.

Problematic Internet Surveys

The worst type of Internet polls has already been mentioned: click-in polls. Some popular Web portals conduct daily polls. Respondents to click-in polls are volunteers, and there is no reason to think that they are a representative sample of a broader population. Interest groups can alert their membership when a click-in poll topic is of importance to them, in an attempt to boost the desired numbers. Further, a person can respond repeatedly on these polls.³ Having said that, even those types of Internet polls have their defenders, who ask whether telephone polls with 50% response rates are really any more scientific than click-in polls. The issues involved in click-in polls will be considered repeatedly throughout this book.

There are several Internet sites that connect people with Web surveys that pay for participation. This leads to volunteer samples of respondents who are more concerned with earning money for participation than with giving honest answers. While it is increasingly common to provide incentives for participants in all types of surveys, the sample's representative-

ness becomes dubious when many respondents are people who frequent pay sites regularly.

Some survey sites allow free posting of surveys, with the Web sites getting their funding from ads shown to respondents. However, these sites have no incentive to facilitate high-quality polls. Surveys posted on such sites are usually static questionnaires that ask respondents all the questions, even if their previous answers show that some of the questions are not appropriate. The respondent is usually expected to skip those questions, though skip instructions are not always clear. If the respondent answered some questions that should have been skipped, the software may still include those responses in tallies of answers.

An inexpensive form of Internet survey is the e-mail survey, in which the questionnaire is simply embedded in an e-mail message. The respondent has to hit the reply button, put X's inside brackets when choosing answers, and then hit the send button. However, it is very easy to get rid of these surveys by hitting the delete button. Yet e-mail is still one of the simplest ways to get surveys to computer-literate respondents.

There is an unfortunate tendency to ignore the usual ethical guidelines when posting surveys on the Web. It is so easy to put up surveys on the Web that appropriate ethical guidelines are often not noticed.

Better Internet Surveys

As an example of a reasonable procedure for Internet surveys, some universities now allow their faculty and students to fill out internal surveys on the Web. Respondents can be notified by e-mail that they have been randomly chosen to participate, with an indication that they can answer the survey on the Web if they prefer that to a phone interview. They are given the Web address for the survey along with a unique password, so that only designated respondents can participate. Using the Web version of the survey can be a popular alternative because university personnel and students are already very familiar with computer use and since many would prefer doing the survey at off hours rather than when an interviewer happens to call them. There can even be incentives to participate online, such as a chance to win a free campus parking permit. The main disadvantage is that people often decide to take the Web version later but then forget to do so, so reminders must be sent after a week, and phone interviewers then must call people who still forget to do the survey.

There are two important innovative national Internet surveys that will be referred to in several later chapters. One is Harris Interactive, which obtains a pool of respondents by sending out large numbers of invitations

offering a chance to win monetary prizes in exchange for participation in polls. Clicking on the ad leads to the Harris Interactive Web site and a short survey that obtains basic demographic data on the person, along with assurances of confidentiality. The person's e-mail address is put into Harris's database of panel members, and those people are then sent frequent notices about particular surveys. The e-mails have subject headings that indicate a chance to win a large prize, such as "Win \$6,000: Evaluate Ecommerce sites" and "Win \$4,000: Home Purchases Survey." Additionally, participants earn points that can be exchanged for merchandise. A good prize requires about 6,000 points, with each survey participation being worth 100 points and panel members being given the opportunity to participate in one or two surveys each month. The e-mail invitations provide the Web address of the survey along with a unique password that gives access to the survey. The known demographics of participants in each survey can then be compared to known population parameters to weight these Internet surveys appropriately (chapter 8). Reminders are sent out in a few days to panel members who have not yet participated in the survey. There are not yet definitive independent studies of the validity of this approach, but it is a highly visible attempt to produce a quality Internet survey.

Knowledge Networks is also doing innovative Internet surveys (Lewis 2000). It obtains a pool of respondents by conducting a random national telephone survey in which respondents are offered a Web TV receiver and free Internet connectivity in exchange for participating in its Internet surveys. This approach minimizes the coverage problem of many homes' not having computers or Internet access by providing both to respondents. Therefore its sample does not have the middle-class bias that an Internet survey might otherwise have. A free Web TV and Internet access are an attractive incentive for many people, but there are still some potential coverage issues, such as the unwillingness of some older people to learn how to use a computer. Once the Web TV device is installed in the home, respondents answer profile surveys about the demographics of each household member. Each panel member is expected to answer one short survey per week; the Web TV receiver is removed from the home if respondents skip eight surveys in a row. Participants are rewarded with incentives, including cash, gift certificates, movie passes, airline bonus miles, and phone cards. Forty percent of the Knowledge Networks phone sample were contacted for the initial telephone interviews, participated in that survey, agreed to join the panel, and actually had the Web TV receiver installed (Krosnick and Chang 2001, table 1, note a). Again, there are not

yet definitive judgments of the validity of this approach, but it is a serious attempt to harness the Internet for quality survey research.

Some of the most innovative uses of Internet surveys involve multimedia. In particular, survey participants who are on broadband networks can be sent streamed video. For example, they could watch a presidential speech as it occurs and continually record their reactions to it, and the television networks can relay the compiled response to viewers by the end of the speech. Yet this approach is controversial, as the participants are essentially volunteers and may not be representative of people who were not at home during the speech; further, the sample is biased to Internet users who can afford broadband access.

Responsible Internet surveys may eventually be a major mode of survey research, with the potential of supplanting telephone surveys. The ability to do inexpensive surveys on the Web, however, will continue to encourage low-quality surveys as well. In many ways, the status of Web surveys in the years between 1998 and 2005 resembles the status of telephone surveys in the mid-1970s, when procedures for quality phone surveys were still being tested and perfected.

Mode Limitations

The myth of the "ideal survey mode" is that there can be an ideal survey mode that is always best. In practice, each survey mode has its limitations. Face-to-face surveys over large geographical areas have complicated personnel and organizational needs that lead to large expenditures and a long time to complete the interviewing. Mail questionnaires are much less expensive but have always been plagued by low response rates, which in turn necessitate keeping the study going for several weeks as attempts are made to increase the response rate. CATI surveys are less expensive than face-to-face interviewing because they save transportation costs and can be faster to complete, but in fact they can sometimes be more expensive because of high programming costs; moreover, response rates are falling rapidly on telephone surveys. Internet surveys can be less expensive (except again for programming costs) and quick, but they have real coverage issues. Unfortunately, the mode with the highest response rate (face-to-face interviewing) is also the most expensive and time consuming, whereas the least expensive modes (mail questionnaires and Internet polls) have the lowest response rates and the least satisfactory coverage.

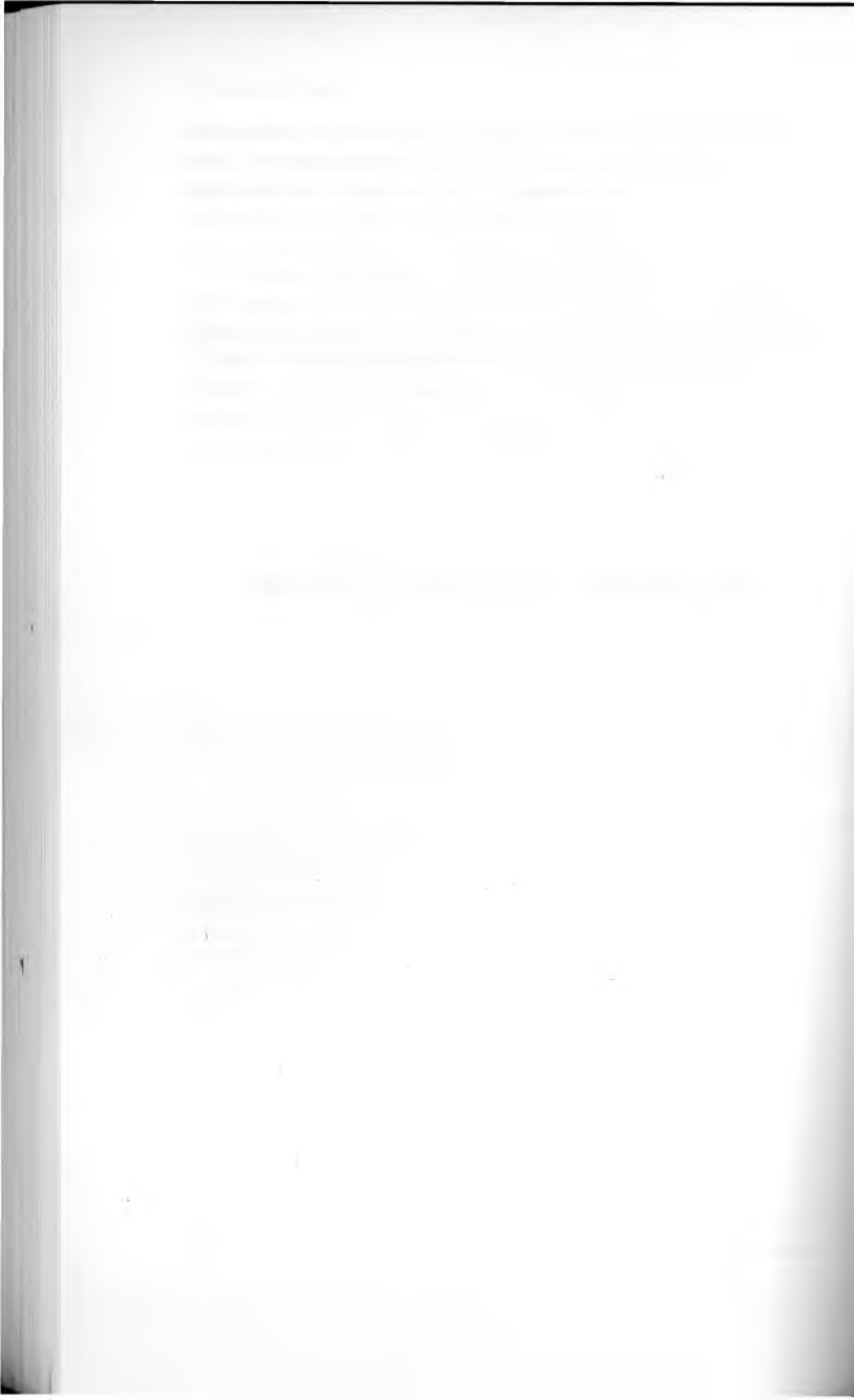
These trade-offs suggest that a multimode approach can sometimes be useful, and more studies are employing multiple modes. Fortunately, most

large survey organizations can conduct surveys using more than one mode. The organizations that do face-to-face interviewing generally also have telephone capabilities too, and phone survey organizations usually are also able to send out mail questionnaires. Similarly, many long-standing survey operations have decided to experiment with Internet surveys. A 1995 survey of a sample of U.S. survey organizations found that all used some mixture of modes, with nearly all using at least one CASIC mode (Groves and Tortora 1998), making true multimode surveys possible.

These various survey modes will be discussed in more detail in later chapters, along with further statements of their advantages and disadvantages. After several types of errors are discussed in the next eight chapters, chapter 12 will explicitly examine the effects of the different modes.

PART II

Response Accuracy Issues



4

Measurement Error Due to Interviewers

THE DEBATE OVER INTERVIEWING STYLE

We begin with measurement error—the error that occurs when the researcher does not obtain accurate measures of the phenomena of interest. Measurement error can be due either to interviewers or to respondents. Interviewer-related error occurs to the extent that responses are different because of the interviewer; the statistical effects of different types of interviewer error will be described in this chapter. Interviewer error is a potential problem in both face-to-face and telephone surveys, but self-administered questionnaires, mail surveys, and Internet surveys are immune to interviewer effects.

Beginning with interviewer error does not mean that interviewers are a large source of error in surveys. Skilled, well-trained interviewers obtain high-quality data. However, interviewers are often not given enough training and supervision, so error due to interviewer effects does occur. Thus, considering interviewer-related error in this chapter provides an opportunity to discuss interviewer management as well as the role of the interviewer more generally. This chapter begins with a discussion of different interviewer styles, followed by treatment of interviewer effects and then interviewer management and supervision.

The Theory of the Interviewer

Understanding interviewer errors requires understanding what an interview is and how to conceive of the role of the interviewer. A lively debate has developed in recent years over the proper interviewing style, with

supporters of the conventional standardized interviewing approach being challenged by advocates of more flexible conversational interviewing.

Two Early Approaches

The earliest view was that the interview is intended to obtain honest reporting of facts and attitudes and that this is facilitated when the interviewer develops a good personal relationship with the respondent. The underlying philosophy was based on psychotherapy, more specifically an analogy with a psychological therapist and a patient, with the patient opening up to the therapist better if they first develop a good personal relationship. The term that is usually used to describe this relationship is *rappport*. A warm professional relationship is necessary, with the interviewer being respected and trusted and acting both accepting and nonjudgmental. Indeed, for many respondents, the interviewer is the reason the respondent is doing the interview. The interviewer does not have to become their close friend, but the interviewer has to develop an easy question style so respondents will feel that they can give candid answers.¹ This view of the interviewer was developed in the days when interviewing was exclusively face to face; such rapport is harder to achieve in telephone interviewing. Rapport became controversial when some said it led to overly personalizing the interviews. Still today, however, many of the best interviewers are those who are able to develop rapport with respondents and use that to conduct high-quality interviews.

A *motivational* approach to interviewing was advanced in a series of research studies by Charles Cannell and his associates (e.g., Cannell, Oksenberg, and Converse 1977). Their emphasis was in motivating respondents to answer accurately through building a sense of commitment. The underlying logic is that being a survey respondent is a rare activity, so respondents do not know what is expected of them, and that it is possible to help them understand how they should behave in this setting. They found four specific techniques useful in training respondents to perform as desired.

First is reinforcement of constructive behavior by the respondent. Cannell's previous research (Marquis, Cannell, and Laurent 1972) had found that interviewers were giving positive feedback to respondents whenever they answered questions, even when their answers were not responsive. Indeed, interviewers sometimes gave the most positive feedback in that situation, accidentally rewarding behavior that was not desired. Instead, the interviewer should train the respondent by reinforcing answering behavior with positive comments only when the respondent answers in a thorough

and thoughtful manner. A second technique they emphasize is having interviewers stress at the beginning of the interview how important it is that answers be complete and accurate. Third, their research suggests that it is useful to have the respondent promise to try to answer the questions completely and accurately. Fourth, they suggest a technique called *modeling*: giving respondents an example of how they are supposed to behave in answering questions, such as first playing a tape recording of a respondent who is concerned about giving complete and accurate answers. Obviously, these techniques cannot all be used in every type of interview; modeling, for example, is not feasible in a phone interview, but it is possible when a roomful of people is instructed in how to answer a survey.

The Standardized Approach

Cannell's motivational approach is considered insightful, but the dominant view of the role of the interviewer today is that the interviewer should ask questions in a *standardized* manner. The standardized approach to interviewing the mass public emphasizes the importance of the interviewer's establishing a professional relationship with respondents, rather than developing comfortable rapport with them. The underlying principle is that every respondent should be asked the identical questions in the identical manner and that the interviewer's role is to do just that.²

The interviewer in standardized interviewing is basically a "relayer of questions" (Houtkoop-Steenstra 2000, 45). Lavrakas (1993, 132) likens the interviewer to an "intelligent automaton," and that description aptly conveys the view of the interviewer's role from the standpoint of standardized interviewing. Indeed, the emphasis often is more on the interviewer as automaton than as intelligent. For example, in standardized interviewing an interviewer is supposed to repeat all answer categories to a respondent whenever it is not clear exactly which category the respondent has chosen. Say approval of something is being measured: the interviewer must repeat, "Do you strongly approve, somewhat approve, somewhat disapprove, or strongly disapprove?" even if it is clear that the respondent is just deciding between "strongly approve" and "somewhat approve."

Rules for Standardized Interviewing

1. The interviewer should read the questions exactly as written.
2. Probing of incomplete answers should be neutral.
3. Answers should be recorded verbatim.
4. Interviewers should be nonjudgmental about answers.

Having interviewers follow a script exactly is particularly important for survey operations that hire hundreds of telephone interviewers but want to minimize training costs. Too often, however, a respondent can recognize when the interviewer is merely reading from a script, especially when the script is read poorly with words mispronounced and the interviewer is seemingly lost whenever the respondent asks about the meaning of a question. Respondents are less likely to take interviews seriously when the interviewer does not seem fully competent. Obviously this is not a problem with standardized interviewing *per se*, but it emphasizes the challenge that standardized interviewing operations face in making sure that their interviewers sound professional.

Several procedures have been developed for conducting standardized interviewing. The first is to establish the proper context with the respondent. This includes explaining the purpose of the interview, setting the proper tone, and communicating the goal clearly (Fowler and Mangione 1990, chap. 4). The interviewer should be professional, while encouraging a free flow of communication in order to obtain high-quality answers. This approach is based on Kelman's (1953) work on three ways to achieve influence relationships: compliance (through reward or punishment), internalization (linking the desired behavior with the person's own values), and identification (behaving consistently with role expectations). Achieving the interview through a compliance relationship (as when a boss orders a subordinate to answer the questions in a survey) is thought to lead to poor interviews if respondents do not see any reason to give good answers. Internalization can be more effective for producing accurate answers if the interviewer can establish a link between the survey and the respondent's own values and goals. The identification route requires respondents to learn a role in which good answers are provided, as expectations of what good answers should be like are supplied—such as answering closed-ended questions with one of the offered choices.

Maynard and Schaeffer (2002a) refer to survey interviews as “messy interactions,” though hopefully good enough to obtain the needed information within available cost constraints. The “messy interactions” phrase usefully points out that interviewing is never perfect; standardization is often stated as an ideal, but it is often not achieved. A study of twelve university survey centers that purported to use the standardized approach found that only one consistently followed standardized interviewing procedures (Viterna and Maynard 2002).

One instance of the problem with standardized interviewing is what Houtkoop-Steenstra (2000, chap. 7) describes as the tendency of inter-

viewers to behave like "ordinary conversationalists" by "normalizing" awkward situations. Her example involves a literacy survey in which respondents were to talk about their reading incompetence. In order to maintain rapport with respondents, the interviewers often rephrased questions, praised indications of progress in reading, made comments about how the respondents' literacy problems were normal, and sometimes indicated that they shared some of the respondents' problems. This may be an extreme case, but it shows how interviewers will often help respondents save face, even when the interview is supposed to be standardized. The general point is that interviewers tend to work through awkward interview situations, finding ways to make them less awkward, even if this departs from standardized interviewing. This situation may not be desirable, but it emphasizes the need for researchers to take such potential problems into account in writing their questions so there is no need for interviewers to do this.

The Conversational Approach

Standardized interviewing is now the dominant approach, but some researchers instead favor *conversational interviewing* (or *flexible interviewing*). Advocates of this approach argue that there will always be ambiguities in question wording, so it should be the job of the interviewer to elicit answers that will be as close as possible to what the researcher is really trying to measure. The goal is that all respondents understand the questions in the same way, but the inevitable differences between respondents suggest that interviewers will have to deal with different respondents differently to achieve that standardized understanding. The result is to give the interviewer a more active role than under standardized interviewing.

The conversational approach harks back to an early description of an interview as a "conversation with a purpose" (Bingham and Moore 1924). Conversational interviewing emphasizes the need to work with the respondent to accomplish that purpose. Suchman and Jordan (1990) particularly argue that standardization prevents interviewers from handling unpredictable situations and that more valid results are obtained when the *meaning* of questions is standardized rather than their *wording*.

The differences between standardized and conversational interviewing styles are most apparent when respondents indicate that they do not know how to interpret a question. In a standardized approach, the interviewer explains that respondents should interpret the questions themselves, whereas in a conversational approach the interviewer is supposed to collaborate with respondents in achieving the intended meaning of the

question. Houtkoop-Steenstra (2000, 61) usefully describes this as the interviewer acting as a "spokesperson for the researcher," obtaining the information the researcher really wants, even if that means departing from standardized wording when the respondent misunderstands the question. As Maynard and Schaeffer (2002a) argue, interviewers must employ their "tacit knowledge" in implementing rules for interviewing. Commonsense understandings are achieved between the interviewer and the respondent, understandings that go beyond pure standardization. The notion of tacit knowledge incorporates the view that interviewing is partly art, depending on the capabilities of the interviewer.

Conrad and Schober (2000) report an experiment in a fact-based national telephone survey involving purchasing behavior, where respondents were interviewed twice, with the interviews taken a week apart. Respondents were about their housing and household purchases, using questions from the federal government's Consumer Price Index-Housing Survey and Current Point of Purchase Survey. The first interview was always standardized, whereas the second was standardized for a random half of the sample and conversational for the other half. Conrad and Schober looked to see how answers changed between these interviews and which best fit the official definitions for terms in the study, such as that buying a floor lamp does not constitute buying household furniture. Respondents were twice as likely to change their answers from the first interview when the conversational approach was used for the reinterview (22% of answers changing versus 11% when the second interview was also standardized), and the changed responses were more likely to fit the intended definitions when conversational interviewing was used.³ For example, 57% of answers about purchases were in accord with official definitions in the standardized interviews, versus 95% in the conversational. Conrad and Schober interpret this evidence as demonstrating that conversational interviewing produces more valid answers, at least on factual material. As should be expected, the added cost of this approach was in terms of longer interviews (from a median of 5 minutes to a median of 9), due both to explaining the conversational procedure to respondents at the beginning of the interview and to the time required to help respondents understand what was being asked in each specific question.

The Choice of Interviewing Approach

The advocates of standardized interviewing would strongly attack the conversational approach as leading to data of questionable validity. They would argue that interviewer effects would be larger with this approach,

since interviewers would likely each develop their own styles which would affect all interviews they conduct. The data from conversational interviewing would not be comparable between respondents because they are given different stimuli. As a result, those who favor standardized interviewing would contend that greater question clarity is what is needed, combined with complete standardization; better question wording would make conversational interviewing unnecessary.

In this vein, Houtkoop-Steenstra (2000, chap. 1) contends that "interviewer effects" may really be "questionnaire effects," so that, by extension, interviewer errors may really be questionnaire construction errors. The implication is that researchers should take the nature of interviews as conversations into account when designing their questions. Interviewers depart from standardized interviewing practices when the questionnaire does not make adequate allowance for the reality of interviewing. Thus, flexible interviewing might not be necessary if the questions were formulated so that interviewers and respondents could feel more comfortable with a fully standardized approach.

The standardized and conversational interviewing styles play through differently for fact-based and attitude surveys. In fact-based surveys, it is important that respondents correctly understand what facts they are supposed to provide. Take the question "How many credit cards do you have?" The question seems simple, but respondents may not fully understand what to include in their answer. Should they include debit cards in the count? What about cards that their spouse has, or that they and their spouse have together? Should store cards be included or only bank cards? What about cards they have but never use? Wording the question more carefully could address some of these concerns: "Not counting debit cards that you use only for withdrawals, how many bank credit cards did you yourself use last month?"

Survey organizations differ in the instructions they give to interviewers as to how to handle questions from respondents on the meaning of fact-based questions. Two intermediate approaches can be delineated between the standardized approach of just repeating the question and flexibly working with respondents to develop a proper understanding of the question. A semistandardized approach would be to give interviewers the exact replies they are supposed to give to particular questions that respondents might ask, such as whether to include credit cards used jointly with a spouse. A semiconversational approach would be to allow interviewers to use their best judgment in responding to the respondent's questions but not to initiate clarifications themselves.

It is more awkward for the interviewer to give clarifying information to respondents in an attitude survey than in a fact-based survey. Take the question "In politics today, do you generally consider yourself a liberal, a conservative, or a moderate?" If the respondent asks what those terms mean, it would be hard for an interviewer to respond in a nonbiased manner. A question like this one is usually intended to find out the respondent's ideology, based on how the respondent interprets the terms. There are no standard meanings to the terms, and spontaneous efforts by the interviewer to provide meanings would only throw off the results.

Interviewing approaches also differ in their *probing* styles. Respondents often give answers that are incomplete, unclear, or not fully responsive. Interviewers should be trained to encourage the respondents to improve their answers, without making respondents feel that their answers are inadequate. The best probes are neutral and nondirective. Typical nondirective probes for open-ended questions would be "How do you mean that?" "Tell me more about that," or "Anything else?" Or the interviewer might repeat the original question or simply remain silent. Directive probes that can be answered "yes" or "no" are particularly frowned upon, given that respondents are more likely to answer "yes." With closed-ended questions—items that give respondents a limited set of response options—the proper standardized response to a request for clarification is to repeat all the answer choices, whereas conversational interviewing would permit the interviewer to help the respondent understand the choices. Some advocates of standardized interviewing would argue that probing detracts from standardization (Foddy 1995), since some respondents are being given different stimuli than others.

The two approaches also differ in handling questions that ask for material that has already been provided in the interview. For example, what if the respondent has already indicated his or her birthplace before it was asked? Standardization normally requires that respondents be asked a question even when they have happened to answer it in response to an earlier question. The Census Bureau and some other survey organizations instead permit the interviewer to ask the respondent to confirm their earlier answer, a process known as *verification*. For example, the interviewer might be permitted to ask, "I think you said earlier that you were born in Minnesota; is that correct?" and go on to the next question after a short pause unless the respondent corrects that information. This departs from standardized interviewing, but it is a way of showing the respondent that the interviewer is paying attention to the conversation rather than being an "incompetent conversationalist" (Houtkoop-Steenstra 2000, 85) who

does not listen to what has been said. The conversational interviewer might be allowed to skip the question entirely and just record the previously offered information.

The success of the conversational approach depends on high-quality interviewers who have undergone considerable training on the questionnaire so they understand the information that each question is meant to elicit. By contrast, standardized interviewing is the only option for market opinion research firms that hire large numbers of inexperienced interviewers and give them minimal training.

The debate between advocates of standardized and conversational interviewing is presented in an edited volume (Maynard et al. 2002). That volume also nicely illustrates the use of conversational analysis (Houtkoop-Steenstra 2000) and ethnomethodology in analyzing taped interactions between interviewers and respondents. The standardized approach remains dominant today, but conversational interviewing has attracted interest as researchers realize that interviewers often depart from fully standardized interviewing.

Interviewer Effects

The use of interviewers affects answers given by respondents in several ways (Groves 1989, chap. 8). Interviewers can administer the questionnaire in different ways, such as by rewording a question rather than asking it as written. They can emphasize different words or use different intonations in reading the question. They can probe differently with different respondents. Finally, interviewer demographic characteristics can affect the answers that are given, such as if people give more candid answers to interviewers of their own social group.

Interviewer Errors

Interviewers inevitably make some actual errors. Some are mechanical slips, such as the *skip errors* in pencil-and-paper surveys that occur when interviewers accidentally skip questions or do not record answers. One study (Sebestik et al. 1988) found that 90% of interviewer mistakes in pencil-and-paper surveys involved not recording answers. Skip errors are particularly common when the interviewer is supposed to skip some questions because of an answer to a previous question, such as skipping questions about experiences in the military for respondents who have never served. This problem is usually due to the layout of the written questionnaire, with unclear skip instructions making it difficult for interviewers to

follow the question order correctly. Analyzing Detroit Area Studies surveys, Sanchez (1992) finds that error in skipping questions is due to the design format of the questionnaire and is not diminished by interviewer experience. Computer-assisted interviewing avoids such errors by making sure that the interviewer records the answer to a question before going on to the next question.

Another mechanical source of interviewer error involves recording the answer incorrectly. *Recording errors* do occur, but their rate is usually thought to be fairly low. Lepkowski, Sadosky, and Weiss (1998) construct a model in which recording accuracy is a function of question characteristics, respondent behavior, interviewer behavior, and the conversational dynamics of the interview. They checked the responses on 200 interviews that they had tape recorded. In simple exchanges (such as closed-ended questions with "yes" and "no" answers), they found that only 0.1% of the answers were recorded inaccurately, with the same rate for CATI and non-CATI interviews. Most of the inaccuracies involved recording "don't know" when the person gave a "yes" or "no" answer to a question, but the researchers deemed these to be random errors. In more complex interviewing exchanges, the error rate was somewhat higher for non-CATI phone interviews than for the CATI interviews, 6.3% versus 4.5%, with CATI interviewers being less likely to not record an answer or to record the wrong answer. They also found that recording errors were mainly associated with the behavior of the respondents, particularly when they gave an answer outside of the actual response choices, with fewer recording errors being associated with the behavior of the interviewers. Recording errors involving entering a different response from the one the person actually gave are usually random errors. However, sometimes interviewers choose to record a different answer that allows them to skip follow-up questions and therefore finish the interview more quickly; this would be considered interviewer falsification and research misconduct, as discussed later in this chapter.

The more serious type of interviewer error has to do with how questions are asked. Interviewers are supposed to ask questions as written, but studies often find that a large proportion of questions are not asked as written. Studies counting the extent to which interviewers change question wording obtain estimates ranging from 20–40% of the time (Fowler and Mangione 1990, 34) to 30–95% (Groves 1989, 383). Interviewer training and experience are not highly predictive of whether questions are read exactly, but interviewers vary considerably. The four best interviewers in one study delivered 11% of questions with minor changes, whereas

the four worst interviewers averaged 67% (Oksenberg 1981). These changes are usually random and therefore do not usually bias results. A later study of 79 interviews at the University of Maryland Survey Research Center in 1989 (Presser and Zhao 1992) found that 91% of the questions were asked exactly as worded, and 8% involved only minor alterations, but error was greater for longer questions and for questions asked in series. Interviewer experience was again, surprisingly, not found to make a difference. Whereas most accounts emphasize that rewording questions can affect the results, Groves (1989, 387) maintains that rewording of questions by interviewers does not necessarily increase measurement error, since answers to questions often do not vary with minor changes in question wording.

The main interviewer-related recording errors found by Lepkowski, Sadosky, and Weiss (1998) occurred when incorrect probing techniques were used and, in particular, when the interviewer did not probe. This emphasizes the importance of training interviewers on desired probing techniques.

Statistical Consequences of Interviewer Effects

As described in chapter 3, a key distinction is between uncorrelated and correlated measurement error. Uncorrelated error occurs when an interviewer happens to ask questions somewhat differently from interview to interview in a random fashion, which increases the variance of those items and therefore decreases their reliabilities. Correlated error occurs if each interviewer asks questions in an idiosyncratic manner across all of her interviews, and this has more serious statistical consequences.

The *intraclass correlation*, ρ (ρ), is commonly used as a measure of the variance associated with interviewers (Kish 1965).⁴ Computing ρ is equivalent to computing an analysis of variance in which the treatment is the interviewer, thus obtaining a measure of how much of the variance is due to the interviewers. ρ has a maximum value of 1.0, with lower values indicating that less of the variance is associated with the interviewers. This estimation is accurate so long as the assignment of respondents to interviewers is random (known as an *interpenetrated design*), so there is no reason to expect that respondents assigned to different interviewers should vary.

ρ is used as a part of a multiplier, inflating the variance of a variable. Even a small ρ can lead to a large multiplier. The statistic used here is called a *design effect* (or "deff"), and it equals $1 + (\rho(\bar{n} - 1))$, where \bar{n} is the

average number of completed interviews per interviewer. The design effect is evaluated in terms of the variance, so the appropriate multiplier for the standard error is its square root, commonly referred to as a "deft": $\sqrt{1 + (\rho(\bar{n} - 1))}$. The deft shows how much the standard error of a variable is increased by correlated measurement error.

As an example of the insidious nature of correlated measurement error, say that the rho is a seemingly trivial 0.01. If the average number of interviews conducted by an interviewer were 34, the design effect would be 1.33, increasing the variance of a variable by a third. The corresponding deft is 1.15, a 15% increase in its standard error, which makes it harder for results to appear statistically significant. The rho will vary between questions in a survey, and obviously this is a problem only on questions for which there are interviewer effects. There is no reason, for example, to expect an increase in the standard error on determining the respondent's gender, but there is likely to be an increase on attitudinal questions.

Consider a situation in which a survey organization is deciding whether to assign an average of 20 or 30 interviews per interviewer. Let's still assume a small intraclass correlation, 0.01, to see what damage even a small rho value can do. The deft with 30 interviews would be 1.136, whereas it would be 1.091 with 20 interviews, which amounts to a third less inflation of the standard error. If 300 interviews were to be taken, having interviewers take only 20 apiece would require a staff of 15 interviewers, rather than just the 10 that would be needed if each interviewer were assigned 30 interviews. Hiring 50% more interviewers (15 rather than 10), though, would cut the inflation of standard errors on variables susceptible to this effect by one-third, though the actual differences in the two defts is small enough to not make a big difference. Thus when survey organizations decide how many interviewers to hire for a project, they are implicitly deciding on the trade-off they can accept between the number of interviewers to train and the amount that some standard errors will be inflated. As Biemer and Lyberg point out, there is a general "lack of information in the research community regarding the damaging effects of [correlated] interviewer error" (2003, 168); interviewer load has direct implications for survey error that should be taken into account.

Varying Interviewer Effects

Interviewer effects can vary across modes of interviewing, across types of questions, between types of respondents, and because of interviewer

characteristics. This section will summarize the differences that have been reported along these four different dimensions.

As to *mode differences*, measurement error due to interviewers has been found to be greater with face-to-face interviews than phone surveys. Groves (1989, chap. 8) reports that personal interviews from the University of Michigan's Survey Research Center have an average rho of 0.031. Actually most of the questions have ρ 's below 0.02, but the ρ values were higher on sensitive topics. He also reports an average 0.01 value for phone surveys, with the highest ρ being 0.0184, but even the 0.01 rho for phone surveys means an increase in 10% in the variance for every 10 interviews added to an interviewer's workload.

Using a survey mode that does not use interviewers will, by definition, completely eliminate interviewer error. Indeed, the Census Bureau's concern over the effect of census enumerators on the 1950 census was one factor that led to its adopting self-enumeration in subsequent censuses, since the measurement error between households is uncorrelated when people fill out the census forms themselves (Hansen, Hurwitz and Bershad 1961). Doing away with interviewers is a drastic method for avoiding measurement error due to interviewers, but this consideration does explain why some researchers prefer self-administered surveys (as through the Internet) rather than telephone surveys.

The evidence is mixed on whether interviewer effects are greater on some *types of questions*. Fowler's review of this literature concludes that "on average, there is very little evidence that either the content or the form of the question *per se* is a good predictor of where [i.e., the types of questions on which] interviewer-related error is most likely to be found" (1991, 275). Several studies find smaller interviewer effects on factual items than on attitude questions, but in nine University of Michigan Survey Research Center studies the interviewer effects were not different on the two types of questions (Groves 1989, 375). Also in these studies, open-ended questions that respondents answer in their own words were not subject to larger interviewer effects than closed-ended questions (respondents choose among response options that are part of the question). The effects that were found have to do with giving a second mention on open-ended questions, supporting Fowler and Mangione (1990) on failure to probe open-ended questions.⁵

Interviewer effects could also vary between *types of respondents*. Groves (1989, 379-80) expected poorly educated respondents to be more affected by the behavior of the interviewer than highly educated respondents, but the rho he obtained was not larger for less educated respondents. He also

expected older respondents to be more influenced by the interviewer. The Michigan Survey Research Center data do show a higher rho for older people. However, that measures variation in nonresponse as well as response errors over interviewers, so the higher rho could just reflect higher non-response variation among elderly respondents.

Interviewer characteristics have generally not been found to affect data quality (Fowler and Mangione 1990, chap. 6). Because there are not consistent correlates of interviewer quality, interviewer-related error cannot be minimized by selecting interviewers on the basis of demographic or psychological factors (Hox, de Leeuw and Kreft 1991). The common exception has to do with younger interviewers' producing poorer quality data, but that is probably due simply to insufficient training of student interviewers. Fowler and Mangione (1990) conclude that interviewers do not need specialized knowledge of the topic to get quality data; indeed Fowler (2002, 122) argues that interviewers with such knowledge are not desirable, since they are likely to read more into unclear answers instead of probing.

An extreme example of respondents' tailoring their answers to the apparent sympathies of interviewers comes from the famous "pen experiment" conducted by Howard Schuman (1990) in an attempt to explain the very disparate results of polls conducted before an election in Nicaragua at the time of that country's civil war (Miller 1991). Interviewers randomly used pens showing the leftist Sandinista government logo, the opposition logo, or a neutral design. The results confirmed that clues about the political leaning of the pollster can affect responses, at least in civil war conditions.

Still, some differences related to interviewers have been found, often when interview characteristics relate to questions in the survey. For example, respondents are less likely to express anti-Jewish feelings if the interviewer is identifiable as being Jewish. Interviewer effects are most likely, of course, when respondents know (or think they know) the demographics of the interviewer. Thus, gender effects and race effects are both likely in face-to-face interviews; race may matter less in phone interviews if respondents do not know the interviewer's race.

There is evidence of some gender effects, but the findings are sometimes inconsistent. Female respondents gave more feminist answers to male interviewers than to female interviewers on women's rights and roles, whereas male respondents gave more conservative opinions to male interviewers (Groves 1989, 401).⁶ In a consumer attitudes survey, respondents were more optimistic when the interviewer was male, regard-

less of respondent gender (Groves and Fultz 1985). Female interviewers are more likely to be rated excellent, friendly, and professional (Fowler and Mangione 1990, 104). Huddy et al. (1997) found differences of about 5% in answers on a wide variety of gender-related survey questions administered by female and male interviewers. Kane and Macaulay (1993) report that people are more likely to give egalitarian gender-related attitudes and more criticism of gender inequalities to female interviewers, though interactions between interviewer and respondent gender tend not to be significant in multivariate models. Catania et al. (1996) find that respondents (especially men) indicate higher rates—and presumably more truthful responses—on some sexual behavior questions (such as same-gender sex, extramarital sex, and sexual problems) to interviewers of their gender. Yet the evidence on effects of matching the education and social-economic status of respondents and interviewers is not conclusive, and there is no evidence of age-of-interviewer effects.

The most interesting research on interviewer characteristics has to do with race. Starting with Hyman's (1954) early research, several studies have found more effects of race of interviewer on racial attitudes than on other questions. Respondents have been found to be less likely to express feelings against the interviewer's race. The differences are usually interpreted to show that respondents are more open with interviewers of their own race. Most of the research involves face-to-face interviewing, but some similar effects have been found with perceived race of interviewer in telephone surveys.

Table 4.1 shows differences in responses by whites to racial attitude questions, with a 46% difference in response on attitudes on interracial marriage (Hatchett and Schuman 1975–76). This example is old, but it shows that the race of the interviewer can have a strong effect on face-to-face interviews. There are also social desirability effects due to respondents' trying to save face with the interviewer, especially in face-to-face interviews, with blacks overreporting voting turnout to black interviewers (Anderson, Silver, and Abramson 1988a) and those living in predom-

Table 4.1 Support of Whites for Integration by Race of Interviewer, Adapted from Hatchett and Schuman (1975–76)

	Black Interviewers	White Interviewers	Difference
Interracial marriage	72%	26%	46%
School integration	91%	56%	35%
Residential integration	100%	69%	31%

inantly black areas who were interviewed by blacks in a preelection survey actually being more likely to vote.

Table 4.2 reproduces a classic result from Schuman and Converse (1971), showing that more than 2% of the variance in answers by black respondents in Detroit was accounted for by race of interviewer on 32% of racial attitude questions, whereas that much of the variance was accounted for by race of interviewer on only 3% of nonracial attitude questions. Differences of 13–28% were obtained on questions relating to militant protest and hostility toward whites. Although these differences look large, it turns out that it may partly reflect other differences between the black and white interviewers. The white interviewers in this study were graduate students doing interviews for the first time, whereas the black interviewers were professional interviewers (Groves 1989, 399–400). Schuman and Converse state their belief that the same race-of-interviewer result would have held in the year of the study (1968, after Martin Luther King's assassination) regardless of interviewer training. Still, this example illustrates the importance of looking for alternative explanations of effects that seem to be due to characteristics of interviewers and testing, to make sure that effects are not due to differences in how interviewers are selected.

As another example, Anderson, Silver, and Abramson (1988b) find that African Americans interviewed by whites in face-to-face National Election Study surveys were more likely to give responses indicating warmth toward whites than were African Americans interviewed by African Americans. Also, they report that blacks were more likely to say to a black interviewer that civil rights leaders were pushing too slowly and were more likely to favor increased spending on blacks, welfare, and food stamps; there were not significant differences on several other issues. Caution is appropriate in evaluating this finding, for respondents were probably not assigned randomly to interviewers, leading to the likelihood that black interviewers were more likely to be interviewing in areas in which more blacks lived. Thus some of the differences could be due to neighborhood effects rather than interviewer effects.

Table 4.2 Percentage of Questions with 2% of Their Variance Explained by Race of Interviewer, by Type of Question, in Schuman and Converse (1971) Study (with Number of Questions of Each Type Shown in Parentheses)

Type of Question	Racial Topic	Nonracial Topic
Attitude question	32% (40 Qs)	3% (29 Qs)
Fact question	14% (14 Qs)	8% (47 Qs)

An important study by Darren Davis (1997) documents significant race-of-interviewer effects on over 60% of the attitude questions in the 1984 National Black Election Study survey. He used a two-stage least squares procedure to correct for race of interviewer leading to correlated errors. As an example, other things being equal, African Americans interviewed by African Americans were nearly twice as likely to say they were planning to vote for Jesse Jackson than ones interviewed by whites. Yet it is not clear whether this means they were underreporting their plans to vote for him when interviewed by a white or overreporting when interviewed by a black.

One possible interpretation of these race-of-interviewer effects is that the validity of the data would be greater if black interviewers were assigned black respondents and white interviewers were assigned white respondents—race matching. However, this ignores the real possibility that interviews among same-race actors may yield overreports of extreme positions on racial attitude questions, which would become a bias in the study. Given that there is not necessarily a “correct” answer on racial attitude questions, it is important to be able to assess the extent of interaction effects by having some same-race and some opposite-race interviews. Advocates of standardized interviewing draw yet a different inference. Fowler and Mangione conclude: “The fact that interviewer race can predict answers means that the measurement is not standardized across race of interviewers. When answers are significantly affected by interviewer race, measurement error would be reduced if interviewers were all of the same race” (1990, 101). Although the implication that Fowler and Mangione derive in that last sentence seems to be a way of minimizing measurement error, it also could be read as inadvertently justifying discrimination in hiring interviewers. Furthermore, standardizing on race will not diminish measurement error if people are more likely to give valid answers to interviewers of their own race. Given the several different interpretations of this effect that are possible (Rhodes 1994), it is critical that further research be done before accepting it as true and using it as a basis of interviewer hiring and assignment.

Interviewer matching is rarely used in the United States, except when it is necessary to use interviewers who can speak another language to interview non-English-speaking respondents. Interviewer matching is more necessary in some other countries, as in Arab countries where it would be considered inappropriate for an interviewer of one gender to speak with a respondent of the other gender.

One other interviewer-related problem occurs when panel surveys have the same interviewer repeatedly interview a respondent. Mensch

and Kandel (1988) find an *interviewer familiarity effect*, with people interviewed previously by the same interviewer in the National Longitudinal Study of Youth (NLSY) being less likely to admit to having used marijuana and cocaine. Presumably the respondents were less willing to admit to these behaviors with an interviewer whom they already knew. The use of the same interviewers in long-term panel surveys is usually seen as useful for increasing the retention rate of respondents, but this finding shows that it can negatively affect the accuracy of reporting on sensitive questions.

Another interviewer-related issue is *interviewer experience*. Paradoxically, some studies of face-to-face interviewing seem to find that experienced interviewers do not obtain as complete information from respondents as new interviewers (e.g. Hughes et al. 2002). This fits with a study finding that interviewers who completed more interviews obtained fewer reports of known hospitalizations than those with smaller interview loads (Cannell, Marquis, and Laurent 1977). Groves et al. (2004, 273) suggest that the mechanism for these results may be that more experienced interviewers may try to maximize response rates and productivity more than answer quality.

There could also be social psychological and sociological explanations for response error associated with interviewers, such as the role of *interviewer expectations*. Empirical studies find little support for these expectations' affecting answers, but the studies have their own limitations, such as looking at sets of questions rather than individual questions (Groves 1989). Leal and Hess (1999) find that interviewers can be biased in what they think of respondents. Looking at the American Citizen Participation Survey, the Latino National Political survey, and the 1996 National Election Studies survey in which interviewers were asked to evaluate respondents, they find that even with controls for the respondent's actual level of political information, interviewers considered respondents of lower socioeconomic status to be less informed and intelligent. There were also some racial effects, with blacks being seen as less intelligent by interviewers in the participation and NES surveys, again with controls for objective levels of information. However, there is no direct evidence of how these biases translate into survey error.

A possible implication of the research on interviewer effects is that the selection of interviewers matters, but there are too few empirical studies, and many of those studies have methodological problems. There is need for more careful research designs in which comparable sets of respondents are assigned to different types of interviewers, but in the end it may be

impossible to obtain a definitive determination of the importance of interviewer effects.

Interviewer Management Issues

Fowler and Mangione (1990, 105) argue that the focus should be on interviewer training and supervision more than on interviewer selection. Interviewer-related error can be minimized by a survey organization through its procedures for managing interviewers, including training procedures, supervision, and incentives for interviewers.

Interviewer Training

A common question is what types of interviewers should be hired. Some survey organizations begin the hiring process by testing the ability of prospective employees to read survey questions exactly as written. Additionally, it is important to hire people who will show up reliably. Beyond these minimal requirements, training may be more important than hiring. This is apparent in the choice some organizations face between hiring professional and student interviewers. There are benefits to both, assuming the students are given adequate training. Having experienced interviewers is important to professional survey organizations. Yet student interviewers are often easier to schedule for part-time work and easier to add (or drop) with the ebb and flow of survey business, but more money is spent on training since student interviewers rarely last longer than a few semesters. Many academic survey organizations use a mixture of the two types of interviewers, often with the hope that the presence of professional interviewers will help give student interviewers the appropriate attitude toward their job. Even a few large commercial organizations such as Gallup supplement their professional staff with student interviewers, in the belief that proper interviewer training is what is really important.

Interviewer training is particularly urgent when interviewing is not carried out on a regular basis. A survey organization that does interviewing every week of the year is better able to maintain a well-trained staff than one that conducts a survey only once every few months. Election exit polls pose the most extreme training problem: they generally occur only once every other year in the United States, which essentially requires hiring and training a new interviewing staff each time. Thus it is not surprising that some of the blame for exit poll errors in the 2000–2004 U.S. elections was ascribed to inadequate interviewer training. Indeed, exit

polls should be considered the most difficult type of survey to conduct well, because of the need to start virtually from scratch every few years.

Training Guidelines

Since measurement error due to interviewers is an important component of survey error, it is essential to train interviewers so as to minimize interviewer-related error. Interviewer training includes both training new interviewers in standard procedures and giving them training on new surveys. Every survey organization handles this training somewhat differently, but high-quality survey organizations all emphasize the training they give to interviewers.

The Council of American Survey Research Organizations (CASRO) has developed a set of guidelines, available on its Web site, <http://www.casro.org>, as to what should be covered in general interviewer training. They include an explanation of the research process, the importance of professionalism of the interviewer, the need for confidentiality, and the importance of maintaining positive feelings by the respondent about survey research.

The general training given interviewers includes how to obtain interviews and how to conduct them. Obtaining interviews is not easy, especially telephone interviews, since people find it easy to hang up when they hear that they are about to be interviewed. As will be seen in chapter 8, survey organizations develop standard responses that interviewers are supposed to give when people ask particular questions such as "How did you get my number?"

The training on conducting interviews emphasizes the interviewing philosophy of the survey organization. Interviewers will be told to read questions exactly as they are written, but differences arise in how interviewers are told to handle questions from the respondent. Organizations that follow a standardized approach will emphasize the need to repeat the question as originally worded without ad-libbing. Organizations that follow a conversational approach will instead explain how to collaborate with the respondent in establishing a meaning for the question that fits the research objectives. Similarly, the interviewer training should explain the type of probing that is used by the survey organization. In any case, the importance of recording the respondent's answers verbatim is stressed. Academic survey organizations that employ the total survey error approach often include an introduction to that approach in their interviewer training, so that the interviewers recognize the importance of their effort to the overall goal of minimizing survey error.

Some survey organizations use videotapes as part of their training. Such tapes can include an introduction to the organization and its history, a statement of the importance of conducting high-quality interviews, and an explanation of the interviewing philosophy that is employed. Training tapes often include examples of how not to interview as well as examples of interviews that are conducted properly.

Interviewer training must also include specific training with any technology that is used in the interviewing. Phone surveys nowadays are generally computer assisted, as are many face-to-face interviews, so the training for such interviewing must include familiarization with the hardware and software to be used.

Interviewer training should include opportunities to take practice interviews. The first practice interviews are typically with fellow trainees in a role-taking mode, with one trainee sometimes taking the role of an uncooperative respondent. Later practice interviews are usually taken with real people, with these interviews not included in the final data.

In addition to general training, it is important to train interviewers on new surveys. Such training includes both going over the purposes of the survey and giving interviewers a chance to practice the survey. Again, most survey organizations have interviewers take a few practice interviews with real people who are not part of the sample so as to learn the interview schedule.

CASRO lists several specific points to address when briefing interviewers on new projects. Interviewers are to be instructed on the sampling method that was employed, how they should handle materials, the anticipated length of interviews, skip patterns for controlling the question order, specific directions on clarifying questions and on probing, and the approach to validating interviews. Additionally, the briefing sessions should include reading the interview instructions and reading the questionnaire itself, followed by practice interviews. Even an experienced interviewer cannot be expected to get a high-quality interview the first time on a new questionnaire, so practice is always important.

Studies of Interviewer Training

Several studies have examined the effectiveness of interviewer training. A Belgian study (Billiet and Loosveldt 1988) showed that interviews conducted by trained interviewers had less missing data and more probing. Fowler and Mangione (1990, chap. 7) conducted an extensive experiment with different amounts of training for face-to-face interviewers. Interviewers with less than one day of training did worst. There was little

difference between interviewers with two to ten days of training, except for probing. Probing is one of the hardest parts of interviewing, so training pays particular benefits in improving the ability to probe. Also, interviewers with more training were seen as more task oriented, whereas those with less training were seen as more interpersonally oriented; however, interviewers with the least training were marginally more likely to be rated by researchers as excellent. Still, two and five days of training produced more standardized results than one or ten days. Fowler and Mangione concluded that ten-day training is too long, with interviewers becoming tired of being trained, but two to five days is appropriate. They generally found better quality when interviews were tape recorded. Only the largest survey organizations would even consider having five to ten days of interviewer training, but this study usefully implies that even small survey organizations would benefit from doing more than one day of training. Additionally, interview supervisors are often instructed to give special attention to new interviewers, so that, in effect, the interviewers get additional training in their first few days on the job.

As a means of measuring interviewer compliance with training guidelines, Cannell, Lawson, and Hausser (1975) developed several categories for "interaction coding" to keep track of the interactions between interviewers and respondents. Interviews are monitored or tape recorded, and then tallies are made to count how often the interviewer asks questions exactly as written, questions are repeated correctly, definitions and clarifications are given correctly, the interviewer delivers short and long feedback correctly, and the pacing of the interview is correct. The interviewer's overall clarity is also recorded, in terms of how naturally questions are read (unnatural reading could involve awkward inflection, wrong emphasis, monotone voice, or wooden delivery) and mispronunciations.⁷ Interaction coding can be used to identify questions that are problematic, as well as interviewers who do not perform in a standard manner. However, Groves (1989, 389) cautions that there is no direct evidence that behaviors measured in interaction coding are related to measurement error.

Interviewer Supervising

Interviewer supervision provides another opportunity to work on the quality of the interviewing so as to minimize measurement error due to interviewers. The organization of supervision varies between survey organizations, often with first-level supervisors being responsible for shifts and reporting to higher-level supervisors or directly to a field manager.

First-level supervision consists of many tasks, from scheduling and problem solving to performance monitoring and interview validation. The goals of a supervisor include maximizing productivity, increasing the data quality, and answering questions raised by interviewers.

Supervisors for both telephone and face-to-face interviewing are responsible for assigning interviews to interviewers when computer programs do not handle the assignment. In phone interviewing, they get the site ready before the interviewers come in for each shift, they keep track of forms filled out by interviewers when people refuse to be interviewed or ask to be called back later, and they deal with problems that come up during their shift. If the survey is on paper (rather than on a computer), the supervisors collect the survey forms after they are filled out and check that the interviewers are doing a good job. In phone surveys, the supervisors make sure that the interviewers are calling continuously and that interviewers keep appointments (as when a respondent indicates that she can do the interview at a specific time). At the end of a shift, supervisors submit a shift report to the field manager, including the numbers of completed interviews and refusals for each survey, as well as which interviewers did not show up for their scheduled work time.

Interviewer supervision also provides an opportunity for performance monitoring. Fowler and Mangione (1990) suggest that a trained supervisor monitor one in ten interviews. Many telephone survey organizations have developed standardized forms for monitors to fill out so as to provide systematic evaluation of the interviewing. As an example, the interviewer monitoring form (IMF) used by Ohio State University's Center for Survey Research has a supervisor record how the interviewer did in the introduction to the survey, including information given to the respondent, handling the request for the interview, enthusiasm, confidence, sincerity, answering respondent questions, and types of questions asked for scheduling the interview. The interviewer's performance in the actual interview is scored in terms of assertiveness, professionalism, pace, clarity, tone, and overall performance. Additionally, the supervisor tallies errors made by the interviewers in terms of minor or major deviations from question wording, not clarifying ambiguous responses, interpreting questions for the respondent, making extraneous remarks, offering response options improperly, and failing to probe open-ended responses. Ideally the supervisor meets with the interviewer immediately after the interview to go over the form together. Monitoring is often more informal than this, with the supervisor listening in and keeping only a mental note of items to mention to the interviewer afterward. In practice, these are often supportive

comments, such as consoling an interviewer who has just survived a difficult respondent.

Some organizations have had good experiences with tape-recording interviews. An experiment by Fowler and Mangione showed that taping the interviews did not have negative effects on the interviewer's feelings and had no effect on the respondent's feelings, but it did result in better probing for one of the groups as well as less bias. In computer-assisted personal interviews using laptop computers, it is now also feasible to tape record interviews with the computer's built-in microphone, thereby improving quality control. Taping and supervision are considered to have positive effects even for the best and most experienced interviewers, partly to correct their tendencies to feel they have developed enough skill to substitute their own judgment for standard procedures.

Interviewer supervision is most effective with phone interviews from a single central location. It is least effective when the interviewing is dispersed, as in national election exit polls where one interviewer is assigned to each sampled precinct, making it impossible to determine how well (or even if) interviewers are doing their job. Some survey organizations now conduct telephone interviews through a "distributed network" with CATI interviewers phoning from their homes (or outsourced to installations in another country) while using a broadband computer connection to the Internet to access the survey. This innovative procedure may make it easier to recruit interviewers and/or lower interviewer costs, but it requires further efforts to maintain effective interviewer supervision.

Supervision should include some interview verification, to make sure that the interviewer is not faking interviews. Interviewers sometimes pretend to be doing interviews, expecting that they will not be caught. Many telephone interviewing operations have the capability of a central supervisor's listening in on the interview, which provides one means of assuring that real interviews are occurring. Modern monitoring programs actually keep track of some aspects of the phone calls, such as how long the person is on the phone and which question is being asked, and this allows supervisors to make sure that appropriate progress is being made. Tape recording interviews is another means of verification. Some centralized facilities use video monitoring of interviewers. Validation through recontacting respondents is common also for face-to-face interviewing, with supervisors phoning a sample of respondents to make sure that they were actually interviewed. CASRO guidelines state that "standard industry practice" involves validating 15% of the interviews. Large survey firms sometimes employ an independent organization to validate interviews in order to

assure the client that the validation was real. Obviously interviewers who are found to fabricate interviews are terminated, and interviews that they conducted are closely examined to determine which must be discarded. In an unusual case, several students were disciplined by California State University–Stanislaus after they admitted fabricating interviews in a class survey on the bias of prospective jurors that was used by a judge in deciding to move the 2004 Scott Peterson murder trial out of Modesto, California.

The Office of Research Integrity of the U.S. Department of Health and Human Services has ruled that falsification by survey interviewers is an act of scientific misconduct. That action has led to work on a best practices document on interviewer falsification (American Statistical Association 2003). In addition to fabricating interviews, the document includes as falsification the deliberate misreporting of disposition codes (such as recording a refusal as ineligible so as to make the refusal rate look low), deliberate miscoding of answers (as when the interviewer does not want to ask follow-up questions), and deliberate interviewing of a person who was not sampled (so as to increase the number of interviews obtained). The document emphasizes the importance of promotion of the proper values in interviewer hiring and training, adequate supervision, quality control, avoidance of excessive workloads, and a pay structure that is not based exclusively on number of completed interviews.

Interviewer Incentives

Interviewer pay is one of the largest expenses for surveys. Some survey organizations skimp on pay, but most higher-quality organizations believe that it is better to pay more—well above minimum wage—and hire fewer but better interviewers.

Survey organizations have tried different motivating systems, some emphasizing incentives and others emphasizing demerits. Given how difficult it can be to get respondents to cooperate for phone interviews, giving incentives for completed interviews is a good way to keep interviewers motivated—either extra pay or points that can be accumulated for prizes. Overreliance on incentives for completed incentives, however, gives interviewers a reason to fake interviews. Demerit systems are especially common when using student interviewers, with supervisors keeping track of shifts missed without calling in and interviewers being fired if they get a certain number of demerits a month. Some survey organizations try to motivate the callers by getting them involved in the survey process, but

most maintain a clear distinction between the interviewing and the rest of the operation.

Interviewer-Related Error and Survey Constraints

The interviewers' job is difficult because it is multifaceted. They must first convince the respondents to participate, then motivate the respondents to give full answers, and at the same time record their answers accurately. This requires a combination of human skills, quick thinking, and organization, as well as a personality that does not get depressed by refusals.

At the same time, it is not possible to do away with interviewer effects. Some interviewer error is inevitable, as illustrated by studies showing that even standardized interviewers routinely deviate from the exact question wording. Additionally, there can be effects associated with the characteristics of the interviewer, effects that cannot be eliminated since no interviewer characteristics are entirely neutral.

Interviewer-related error is one of the few sources of survey error that are under the direct control of a survey organization. Proper training of interviewers, appropriate supervision, and good pay all can help keep the quality of interviews high. Each of these is, however, costly, which illustrates the continual trade-off between survey errors and costs. When a researcher gives a survey contract to the lowest bidder, the low bid is often due to cutting corners in these respects, which can lead to lower data quality. The trade-off is probably most severe in attitude surveys that include many open-ended questions requiring skillful interviewing.

As an example of the trade-offs that are involved between survey errors and costs, Fowler and Mangione (1990, 144) estimate the effects of different ways to reduce survey errors. Say that the goal is to decrease the standard error on variables by 10%. The standard way to do so is to increase the sample size, which would require taking about 20% more interviews, which increases interviewing costs by 20%. Alternatively, they suggest that the standard error on questions that are most affected by interviewers could be cut by 10% through use of three procedures that have been discussed in this chapter: adding another day to interviewer training, tape recording interviews, and/or reducing the number of interviews assigned to each interviewer. These procedures are not costless, but they increase cost by much less than 20%.

Time is also a relevant constraint. Hiring more interviewers, for example, allows more interviews to be done in a shorter period of time, which will be important to some clients. On the other hand, extensive inter-

viewer training takes time—and some interviewers would not bother returning for lengthy training before conducting actual interviews.

Ethics pose additional constraints. Interviewers can fake interviews; this makes it important to motivate interviewers with information about the importance of the research and through good pay so that they will have incentives to do good work. Additionally, emphasizing the role of the interviewer in helping minimize total survey error is a way of making interviewers understand the importance of their contribution to the research effort. At the same time, validating some of the interviews—with interviewers knowing that this is done—lessens the likelihood of interviews' being faked. Interviewers also face conflicts between increasing response and honoring the respondents' right to voluntary participation; the temptation is to push the person into participating so as to maximize response. Another important ethical requirement is that interviewers are supposed to keep information they obtain from respondents confidential, not telling friends the names of respondents or what they said.

This chapter's consideration of interviewer-related error led to consideration of how to view the interviewer's role: whether the interviewer should act like a therapist, friend, automaton, or partner. Chapter 5 turns to the other side of the coin: the role of the respondent.