An Introduction to Survey Research

Ernest L. Cowles
Edward Nelson

An Introduction to Survey Research

An Introduction to Survey Research

Ernest L. Cowles and Edward Nelson

Abstract

This book is an introduction to survey research for those who want an overview of the survey process. It is intended to describe fundamental survey components to help both students and managers understand and use surveys effectively and avoid the pitfalls stemming from bad survey construction and inappropriate methods. We begin by talking about how best to identify the information needed and the best approach to get that information. We then discuss the processes commonly involved in conducting a survey including the value of both obtaining a representative sample and dealing with the types of errors that can distort results. Next, each chapter focuses on one of the key components of constructing and carrying out a survey, including the elements to consider when developing a survey, the modes of survey delivery, writing good questions, conducting the survey, and presenting the results. Each chapter concludes with a summary of important points contained in the chapter and an annotated set of references indicating where readers can go for more information on chapter topics.

Keywords

Contents

Preface

Survey research is a widely used data collection method that involves getting information from people typically by asking them questions and collecting and analyzing the answers. Such data can then be used to understand individuals' views, attitudes, and behaviors in a variety of areas such as political issues, quality of life at both the community and individual levels, and satisfaction with services and products, to name but a few. Decision makers in both the public and private sectors use survey results to understand past efforts and guide future direction. Yet there are many misperceptions regarding what is required to conduct a good survey. Poorly conceived, designed, and executed surveys often produce results that are meaningless, at best, and misleading or inaccurate, at worst. The resultant costs in both economic and human terms are enormous.

Our purpose of writing this book is to provide an introduction and overview of survey research. We begin our exploration at the foundation of gathering information about the world—observation and questioning. We talk about identifying the information needed and the best approach to get that information. We discuss the processes commonly involved in conducting a survey including both obtaining a representative sample and dealing with the types of errors that can distort results. Next, we focus on the components of constructing and carrying out a survey, including the elements to consider when developing a survey, the modes of survey delivery, writing good questions, conducting the survey, and presenting the results.

Making use of what people tell us in surveys depends on a number of factors. The mechanics of putting a survey together and proper survey administration determine whether useful information is obtained or whether the results are meaningless or even misleading. The way questions are worded and the order in which questions are asked affect what respondents tell us. The way the researcher interacts with survey participants can influence not only what people tell us but even whether they will respond.

Factors such as economic status, gender, race and ethnicity, and age can also influence how people respond to questions.

This book describes fundamental survey components to help both students and managers understand and use surveys effectively and avoid the pitfalls stemming from bad survey construction and inappropriate methods. Each chapter focuses on a key component and has an annotated set of references indicating where readers can go for more information.

Acknowledgments

CHAPTER 1

Introduction

Research starts with a question. Sometimes these are *why questions*. Why do some people vote Democrat and others vote Republican? Why do some people purchase health insurance and others do not? Why do some people buy a particular product and others buy different products? Why do some people favor same-sex marriage and others oppose it? Why do some people go to college and others do not? Other times they are *how questions*. If you are a campaign manager, how can you get people to vote for your candidate? How could we get more people to purchase health insurance? How could you get customers to buy your product? How could we convince more people to go to college? But regardless, research starts with a question.

Think about how we go about answering questions in everyday life? Sometimes we rely on what people in authority tell us. Other times we rely on tradition. Sometimes we use what we think is our common sense. And still other times we rely on what our gut tells us. But another way we try to answer questions is to use the *scientific approach*.

Duane Monette and his associates suggest that one of the characteristics of the scientific approach is that science relies on systematic observations.[1] We often call these observations *data* and say that science is *empirical*. That means it is data based. However, the scientific approach doesn't help you answer every question. For example, you might ask whether there is a God or you might ask whether the death penalty is right or wrong. These types of questions can't be answered empirically. But if you want to know why some people vote Democrat and others vote Republican, the scientific method is clearly the best approach. Relying on what people in authority tell you or what tradition tells you or your gut won't work.

Research Design

Your *research design* is your plan of action. It's how you plan to answer your research questions. The research design consists of four main parts—*measurement*, *sampling*, *data collection,* and *data analysis*. Measurement is how you will measure each of the variables in your study. Sampling refers to how you will select the cases for your study. Data collection is how you plan to collect the information that you will need to answer the research questions. And data analysis is how you plan to analyze the data. You need to be careful to decide on your research design before you collect your data.

In this book, we're going to focus on data collection and specifically on surveys. We'll talk about sampling, survey error, factors to consider when planning a survey, the different types of surveys you might use, writing good questions, the actual carrying out of surveys, and survey reporting.

Observation and Questioning

Irwin Deutscher succinctly summarizes the different ways we collect data—"(1) we can observe it in process; (2) we can view the records men [and women] leave behind…; and (3) we can ask questions and listen to answers."[2] In this chapter, we're interested in two of these approaches—*observation* and *questioning*.

Matilda White Riley makes the following comment about observation and questioning noting that one method isn't inherently superior to the other but that observation and questioning focus on different aspects of the social setting we are studying.[3]

> Researchers sometimes feel—mistakenly, we believe—that they can obtain a true picture of a social phenomenon *only* if they observe it with their own eyes. To be sure observation and questioning often give different results; but this occurs, not because one method is more valid than the other, but because the two focus … on different sets of social system properties.

Observation and questioning give us different information about what is going on in the world. Observation gives us information about

what people do. Questioning gives us information about what people say and the context to help interpret their observations.* This suggests that we often need both observation and questioning to give us a complete picture of what is happening and why it happens.

Elliot Liebow in his book, *Tally's Corner*, provides a clear example of these two different approaches to data collection.[4] Liebow studied a group of men who hung out on street corners in Washington, DC. He notes that "men and women talk of themselves and others as cynical, self-serving marauders, ceaselessly exploiting one another as use objects or objects of income."[5] The men in Liebow's study "are eager to present themselves as exploiters to women as well as to men."[6] In other words, this is what they say. He goes on to say that "in practice, in their real relationships with real women, the men frequently gave the lie to their own words."[7] This is what the men do. So how does Liebow explain this apparent contradiction between what men say and what they do? He suggests that there are two opposing impulses at work. "The impulse to use women as objects of economic or sexual exploitation is deflected by countervailing impulses and goals, especially the desire to build personal, intimate relationships based on mutual liking and love."[8] The apparent contradiction between what the men say and what they do is explained by the "interplay of these opposing impulses."[9]

Let's consider another example. You're doing a market research survey for a company that manufactures condoms. You want to know whether people purchase condoms and the particular brands they buy. It's easy to imagine a discrepancy between what people say and what they do. Some people might be embarrassed to give you this information and others might feel that it's none of your business. What people say might not accurately reflect what they do.

Even though we see that observation and questioning give us different information about the world, we are still surprised when there is a lack of consistency between what we learn from observation and from

---

* Someone might point out that we can ask people questions about their behavior. For example, we might ask what products they bought or whether they voted in the last presidential election. Certainly we are getting information about how people report their behavior. But we know that people overreport how often they vote and underreport other things such as illegal or deviant behavior. We're still getting their subjective reports of their own behavior.

questioning. Deutscher in his book, *What We Say/What We Do*, describes an early study by Richard LaPiere.[10] In the early 1930s, LaPiere travelled across the United States with a Chinese couple. They ate together in restaurants and stayed at hotels and auto camps and were refused service only once, and this was during a time in the United States when there was considerable prejudice toward Chinese. Six months later, LaPiere sent a questionnaire to these same hotels and restaurants asking the following question—"Will you accept members of the Chinese race as guests in your establishment?"[11] He describes the results of his survey as follows:

> With persistence, completed replies were obtained from 128 of the establishments we had visited; 81 restaurants and cafes and 47 hotels, auto-camps, and "Tourist Homes." In response to the relevant question, 92 percent of the former and 91 percent of the latter replied "No." The remainder replied "Uncertain, depends upon circumstances."[12]

So what are we to make of this? Is this an example of the inconsistency between what people say and what they do? Or does it simply reflect that observation and questioning are telling us different things about the world? LaPiere's classic study sparked a great deal of interest and follow-up studies. Howard Schuman in his book, *Method and Meaning in Polls and Surveys*, describes a study that he and Robert Brannon carried out in 1969.[13] He refers to this as "an attitude-behavior field study."[14] In a survey, respondents were asked their opinion of open-housing laws. Here's the question they were asked and the percent of respondents giving each answer.[15] (DK stands for don't know and NA for no answer.)

Suppose there is a community-wide vote on the general housing issue. There are two possible laws to vote on.... Which law would you vote for?		
1.	One law says that a homeowner can decide for himself who to sell his house to or, even if he prefers not to sell to blacks.	82%
2.	The second law says that a homeowner cannot refuse to sell to someone because of their race or color.	16%
DK, Neither, NA		2%
Total		100%
N		(640)

Three months after the survey the same respondents were asked to sign a petition. One of the petitions supported the first law and the second petition supported the other law. Those who said they would sign the petition were then asked if they would be willing to have their name appear in the newspaper as a supporter of that petition. Schuman summarizes the overall consistency between what people said and what they were willing to do for those opposed to the open-housing law—"… 85 percent were consistent in signing the Owner's Rights petition, and 78 percent were consistent in refusing to sign the Open Housing petition which gives an overall average of 82 percent consistency."[16] The same type of consistency was also found for those who supported open housing. Schuman concludes that in this study "attitudes can predict behavior to a reasonable extent, though of course not perfectly."[17]

In a more recent study, Eleanor Singer and her associates studied the "impact of privacy and confidentiality concerns on participation in the 2000 Census."[18] This is another example of what Schuman referred to as the attitude–behavior question. Their analysis found that attitudes toward confidentiality and privacy were significantly related to behavior (i.e., returning the census form). It's interesting that they also report that other researchers found that "many more people … say they would not provide their SSN [Social Security number] to the Census Bureau than actually fail to provide it when it is asked for on their census form."[19]

There are many more examples of the attitude–behavior issue, but these are sufficient to show that sometimes people behave in a way that is consistent with what they say and other times what they say is different from what they do. As Liebow pointed out, there are various factors affecting both what people say and what they do, and it is the interplay of these factors that eventually determines the outcome. For our purposes, it is important to keep in mind that observation and questioning provide us with different information. Questioning tells us how people feel and observation provides us with information about how people behave. It's not surprising that sometimes these two types of information are consistent with each other and other times they are not. The focus in this book is on questioning and how we carry out surveys. We'll cover these topics in Chapters 4 (Factors to Consider When Thinking About Surveys), 5 (Modes of Survey Delivery), 6 (Writing Good Questions), and 7 (Carrying Out the Survey). But we should never lose sight of the fact

that this is not the same thing as observing how people behave and interact with each other.

Triangulation

Triangulation refers to the use of data from different sources and methods of data collection. All data suffer from different types of error and error is inevitable. We'll have more to say about that in Chapter 3 on Total Survey Error. It follows then that using data from different sources and methods of data collection is a powerful research strategy. Eugene Webb and his associates in their book on *Nonreactive Measures in the Social Sciences* put it this way—"If a proposition can survive the onslaught of a series of imperfect measures, with all their irrelevant error, confidence should be placed in it."[20] In other words, if data from different sources and methods of data collection lead to the same conclusion, then we can have more confidence in that finding.

An informative example of triangulation is William Foote Whyte's work on community conflict and cooperation in Peru.[21] Whyte studied 12 communities using both surveys and observation in the 1960s. In 1964 and 1969, the questions used to measure conflict were slightly different. In 1964 the survey asked "is there much conflict or division among the people of this village?" In 1969 the question was "is there much conflict in this village between the people who want to change things and the people who want to keep to the old ways?" The question used to measure cooperation was the same in both years: "When it comes to cooperating on some project for the community, how well do the people cooperate?"[22]

Whyte's analysis of the survey data focused on change over the five-year period. He found that four communities shifted from high conflict and low cooperation to low conflict and low cooperation. Only one community, Huayopampa, shifted from low conflict and high cooperation to high conflict and low cooperation. This single piece of data would probably have been dismissed were it not for the fact that Whyte also had observational data from graduate students who observed behavior in these communities. The data from the observers corroborated the survey findings. This led Whyte to reject the commonly held belief that conflict and cooperation were different ends of a single continuum and to assert that it was possible for communities to be high or low in both conflict

and cooperation, an important theoretical breakthrough. Whyte concludes that this "theoretical breakthrough… would have been impossible without the *combination* of research methods used in our program."[23]

This example of triangulation focuses on conflict and cooperation at the community level. But it is easy to see how this could be relevant for other studies of conflict and cooperation at the organizational level. Imagine that you were studying hospitals or multinational business corporations and you wanted to study conflict and cooperation among staff. You might employ the same strategy of using both survey and observational data to achieve triangulation.

Questioning (Interviewing) as a Social Process*

Interviewing is in some ways similar to the types of conversations we engage in daily but in other ways it's very different. For example, the interviewer takes the lead in asking the questions and the respondent has little opportunity to ask the interviewer questions. Once the respondent has consented to be interviewed, the interviewer has more control over the process than does the respondent. However, it is the respondent who has control over the consent process and it is the respondent who determines if and when to terminate the interview. We'll discuss nonresponse in Chapter 3 on Total Survey Error and in Chapter 7 on Carrying Out the Survey.

Raymond Gorden has provided a useful framework for viewing the interview as a social process involving communication. Gordon says that this communication process depends on three factors—"the *interviewer*, the *respondent*, and the *questions* asked."[24] For example, the race and gender of the interviewer relative to that of the respondent can influence what people tell us and we know that the wording and order of questions can also influence what people tell us. We'll discuss these considerations in Chapter 3 on Total Survey Error.

Gorden goes on to suggest that the interaction of interviewer, respondent, and questions exists within the context of the interview situation.[25] For example, are we interviewing people one-on-one or in a group setting? Many job interviews occur in a one-on-one setting, but

---

* From now on we will refer to questioning as interviewing.

one of the authors recalls a time when he was one of several job applicants who were interviewed in a group setting involving other applicants. Rest assured that this affected him and the other applicants. Are we interviewing people in their homes or in another setting? Think of what happens in court when witnesses are questioned in a courtroom setting. That clearly affects their comfort level and what they say.

Gorden notes that the interview and the interview situation exist within the context of the culture, the society, and the community.[26] There may be certain topics such as religion and sexual behavior that are difficult to talk about in certain cultures. Norms of reciprocity may vary from culture to culture. Occupational subcultures, for example, the subcultures of computer programmers and lawyers, often have their own language.

It's helpful to keep in mind that the interview can be viewed as a social setting that is affected by other factors just as any social setting is influenced by many factors. In this book, we will be looking at many of the factors that affect the interview. We'll look at the research that has been done and how we can use this research to better conduct our interviews.

Brief History of Surveys

Don Dillman and his associates have an excellent and brief history of *surveys* in their book *Internet, Mail and Mixed-Mode Surveys – The Tailored Design Method*. Dillman says that "during the first two thirds of the 20th century, there existed only one generally accepted mode for conducting surveys: the in-person interview."[27] Edith DeLeeuw reports that the "first scientific face-to-face survey" was a "study of working-class conditions in five British cities" in 1912[28] and "the first documented mail survey" was in 1788 sent to "ministers of all parishes of the Church of Scotland."[29] By 1980, Dillman says that the mail survey was commonly used, and by the early 1980s, the phone survey became a dominant mode. By the mid-1990s, another form of surveying had emerged—the web survey.

One of the factors that influenced surveys was *technological change*.[30] By 1970 almost 90 percent of households had a landline phone.[31] By the early 2000s, the cell phone was commonly used, particularly by young males. Caller ID made it easier for people to screen their calls and not take unwanted calls. E-mail and the web created a new medium of

communication. By the mid to late 2000s web surveys were common. Today there has clearly been a shift to what are called *mixed-mode surveys*, which rely on a combination of face-to-face, mail, phone, and web-based surveys along with new technologies that have appeared such as the Interactive Voice Response survey where respondents use their touch-tone phone to record their answers and Audio-Computer-Assisted Self-Interviews, which are respondent-administered surveys on a computer.

During the last 50 years, there were shifts in the way people used these technologies. Response rates to phone surveys began to drop because of the difficulty in contacting respondents and survey refusal. Roger Tourangeau and Thomas Plewes conducted an extensive review that looked at nonresponse in a number of large surveys. They conclude that "the experience of this illustrative set of surveys provide evidence that nonresponse rates continue to increase in all types of cross-sectional surveys, with little to suggest that the trend has plateaued."[32] They go on to point out that the increase in nonresponse rates for phone surveys has been particularly large.

Two events are worth noting in the history of surveys. In 1936, *The Literary Digest* conducted a mailed survey of potential voters in the presidential election which pitted the Democrat incumbent President Franklin Delano Roosevelt against the Republican Governor Alf Landon.[33] They sampled phone numbers in telephone directories and names in state automobile registration lists. Their sample was extremely large totaling over two million potential voters. However, it did not include those without phones and automobiles and clearly did not adequately represent the population. *The Literary Digest* survey predicted a victory by Landon but Roosevelt won in a landslide. This clearly demonstrated the danger of using lists that were biased in terms of variables such as education and income.

Another significant event occurred in the 1948 presidential contest between Democrat incumbent President Harry Truman and the Republican Governor Tom Dewey.[34] The major polls such as Gallup and Roper predicted that Dewey would defeat Truman but Truman prevailed in the election. This brought about changes in surveys such as the decline of quota sampling and the practice of continuing polling right up until Election Day.

The history of surveys demonstrates the influence of societal changes on surveying. The dramatic increase in landlines in households by 1970 and the rapid increase in the 2000s in the use of cell phones along with the development of a significant segment of individuals who use only cell phones have changed surveys drastically.[35] It is likely that this will continue in the future. Address-based sampling is another development that is becoming common where residential addresses are sampled from the U.S. Postal Service's Computerized Data Sequence File and made available through third-party providers.

The Rest of the Book

Here's a brief summary of what we will cover in the rest of the book.

- Chapter 2—Sampling—What are samples and why are they used? In this chapter, we discuss why we use sampling in survey research, and why probability sampling is so important. Common types of samples are discussed along with information on choosing the correct sample size.
- Chapter 3—Total Survey Error—Error is inevitable in every scientific study. We discuss the four types of survey error—sampling, coverage, nonresponse, and measurement error focusing on how we can best minimize it.
- Chapter 4—Factors to Consider When Thinking About Surveys—In this chapter some of the fundamental considerations about surveys are presented: the stakeholders and their roles in the survey process, ethical issues that impact surveys, factors that determine the scope of the survey, and how the scope, in turn, impacts the time, effort, and cost of doing a survey.
- Chapter 5—Modes of Survey Delivery—There are four basic modes of survey delivery—face-to-face, mailed, telephone, and web delivery. We focus on the critical differences among these different modes of delivery and the relative advantages and disadvantages of each. We also discuss mixed-mode surveys, which combine two or more of these delivery modes.

- Chapter 6—Writing Good Questions—Here we look at survey questions from the perspective of the researchers and the survey participants. We focus on the fundamentals of the design, formatting, and wording of open- and closed-ended questions, and discuss some of the most commonly used formats used in survey instruments.
- Chapter 7—Carrying Out the Survey—Every survey goes through different stages including developing the survey, pretesting the survey, administering the survey, processing and analyzing the data, and reporting the results. Surveys administered by an interviewer must also pay particular attention to interviewer training.
- Chapter 8—Presenting Survey Results—This final chapter talks about the last step in the survey process—presenting the survey findings. Three major areas, the audience, content, and expression (how we present the survey), which shape the style and format of the presentation, are each discussed along with their importance in the creation of the presentation. The chapter concludes with a discussion on how to structure different types of presentations such as reports, executive summaries, and PowerPoints, and how to effectively present survey data and results.

Annotated Bibliography

Research Design

- Matilda White Riley, *Sociological Research I: A Case Approach*, is an early but excellent discussion of research design.[36] Her paradigm of the 12 decisions that must be made in constructing a research design includes the alternative methods of collecting data—observation, questioning, and the combined use of observation and questioning.
- Earl Babbie, *The Practice of Social Research*, is a more recent introduction to the process of constructing a research design.[37]

- Delbert Miller and Neil Salkind, *Handbook of Research Design & Social Measurement*, provides many examples of the components of the research design.[38]

Attitudes and Behavior

- Irwin Deutscher, *What We Say/What We Do*, is an excellent discussion of the age-old question of the consistency between what people say and what people do.[39]
- Richard LaPiere, *Attitudes vs. Actions*, is the classic example of the inconsistency between attitudes and behavior.[40]

Triangulation

- William Foote Whyte is the author of *Street Corner Society*, which is one of the classic examples of participant observation that relies not only on observation but also on questioning.[41] His study of 12 rural communities in Peru is an excellent example of how observation and questioning can be combined in the research design.[42]

Questioning (Interviewing) as a Social Process

- Raymond Gorden, *Interviewing: Strategy, Techniques, and Tactics*, is one of the clearest discussions of the communication process and the factors that affect this process.[43]

Archival or Available Data

- There are different sources of data—observation, questioning, and archival or available data. Although archival or available data is not something that we're going to discuss in this book, if you want to learn more about it, the classic introduction is found in Eugene Webb and his associates' book on *Nonreactive Measures in the Social Sciences*.[44]

CHAPTER 2

Probability Sampling

Did you ever go into a coffee shop and discover they were giving away samples of a new blend they were serving? Or how about going to a grocery store and finding a person in one of the aisles distributing hot samples of a new pizza bread snack in the frozen food section? Beyond the obvious purpose of enticing you to buy the drink or food being offered, you probably never gave much thought to the small portion of the drink or food you were offered. Whether you would consider giving the merchant your hard-earned money for a full cup or a large package of pizza bread depends not only on your food preferences and whether you liked the taste, but also on an underlying assumption that you make regarding the sample you were given—namely that it is *representative* of the gallons of that particular coffee blend that is brewed, or that the large boxes of pizza bread snacks found in the freezer case are the same as the bite size piece you were offered. Imagine the disaster for both merchants and customers if it were not. This little example illustrates the important elements that pertain to the kind of sampling done in conjunction with surveys.

If done properly, a sample can adequately replace the need to examine each item or individual in a population (sometimes referred to as the population or universe) in order to determine if a particular characteristic is present. For example, suppose you want to find out if students who attend colleges that are affiliated with particular religious faiths have different attitudes toward the legalization of marijuana than students who attend colleges that have no religious affiliation.[1] It's obvious that if you planned to answer your question by interviewing or providing a written survey to all the students in the country who attended both religiously affiliated and nonaffiliated colleges, it would be impossible due to the cost and time required (along with a host of other practical considerations). Instead, by carefully selecting a manageable sample that actually represents students from both types of educational institutions, it is possible to answer the

question. The key is in getting a sample of individuals that accurately represents all of the individuals in the larger population that you're interested in studying. Now here's a little secret: the only way we can be absolutely certain that we have determined the true picture of the attribute that we are interested in studying, in this case attitudes toward marijuana legalization, is to ask each individual in the population about her or his attitudes toward legalization, then count up, and report the responses that are given; a process called *enumeration*. If we could actually pull this off, we might find, for example, that 86 percent of students at nonreligiously affiliated colleges favored legalization, but only 63 percent of students at religiously affiliated schools were in favor of legalization. We could be very certain that our findings were actually correct, and it really wouldn't take any more complicated math than addition and division to get our percentages. But that creates a dilemma—we've already said that it is not possible to get information from each person for a variety of reasons. As a result, we are forced to draw a sample and be content knowing that there is some probability that our sample will miss the mark in terms of accurately representing the population. This fact has been the force driving two major fields of study regarding research, including survey research. One of these has concerned itself with methodology of sampling and problems due to the way the sample was selected or drawn—which we term *sampling error*. The second area, also part of the larger field of statistics, has focused on determining mathematical probabilities of error and developing mathematical formulas to estimate true values with a variety of data.

In the next chapter we discuss *sampling error* and some ways to deal with it, but for now, we'll simply look at different kinds of sampling approaches and see that how a sample is selected can make a big difference in how much confidence we can have that it actually represents the larger population we want to study.

A few years ago, a large state mental health agency approached one of us about improving the quality of a survey that it administered annually to individuals who were receiving community mental health services throughout the state. The survey was designed to capture the kinds of services used at community mental health facilities and clients' satisfaction with those services. After reviewing the data and the process used to administer the survey, a number of problems came to light. One of the

most serious of these was the manner of selecting clients to participate in the survey. The selection of participants and administration of the survey involved staff members handing the survey to every individual who came into a community mental health facility during one specified two-week period during the year. Can you see any problems with selecting a sample in this way? If your answer included the idea that samples selected this way might not be truly representative of the population using community mental health services, you are right on target. This sample is essentially what is known as a *convenience* sample, and it is a *nonprobability* sample because it relies on individuals who happen to be available to take the survey. Essentially, individuals who did not pay a visit during the two-week sample period had no chance of being included in the sample. Further, while the agency knew the total number of individuals receiving services over the course of a year, they had no way of determining how many would be in the sample beforehand. This is important because as we'll discuss later in the chapter, the size of the sample affects its ability to accurately represent the population from which it is drawn.

What Is a Probability Sample?

The basic idea of survey sampling emerged in Europe at the beginning of the last century,[2] but a major theoretical underpinning for probability sampling seen today is commonly attributed to a Polish born mathematician and statistician who presented a paper entitled "On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection," at the Royal Statistical Society in June 1934.[3] Neyman's influential work would help create an acceptance of probability sampling and shape the basic construction of probability samples that has occurred since. During the ensuing eight decades since his paper, the basic structure of probability sampling has been further developed and refined, but at its core are three assumptions: (1) that a frame of all units in the population can be created (this is called the *sampling frame*), (2) that a person or thing in the frame has a (positive) likelihood of being selected into the sample, and (3) that the likelihood or probability of being selected can be computed for each sampled unit. The advantage of probability sampling over nonprobability methods *is not*

that it eliminates the possibility of picking a sample of individuals or units that doesn't accurately reflect the population, but that it does allow us to estimate the amount of error that may be present in our findings. In other words, we can determine the probability or odds that we have accurately represented the population about which we're trying to draw inferences. In more technical language, we can gauge the precision of our sample estimates. Thus, when you read that a survey found 46 percent of the public approve of the job the president of the United States is doing and the survey had ±3 point margin of error, you can assume (all else being equal) that the actual percentage of public approval has a known probability of being within 43 and 49 percent. We cannot make this type of statement with nonprobability sampling methods because we lack the fundamental ability to calculate the probability that each population element will be chosen, and in fact, we cannot be sure that each individual or thing has a possibility of being selected. Stated in more technical language, with probability sampling we say that each element of the population element has a nonzero probability of being chosen. Thus, while a nonprobability sample may accurately mirror the population from which it is drawn, we don't know whether it does or not.

Returning to the example of the mental health clients' survey, clients who had no visits during the two-week sample period had no chance (a zero probability) of being in the sample. On the other hand, because the survey's selection criterion was a time interval, individuals who visited community mental health centers many times during the year would have a greater likelihood or greater probability of being selected to take the survey. Making this selection approach even more troublesome was the fact that clients were asked to complete a survey each time they came to the community mental health center—thus a single client could have multiple surveys included in the sample. In fact, one individual had completed the survey 16 times during the two-week sample period! The possibility that more frequent visits to the mental health centers might be associated with more chronic or severe mental health issues highlights the possible problems of this sampling approach—the greater likelihood these individuals would be completing surveys might present a different picture of services than the larger population utilizing community mental health services.

The samples that have the greatest likelihood of being representative are those in which the units (in this case people) have an equal probability

of being selected. We term these, *equal probability of selection methods* (EPSEM).

Simple Random Samples

There are many ways to obtain samples in which each person in a population has an equal chance of being selected for a sample; the most straightforward of these is called the *simple random sample*. Simple random samples work when every individual or thing in the population that is to be sampled can be identified. For example, while possible, can you imagine the work you would have to do to compile a complete list with contact information of everyone who had purchased a new Toyota in the past five years? However, if you had such a listing you would be able to use simple random sampling in the same way that you would use it to select a sample of students from three introductory business classes. Once you have a listing of every element (people or things) in the population, you simply need a way to select the number of them you need for your sample in a truly random fashion. The key here is ensuring the randomness of the selection, which gives each element an equal chance or probability of being selected.

A simple random sample method with which most people in this country are familiar (although it is doubtful they think about it in that way) is the ubiquitous state or multistate lottery where five or six ping pong balls are (hypothetically) randomly drawn from a universe of around 50 numbers that are represented by numbered ping pong balls, which are randomly mixed in some type of rotating cage or swirling air container. Each numbered ball should have an equal chance of being selected from the hopper on a given draw except, of course, those already selected, which would entail a different kind of sampling.* The objective is to correctly choose the five or six numbers that are represented on the ping pong balls before they are drawn, which results in winning an enormous *jackpot* usually worth millions of dollars. Individuals paying to play the

---

* The typical sampling done in lotteries is termed sampling *without replacement* because once a numbered ball is selected it is taken from the pool so it cannot be selected again. If it were returned to the pool of balls, the sampling approach would be termed *with replacement*. The odds of correctly choosing the entire sequence of five or six balls needed to win the lottery changes dramatically under the two methods.

lotto hope to preselect the numbers drawn and frequently invent myste-rious and unfathomable methods to choose the right numbers. In reality, because the drawing represents a simple random sample of all the ping pong balls in the lotto universe, they could simply pick any sequence of numbers, say one through five, and they would have just as good a chance of winning as using some supposed system of getting their *lucky* numbers.

Unfortunately, not all probability sampling methods are as easily done as drawing a simple random sample of ping pong balls in a lottery, or pulling random numbers on strips of paper from a hat. After the intro-duction and widespread acceptance of probability sampling, but before researchers had the power of today's computers and statistical software at their fingertips, statisticians and researchers devoted considerable energy to finding consistently reliable ways to randomly select samples. One of the centerpieces of most of the methods was the random number table, which dates back to the early 1900s.[4] When used in conjunction with beginning and ending points, a table of random numbers allows research-ers to select enough random numbers to draw a desired sample. In a book chapter published in the early 1980s, Seymour Sudman reviewed the development of random numbers in sampling. In his discussion, he talks about the lengths undertaken to develop massive random number tables.

> The most convenient and most accurate procedure for obtaining a random process is through the use of tables of random numbers. The largest printed table is *A Million Random Digits* by the Rand Corporation (1955)…. The Rand random digits were generated by an electronic roulette wheel. A random-frequency pulse source passed through a five-binary counter and was then converted to a decimal number. The process continued for 2 months and even with careful tuning of the equipment the numbers produced at the end of the period began to show signs of nonrandomness indicating that the machine was running down.[5]

With the introduction of random number tables, the idea of ran-domly selecting numbers became much easier, however, particularly when needing larger samples, the process could become cumbersome and tedious. A variant of the simple random sample is the *systematic sample*,

which generally maintains the randomness of the selection process, but is easier to use when a very large sample is needed and is often easier to execute without mistakes when manually completed.

Systematic Sampling

Like simple random sampling, systematic sampling starts with a listing that is an identification of the units making up the population. The systematic approach usually starts by first determining the size of the sample needed. For illustration, let's say a researcher is conducting a survey and will require 200 individuals in the sample selected from a population of 2,000 individuals. A calculation is then done by dividing the sample size into the population to determine the interval between individuals to be drawn from the population to complete the sample. Every nth case is then selected, which, in the present example, would be every 10th individual. To inject randomness into the process, the key to the systematic selection is to randomly select a starting point number. Again, random number tables become a handy tool to find this starting point. So if we selected 7 as our starting point using the random number tables, we would draw the 7th individual on the list, the 17th, 27th, 37th, and so forth. In some cases, the systematic selection even allows us to select the cases without numbering them. For example, suppose a medical facility wanted a survey of patients who had been diagnosed with cancer in the past two years.[6] The medical facility is interested in this because it had installed a new imaging scanner a year earlier and it wanted to determine patients' perceptions of the diagnostic process before and after the installation of the new equipment. If the medical researchers used a systematic sample with the same parameters previously mentioned, that is, 2,000 total cases and a sample of 200 patients, they could simply identify the starting point, say the 7th patient in the case records, then select every 10th patient for the survey. Using the case records, the researcher could pull up the case file of every 10th patient after the 7th, and contact that individual about participating in the survey. However, a note of caution should be voiced here about the systematic sample selection process. If the dataset for the population is already sorted by some characteristics (sometimes unknown to the researcher), it can seriously bias the sample and wreak havoc on the study. Suppose in our example that the medical case records

had not been arranged alphabetically by the patients' last names but by the date of their diagnosis (this would be highly correlated with whether they were diagnosed with the old equipment or the new imaging scanner). Can you see a selection problem here? The influence of the present filing system on type of cancer would bias the sample and have serious implications for a survey dealing with diagnostic procedures.

Before moving on to more complex ways of selecting samples, a couple of points are worth noting. First, while some statistics and survey books still contain random number tables, today's students and researchers have a much easier time using random numbers to select samples because many of today's statistical analysis programs incorporate random number generators (subroutines) capable of generating random numbers for simple random and systematic sampling and even more sophisticated sampling designs. However, random number tables are still useful tools and a number of statistics and methodology texts continue to include them and provide instructions on their use.[7]

Second, despite the power of the computer to process samples, a number of problems may still make it impractical or impossible to carry out the process needed for the simple random or systematic selection of individuals from a population. Fortunately, statisticians and survey methodologists have developed a number of sampling designs that allow us to overcome many of these problems and still maintain the integrity of the probability sample. In the following section, we briefly review some of these approaches.

Stratified Sampling

Let's suppose that a large software company based in California with an extensive internship program is interested in how interns of various racial and ethnic backgrounds perceive the value of their internship with the company, such as whether interns believe they have been treated well, and whether they would be interested in seeking full-time employment with the company based on their internship experience. There are 3,060 interns currently with the company, and the company's human resources department estimates that it will have time and resources to conduct interviews with about 306 of them. The breakdown of the race and ethnicity of the interns is as follows:

- 1,200 Caucasians (39.2 percent)
- 660 Chinese/Southeast Asian (21.6 percent)
- 540 East Indian (17.6 percent)
- 240 Latino/Latina or other Hispanic (7.8 percent)
- 180 African/African American (5.9 percent)
- 120 Middle Eastern (3.9 percent)
- 120 Other ethnic groups (3.9 percent)

If we assume that the individuals of various racial and ethnic backgrounds would appear in a random sample in the same proportions as they appear in the population, then we would expect to see the following racial and ethnic distribution in our sample of 306 interns:

- 120 Caucasians (39.2 percent)
- 66 Chinese (21.6 percent)
- 54 East Indian (17.6 percent)
- 24 Latino/Latina or other Hispanic (7.8 percent)
- 18 African/African American (5.9 percent)
- 12 Middle Eastern (3.9 percent)
- 12 Other ethnic groups (3.9 percent)

However, because of the small proportion of certain racial and ethnic backgrounds, it is quite possible that some of the groups might have very few or no interns selected in a simple random sample due to sampling error. This is particularly problematic in our illustration because the researchers are primarily concerned with perceptions by the different racial and ethnic groups. To overcome this problem, we can use a technique termed *stratified sampling*, which works particularly well when we have subgroups within the population that are of very different sizes or small proportions of the population. By dividing the population into homogenous groups or layers called *strata*, then sampling within those strata, we reduce sampling error. In this example we would have seven strata or groups. Once we have the stratum identified, we can then use simple random sampling to select individuals within each stratum.

There are actually two types of stratified samples, proportional and disproportional. In proportional stratified random sampling, the size

of each stratum is proportionate to the population size of the strata. This means that each stratum has the same sampling fraction. In our illustration, there are 180 African American interns and 120 Middle Eastern interns, which are 6 and 4 percent of the total number of interns, respectively, so if our stratified sample is proportional, we would randomly select 18 interns from the 180 African American intern group and 12 interns from the Middle Eastern intern group. On the other hand, if we use a disproportionate stratified sampling method, the number of individuals from each stratum is not proportional to their representation in the total population. Population elements are not given an equal chance to be included in the sample (recall the previous EPSEM discussion). Therefore, while this allows us to *build up* or oversample the individual numbers in each stratum, which otherwise would have low numbers of individuals, it creates a problem if we're trying to generalize back to a population. Suppose in our example we sample disproportionately so that we have approximately 44 interns in each sample. In that case, the responses given by Latino/Latina/Hispanic, African American, Middle Eastern, and our Other category of interns would be overrepresented in the responses, while the responses of Caucasians and Chinese/Southeast Asian and East Indian interns would be underrepresented. To compensate for this, we would need to weight our stratum responses back to the proportions of the strata seen in the populations.

In our example here, we would likely be interested in comparing our strata or conducting what is termed a *between-stratum analysis*.* This would permit us to compare responses on the survey interviews from each of our seven strata against one another.

Cluster Sampling

Another form of sampling that also uses grouping of individuals in the process is called *cluster sampling*. Because both stratified sampling and cluster sampling use groups in their process, they are frequently confused.

---

* To do so, we would use a *balanced allocation* (also termed as factorial allocation), so we would select strata with an equal number of interns in each. Since we have limited our study to 306 individuals and we have seven strata, we would disproportionately sample so we had 44 interns in each stratum.

Recall that a stratified sample begins by placing individuals into groups based on some characteristic such as race and ethnicity, marital status, religious preference, and so forth. In cluster sampling, we begin by first randomly selecting a sample of some naturally occurring or known grouping. For example, we might create a cluster sample by randomly selecting a group of outlet malls. We then take all units from each of our randomly selected clusters for our sample. Thus, we might select all the stores from our randomly selected group of outlet malls. This approach is particularly useful when there is no satisfactory list of individuals or things in a larger population that we want to study, and no way to get at the population directly making it impossible to draw a simple random sample. To illustrate the process, let's consider using this approach to solve the problem of an inability to get a listing of individuals in a population. Suppose the National Collegiate Athletic Associate (NCAA), responding to growing concern with the rising problem of its athletes getting concussions while playing, decides to survey NCAA school athletes. The NCAA thinks that a survey of players would be good to determine how aware players were of the problem of concussions, if they had ever suffered a concussion, and if they had suffered any longer-term effect from a competition-related head injury. Because of the large number of college sports and players, the NCAA decides to start by first conducting surveys of athletes in two sports with higher probabilities of concussions: football and soccer. It contracts with a university to design and conduct two surveys, one for each sport. The league tells the university that it is very difficult to get a single listing of all current players across NCAA football and soccer players from which to pull a simple random sample. This is not an uncommon problem even with well-defined populations such as college sports teams; so can you imagine then the struggle to identify all the members of a less well-defined group such as aerospace workers or the residents of a particular country! Because of this problem, the researchers decide to use *cluster sampling*. Just as with stratified sampling, every member of the population can be a member of one, and only one, group or cluster—in this case one NCAA college or university. The first step is to identify known or accessible clusters, so in our example, the researchers will start by listing NCAA schools (because they are identifiable) across the country, then using a random selection method, they will choose a certain

number of schools that are the clusters from which individual athletes will be drawn. Basically, the researchers would ask for the team rosters from each of the selected schools for each of the two sports in order to produce its final two samples of athletes who will receive a survey.

We can extend the basic ideas of clustering, stratification, and random selection to create much more complex designs to deal with specific issues that might present themselves in sampling. Such designs are commonly referred to as multistage sampling. With *multistage sampling*, we select a sample by using combinations of different sampling methods. For example, because of the large number of student athletes on NCAA college football and soccer teams, the researchers may decide that it's too costly to send a survey to every athlete at schools in the cluster samples. They might then propose a two-stage sampling process. In Stage 1, for example, they might use cluster sampling to choose clusters from the NCAA college and university population. Then, in Stage 2, they might use simple random sampling to select a subset of students from the rosters of each of the chosen cluster schools for the final sample.

As long as every individual (in this case players) can be attached to one of the groups, this sampling approach works very well.[8] As you can see, it is quite easy to include additional clustering criteria in the design.

How Do We Select the Right Sample Size?

As Arlene Fink points out, "The ideal sample is a miniature version of the population."[9] For those pursuing this ideal, there is always a tradeoff in considering sample sizes—what is optimal and what is practical? In sampling methodology terms, the researcher must decide how much sampling error he or she is willing to tolerate, balanced against budget, time, and effort constraints.

When talking about sample sizes, it is important to keep in mind that sample size is based on the number of individuals who respond to the survey, not the number of individuals who initially receive the survey. Thus, the *response rate* needs to be taken into consideration when the sample is being selected. For example, if you have a one in 15 response rate, it means that for a sample of 300 individuals you need to start with an initial sample group of 4,500. If your response rate is one in 20, your initial

sample would need to be 6,000 individuals, so you can see how changing the response rate by only a small amount can have a fairly large impact on the number of individuals you need in your initial sample.

Before continuing to explore sample size determination, let's revisit some fundamental ideas about drawing inferences about a population from a sample. Recall that the only way we can be absolutely certain we have accurately captured a population attribute is to determine the value of the characteristic from every individual or thing in the population. So if we want to determine the average age of an alumni group and be 100 percent certain we are correct, we would need to ask each person's age, then sum up the responses, and divide by the number of individuals, which would give us an arithmetic average (mean). For sake of illustration, let's say that we have 15,000 alumni who range in age from 19 to 73 years and the average age (mean) of the population is 34 years. But, as is the case with most surveys in the real world, we have neither time nor resources to get the information from 15,000 alumni, so we'll draw a sample and use the average age of the sample to estimate the average age of the population. For a little drama, suppose we only ask one person's age, and that person's age is 26 years. This would be a point estimate of our population average age. Simply put, if we use that one person *sample* to estimate the average population of the entire alumni group we will be wrong by quite a bit! In fact, if we repeated our rather absurd sampling of one individual many times over, the only time our sample estimate of the population parameter would be correct is when our sample picked an individual who was exactly 34 years old. This would be a fairly rare event given there is likely to be a very small percent of the alumni population exactly 34 years old. So our sampling error will be very high. Now, suppose that we selected five alumni and determined their average age. You would still have a high likelihood of being off (i.e., have a lot of sampling error) from the population's true mean of 34, but the chances of getting an accurate estimate would slightly improve over the approach where you only asked one person. This is due to the number of different combinations of ages among five individuals. In other words, there is less sampling error with large samples. Stated another way, our sample of five individuals more closely resembles the distribution of age in the population than a sample of one. By contrast, if any of the following three different samples

of five individuals were pulled from the population, each would produce an accurate estimate of the true average age of the alumni population:

Group I: 20, 30, 30, 40, and 50
Group 2: 25, 28, 28, 32, and 57
Group 3: 30, 32, 34, 36, and 38

You may have also noticed that if you had pulled the same 15 individuals as repeated single person samples, you would have seen an accurate estimation of the population's average age only once while 14 of your samples would have yielded an incorrect estimate. You can see how as you increase your sample size, you increase the probability of accurately mirroring or representing the larger population.

Confidence Intervals and Confidence Levels

 If you went to a statistician to ask for help in determining the best sample size for a study you had planned, chances are the statistician would ask you several questions relative to sampling error and the precision of the sampling estimates you would like to have. The two primary questions in this regard would be "What confidence interval would you like to use?" and "What confidence level are you comfortable with?" A *confidence interval* is the range of values around the true population value within which the estimates from our samples are expected to fall a specific percentage of the time. It is also commonly called the *margin of error.* That is why when survey results are reported, they are usually accompanied by a disclaimer about the margin of error (recall the example of the president's approval earlier in the chapter). The statistician would also ask about the *confidence level* to describe uncertainty associated with the interval estimate. For example, the confidence level might be the *95 percent confidence level.* This means that if we used the same sampling method we could potentially create an infinite number of different samples and compute an interval estimate for each sample. We would expect the true population parameter to fall within the interval estimates 95 percent of the time. In survey research, confidence levels are commonly set at the 90th, 95th, or 99th percentile. In the first case, this would mean that you can expect your sample to contain the

true mean value, such as average age, within your confidence interval 90 percent of time. Conversely, 10 percent of the time your sample interval estimates would not contain the true mean. With the 95th percentile, you would reduce the probability of being wrong to 5 percent, and at the 99th you would reduce it even further to 1 percent.* The effect of increasing the confidence level is that the confidence interval becomes wider if the sample size stays constant. One way to counteract that is to increase the size of your sample. While the actual calculations for establishing confidence intervals are outside the scope of the discussion here, many good introductory statistics books will take you through the process. There are also a number of sampling calculators available online, which, once you provide the size of the population, confidence level, and confidence interval parameters, will give the sample size you will need for your survey.† These sample size calculators also will let you see how adjusting the confidence interval and confidence level affects your sample size.

One common mistake that people who are not familiar with survey sampling make is assuming that the size of samples must be proportional to the size of the populations. That is, while small samples work well for small populations, very large samples must be obtained for very large populations. However, the following illustrates how as the population grows, we can still obtain reliable estimates with samples that are modestly bigger. Let's say we are conducting a survey to determine whether voters approve of the president's performance. For sake of simplicity, let's say we start with the assumption that half will approve and half will not (which would result in the largest sample size because this is the maximum variation possible). If we use a standard 5 percent margin of error and a confidence level of 95 percent (both of which are commonly used), our sample sizes would vary across different population sizes as follows:

---

* You may have heard the term *significance level* when we are reporting results of statistical tests for differences between individuals or groups. The significant level is set by the confidence level, so when a researcher reports that there were significant differences between the average ages of the two groups at the $p \leq 0.05$ level, what the person is really saying is that the apparent differences in average age between the groups could be due to chance 5 percent of the time or less.

† See for example, online sample size calculators provided by Creative Research Systems at http://www.surveysystem.com/sscalc.htm, or the one offered by Raosoft Inc. at http://www.raosoft.com/samplesize.html

Size of population	Size of sample needed
200	132
2,000	323
20,000	377
200,000	383
2,000,000	385
200,000,000	385

With a confidence interval of ±5 percent at the 95 percent confidence level.

As you can see, initially our sample size will need to increase fairly significantly as the population gets bigger, however, after a point, the increase in the sample size essentially flattens out, as the population increases by a factor of 10 and finally by a factor of 100.

This is the beauty of probability sampling; if done correctly, we can use rather small representative groups or samples to accurately estimate characteristics of large populations.

Before turning our attention to survey error in the next chapter, a final note on probability sampling is in order. As you may have noted, we have not explored nonprobability sampling methods in this chapter. We chose to omit this because of space limitations, the unevenness of some of the emerging approaches, and the fact that probability sampling has been a tested framework since its inception in the 1930s, chiefly because of its power in generalizing back to a larger population. Recently, however, there has been increased concern about some of the problems facing probability-based sampling designs as rapidly evolving technology alters the landscape of communication essential for conducting surveys. Landline phones are disappearing as cell-phone-only households become commonplace and electronic communication such as e-mail and text messaging is substituted for hard copy. New messaging platforms that integrate rapid transmission of written material, video and still images, and audio are emerging on an almost daily basis. Recent development of address-based sampling frames has allowed researchers the ability to use probability sampling of addresses from a database with near universal coverage of residential homes. The web has become an easily accessible and inexpensive tool for survey delivery, even though a large number of web applications use nonprobability sampling methods, such as survey panels, and therefore are suspect in terms

of generalizing back to a larger population of interest. With these new technologies come sampling problems that affect the reliability and validity of probability sampling methods when they are simply layered over designs created around different data collection methods. The creation of new platforms for survey delivery requires an examination of alternative approaches.[10] Recently the American Association of Public Opinion Research (AAPOR), one of the most respected professional organizations in this area, commissioned a task force to "examine the conditions under which various survey designs that do not use probability samples might still be useful for making inferences to a larger population."[11]

The massive amount of data collected through Internet enterprise has even offered what some see as the elimination of the need to do surveys at all. The ability to collect, store, and analyze so called *Big Data* clearly offers opportunities to look at the relationships between variables (topics of interest) on a scale of populations rather than samples. In so doing, many of the concerns about sampling and sampling error presumably fall away. The use of Big Data also shifts much of the focus in analytic process away from concentrated statistical efforts after data collection is complete to approaches centered around collecting, organizing, and mining of information, "… the fundamental challenge in every Big Data analysis project: collecting the data and setting it up for analysis. The analysis step itself is easy; pre-analysis is the tricky part."[12] However, critics such as Tim Hartford point out that in the rush to use, and sometimes sell, big data approaches, proponents sometimes present "optimistic oversimplifications."[13] Hartford is particularly critical of the notion that theory-free data correlations obtained from Big Data can tell us what we need to know without the need to examine causality further and that those large data sets somehow remove all the analytic pitfalls that are seen in smaller data sets.

Clearly the changing landscape of electronic data collection will impact sampling methodologies used in surveys and create new approaches to get information about a population. Online surveys, opt-in panels, and the use of Big Data techniques serve as examples of the drive to capitalize on electronic data collection and storage capabilities offered through the Internet. However, the technology alone does not automatically improve our ability to make valid and reliable inferences about a larger population.

Summary

Note: In this summary we use the term individuals broadly to refer to the units or elements, which can be people or other individual entities, making up the population from which we draw samples.

- To collect information from a target population, well-done sampling can replace collecting information from each individual in that population (a process called enumeration). Sampling is more practical because it reduces time, effort, and costs needed to gather information.
- There are two basic types of sampling—probability and non-probability.
  ° Nonprobability may be used in some cases, but it has a major limitation; it doesn't allow us to judge how well our sample reflects the larger population about which we are trying to draw inferences. Essentially, we cannot determine sampling error with nonprobability samples.
  ° While probability sampling doesn't eliminate the possibility of picking a sample of individuals that doesn't accurately reflect the larger population, it does allow us to calculate how much error might be present.
- Probability samples have three fundamental elements:
  ° A group or frame of all individuals in the population can be created. This is termed the sampling frame.
  ° Each individual in the population has a positive chance of being selected into the sample.
  ° The probability of an individual being selected can be computed for each individual in the population.
- Simple random sampling is a basic type of probability sampling, which uses techniques to randomly choose individuals from the population for the sample. Each individual in the larger population has an equal chance of being selected for the sample. We call the methods that have this characteristic as Equal Probability of Selection Methods (EPSEM).

- Systematic samples are also EPSEM samples and are very similar to simple random samples, but differ in their approach of selecting the sample. Systematic samples use a process where the sample is formed by dividing the number of individuals needed for the sample into the number of individuals in the population to determine an interval between individuals for selection purposes. A random number located within the first interval is selected as a starting point, and then every subsequent nth individual is selected until the number of individuals needed for the sample is reached.

- Stratified sampling is a more complex sampling strategy that works well when there are subgroups within the population that are of very different sizes or are very small proportions of the larger population. Strata are formed by dividing the population into homogenous groups or layers, and then sampling is done within those strata. This reduces sampling error. We might, for example, create strata based on the racial and ethnic backgrounds of interns in a large company in order to ensure that certain racial/ethnic subgroups are not missed because they make up such a very small part of the intern population.

- Cluster sampling is another form of more complex sampling, which also uses a grouping technique like stratified sampling. With cluster sampling, however, the groups are formed by using some population (known or naturally occurring) characteristic like high schools or organizations such as businesses.

- Multistage sampling extends the basic ideas of stratification, clustering, and random selection to create much more complex designs to deal with specific issues that might present themselves in sampling. When engaging in multistage sampling, we simply break our sampling design down into separate stages, sequencing one method after another.

- Selecting the right size for a sample is a bit of an art and a bit of a science. There is always a tradeoff in considering sample sizes—what is optimal and what is practical.

- ° It is important to keep in mind that the size of the sample refers to the number of individuals that respond to the survey, not the number who initially receive the survey.
- ° Response rate refers to the proportion of individuals who respond out of the number of individuals that are initially selected and asked to participate. For example, a one in 15 response rate means that one out of every 15 individuals asked to participate actually completed a survey.
- ° Confidence intervals, also called the margin of error, refer to the range of values around the true population value within which our samples are expected to fall a specific percentage of the time.
  - ◊ Confidence levels reflect the amount of time that we can expect the values (estimates) derived from our samples to be in error. In social science research, confidence levels are typically set at the 90th (10 percent error), 95th (5 percent error), or 99th (1 percent error) percentiles.
  - ◊ Increasing the confidence level without increasing the sample size widens the confidence interval.
- One of the common mistakes that people not familiar with surveys make is assuming that the size of samples must be proportional to the size of the population. In reality, after population sizes reach a certain level, the sample size only needs to increase a small amount even if the population increases by magnitudes of 10 or 100 or more.
- This is the beauty of probability sampling; if done correctly, we can use rather small representative groups to accurately estimate characteristics of large populations.

Annotated Bibliography

- There are a number of good books and articles on research sampling, from very technical presentations to general introductory discussions.
- A good general introduction to sampling can be found in Earl Babbie's classic *The Practice of Social Research*.[14]

- *The Handbook of Survey Research* edited by Peter H. Rossi, James D. Wright, and Andy B. Anderson provides a comprehensive look at survey methodology and data. Chapter 2 on Sampling Theory by Martin Frankel[15] provides the underlying foundation for survey sampling and Chapter 5 by Seymour Sudman[16] covers the theory of sampling and different sampling approaches.

- If you haven't had statistics or are new to the idea of sampling, some of the self-help websites such as Stat Trek provide overviews to survey methodology on topics such as sampling.[17]

- For insight into more advanced sampling techniques such as *proportionate stratified sampling* and more complex *disproportionate stratification* methods such as disproportionate optimum allocation, see Daniel's *Essentials of Sampling*.[18]

- Our Chapter 2 focuses on probability sampling. Other nonprobability types of sampling are discussed in most statistics or survey research methodology texts. Advances in technology are resulting in rapid change in survey sampling and data collection, such as the use of online survey panels. These methods have come under increasing scrutiny because of questions surrounding their ability to produce representative samples. Information on a recent examination of these by the American Association for Public Opinion Research is available.[19] A nice review of recent advances and how methodology can be improved is provided by Engel and his associates.[20]

CHAPTER 3

Total Survey Error

Error is inevitable and occurs in any survey. If you need perfection, don't bother doing research. What is important is to identify the possible sources of error and then try to minimize these errors.

Herbert Weisberg defines error as the "difference between an obtained value and the true value."[1] Typically we don't know what the true value is but that doesn't change our definition of error. When we discussed sampling in the previous chapter, it was the population value. When we focus on measurement, it is the true or actual value of whatever is being measured. Error is the difference between that true value and whatever the obtained or observed value turns out to be. It's important to keep in mind that error can occur at any point in the process of doing a survey from the initial design of the survey through the writing of the report.

Weisberg points out that error can be *random* or *systematic*.[2] For example, when we select a sample from a population, there will be sampling error. No sample is a perfect representation of the population. Assuming we are using probability sampling, this error will be random. However, sometimes some elements in the population are systematically left out of the sample. For example, if we are doing a phone survey and rely exclusively on landlines that could produce a systematic error because we have left out cell-phone-only households. Systematic error is often referred to as *bias*. We need to be aware of both random and systematic error.

There are many types of error that can occur. In the past, the focus was on sampling error and nonresponse error, which occurred as a result of refusals or the inability to contact respondents. Instead of focusing on just a couple of types of error, we should focus on all possible types of survey error. This is often referred to as total survey error.[3] Paul Biemer defines *total survey error* as the "accumulation of all errors that may arise in the design, collection, processing and analysis of survey data."[4]

There are various ways of categorizing the different types of survey error. Typically we consider the following types of error:[5]

- Sampling error
- Coverage error
- Nonresponse error
- Measurement error

Weisberg also discusses survey administration issues such as the following:

- *Mode effects,* which refers to the fact that different modes of survey delivery such as telephone, face-to-face, mailed, and web surveys sometimes produce different results; and
- *Postsurvey error,* which occurs during the processing and analysis of data.[6]

To this we would add error that occurs in the reporting of survey data.

We're going to look at each of these types of error, discuss some of the research findings about each type, and talk about how you can try to minimize error.

Sampling Error

As discussed in Chapter 2, *sampling error* is one of the issues in sample design and occurs whenever you select a sample from a population. No sample is ever a perfect picture of the population. Let's say that your population is all households in the city in which you live. You select a sample of 500 households from this population.* You're interested in the proportion of households that recycle such things as cans, bottles, and other recyclable materials. It turns out that 45 percent of the sample recycles.

---

* We discussed sampling in Chapter 2 so we're not going to revisit the details of sampling here. You might want to look back at Chapter 2.

That doesn't mean that 45 percent of the population recycles. Why? Sampling always carries with it some amount of sampling error. It's inevitable.

Here's another way to understand sampling error. We can use sample data to estimate population values. If you were to select repeated random samples of the same size from the same population, your sample estimates would vary from sample to sample. If you think about it, this makes sense. Each sample will contain a different set of households. So why would you expect all the samples to give you the same estimate of the households that recycle?

One of the advantages of *probability sampling* is that you can estimate the amount of sampling error there will be from sample to sample. Assuming that you used probability sampling to get your sample and that you properly selected your sample, the resulting sampling error will be random. And to make things even better, there are things you can do to reduce sampling error.

Minimizing Sampling Error

Here are two ways you can reduce sampling error.

- Increase the size of your sample. Of course, there are practical limits to how big a sample you choose. You're limited by the cost and time it will take to collect the data. If you can decide how much sampling error you're willing to tolerate, you can determine the size of the sample that you will need.
- You can also stratify your sample to reduce sampling error. With stratified sampling you start by dividing your population into homogenous groups such as males and females. Then you sample from each of these groups. Often you choose your sample such that the sample has the same proportion of males and females as does your population. If you stratify your sample by some variable that is related to what you want to estimate, then this will reduce sampling error.[7]

Coverage Error

Earl Babbie distinguishes between the *population* and the *study population*. The population is the "theoretically specified aggregation of the elements in a study" while the study population is the "aggregation of elements from which a sample is actually selected."[8] In other words, the population you want to make statements about can be different from the study population from which you draw your sample. The *sampling frame* is the actual list from which the sample is selected.[9]

Coverage error occurs when the sampling frame does not match the population. In other words, sometimes the list from which the sample is selected does not match the population and this produces coverage error. For example, some elements in the population may have been left out of the list from which the sample is selected.* Let's look at some examples.

- The university wants to know how students feel about raising student fees to partially fund a new student center. The population is all students registered at the university during the current semester (or quarter). The list from which the sample is drawn is the registrar's official student roster. In this case, the list from which the sample is drawn almost perfectly matches the population. The only coverage error would be a result of errors in the registrar's list.
- Our research group has a contract to do a consumer attitudes survey in your county. We want to know how people feel about consumer spending, investments, borrowing, and savings. The population is all adults (18 years and over) living in your county at the time of the survey. We decide to do a telephone survey but we're not sure how to select our sample. Here are some possibilities that have been suggested.
  - ° One member of the team suggests that we draw our sample from all individuals listed in the phone directory published by the telephone company. However, this is quickly rejected

---

* Another problem occurs when elements that are not part of the population are included in the sampling frame. Sometimes this can be dealt with by screening. For example, in a phone survey some phone numbers that are outside the geographical area you want to cover might be included in your sampling frame. If you are aware of this possibility, you could include a screening question in which you ask if the household is in the desired geographical area.

when another member points out that this would systematically omit all people with unlisted numbers and those with only cell phones. That would create coverage error since we would be systematically omitting a large proportion of our population and people with listed numbers are systematically different from those with unlisted numbers and with only cell phones.

° Another team member suggests using a random-digit dialing approach in which residential prefixes of landlines in your county are sampled and then random digits are added to these prefixes to produce the list of phone numbers which we would call.[10] This is also rejected when someone points out that while we would be including those with unlisted landlines we would be omitting households which have only cell phones or no landline.

° Then someone tells us that the U.S. Postal Service has a list of residential addresses which is available through commercial providers and which might work for us. This is referred to as address-based sampling. Shelley Roth and associates suggest that this provides "nearly complete coverage of residential addresses in the United States."[11] David McNabb notes that there are some coverage issues that you might encounter when using this approach. The list tends to undercover rural areas and groups such as college students living in dorms. New homes are constantly being built and some homes may be destroyed by fire and natural disasters.[12] So there will still be coverage error but it would be considerably less than in the first two options.[13]

• The General Social Survey (GSS) is a large national probability survey that began in 1972 and is now conducted biannually by the National Opinion Research Center at the University of Chicago.[14] The population is all adults (18 years and over) residing in households in the United States as of a particular date. The sample was originally drawn from all adults who speak English and live in noninstitutionalized settings. In 2006, Spanish-speaking individuals were added

to the sample. That means that individuals living in institutionalized settings are systematically excluded and prior to 2006 non-English speaking individuals were excluded. From 2006 onwards those who didn't speak English or Spanish were excluded. If those who are excluded are a small part of the population, this will probably introduce a small amount of coverage error into the design. Cost considerations may compel the researcher to live with this small amount of bias in order to reduce costs.

• Let's say you want to do a survey of churches in your county. You want to find out why some church members are more active in their church than other members. First, you have to compile a list of all churches in your county. You're surprised to find out that such a list is not immediately available but with some work you assemble the list. Now you select a sample of churches to which you plan to send your survey. You contact the churches in your sample and ask them to send you their membership lists so you can select a sample of members from each church in the sample.* Some churches are not willing to send you their membership list but most offer to send the list on the condition that you do not share it with anyone other than the project staff. However, many of the churches tell you that their membership list is out of date. After more discussion, you find out that there are several problems.

  ° Some members have moved away but are still on the membership list.

  ° Not all new members have been added to the list.

  ° It's possible that some members appear twice on the list.

  You realize this is going to produce coverage error. The best solution is to work with each church in your sample to delete members who are no longer there, add in the new members, and take out the duplicates. It takes some work but it's worth it because it reduces coverage error.

---

* This is often referred to as a multistage cluster sample.

Minimizing Coverage Error

So how do we try to reduce coverage error? First we have to ask the right questions. Don Dillman and his associates suggest that there are five questions that we ought to always consider.[15]

- "Does the list contain everyone in the sample population?"
- "Does the list include names of people who are not in the study population?"
- "Are the same sample units included in the list more than once?"
- "Does the list contain other information that can be used to improve the survey?" This could include information such as phone numbers and e-mail addresses, which could be used to follow-up those who don't respond to our survey.
- "What to do when no list is available?" Dillman and his associates use the example of visitors to a national park. In cases like this, we might sample people as they enter or leave the park.[16]

So the general strategy for dealing with coverage error is to first identify the sources of error. Once we know what the problems are, then we can try to reduce them keeping in mind that eliminating all coverage error is probably not possible. This can be done in several ways.

- We can try to make the list from which we draw our sample more complete by taking out the elements that shouldn't be in the list, adding in the missing elements, and deleting duplicates. Using the example of sampling church members discussed earlier in this chapter, we could work with the staff of the churches to bring their membership lists up to date.
- We can look to see if there are other lists that we can use to improve coverage. For example, churches might have lists of new members or members who have transferred out of the church even though they haven't updated their membership lists.
- Even if we aren't able to completely eliminate all coverage error, we can at least be aware of the error that exists and take

this into account when we write our report. We need to be careful to limit the degree to which we generalize our results. For example, with the GSS discussed previously, we should be careful to only generalize to adults living in noninstitutional-ized settings who speak English or Spanish.

Nonresponse Error

Ideally, we want everyone in our sample to complete the survey but we know that probably isn't possible for two reasons.

- We probably won't be able to contact every person in our sample. For example, in a phone survey, some people are seldom at home or use caller ID to screen their calls and only answer when they know who is calling.
- Some of the people who we do contact will refuse to do the survey. *Refusals* can occur in two ways.
  ° People might completely refuse our request to do the sur-vey. In other words, they don't answer any of our questions. This is sometimes referred to as unit nonresponse. We'll discuss this next.
  ° Other people consent to being interviewed but refuse to answer certain questions such as family income or race. This is often referred to as item nonresponse since they are refusing to answer particular questions or items in our survey.

Theories of Survey Participation

It helps to think about the different reasons that people might agree or refuse to be interviewed. These are often referred to as theories of participation.

- *Social exchange theory*. This approach looks at interviewing as a type of social exchange. Dillman suggests that "people engage in a social exchange with others when the perceived

rewards outweigh the perceived costs."[17] *Costs* include such things as the time it takes to complete the survey or the amount of energy required to do the survey. Individuals also receive *rewards* from doing the survey such as monetary incentives they might be given or the satisfaction of helping. Another important factor affecting participation is the *trust* respondents have that completing the survey will "provide a valued reward in the future."[18] From the perspective of social exchange theory, participation can be encouraged by reducing the costs associated with the survey, increasing the rewards from survey participation, and ensuring the trust of the respondent that rewards will be forthcoming.

- *Leverage-salience theory*. Robert Groves and his associates developed the leverage-salience theory of survey participation. Groves outlines this theory.

Under the theory, different persons place different importance on features of the survey request (e.g., the topic of the survey, how long the interview might take, the sponsor of the survey, what the data will be used for). Some persons may positively value some attributes, others negatively. Of course, these differences are generally unknown to the survey researcher. When the survey approach is made to the sample person, one or more of these attributes are made salient in the interaction with the interviewer or the survey materials provided to the sample person. Depending on what is made salient and how much the person negatively or positively values the attribute, the result could be a refusal or an acceptance.[19]

In other words, different things are important to different people. Some place importance on the length of the survey while others focus more on the topic or incentives. Groves refers to this as *leverage*. Researchers place emphasis on different aspects of the survey. Some emphasize the topic while others focus on the length particularly if it is a short survey. Groves refers to this as *salience*. This approach suggests that we should try to understand what is important to our respondents and emphasize those aspects of the survey. It also suggests that we ought not to focus

on only one aspect of the survey when contacting respondents but we should focus on different aspects which might be important to different respondents.

Nonresponse and Nonresponse Bias

It's clear that *nonresponse* has been increasing and that this is a critical problem for surveys. Roger Tourangeau and Thomas Plewes looked at a number of large national surveys conducted in the United States over the last 35 years and concluded that "nonresponse rates continue to increase in all types of cross-sectional surveys, with little to suggest that that the trend has plateaued."[20]

Edith de Leeuw and her associates focus on the difference in *response rates*[21] for different modes of survey delivery. They concluded that

> in general, face-to-face surveys tend to obtain higher response rates than comparable telephone surveys, and mail surveys tend to have a lower response rate than comparable face-to-face and in lesser degree to telephone surveys. In addition, the response rates for both telephone and face-to-face surveys are declining, although such a trend is not as evident for mail surveys.[22]

But why is nonresponse a critical problem for surveys? One reason is that nonresponse has become sizable and this can increase the risk of nonresponse bias. The other reason is that people who don't respond to surveys are often systematically different from those who do respond and this has the potential for creating bias in our survey data. If the difference between those who respond and don't respond is related to what the survey is about, then bias will occur.

Let's consider some examples of *nonresponses bias*. Andy Peytchev and his associates looked at self-reports of abortion and concluded that "those with a lower likelihood to participate in the survey were also more likely to underreport such experiences."[23] Many researchers have observed that voting in elections tends to be overreported. Roger Tourangeau and his associates note that "nonvoters were both less likely to take part in the survey and more likely to misreport if they did take part."[24] In both these

examples, those who took part in the survey were different from those who did not take part and this difference was related to the focus of the survey.

Another example of nonresponse bias is described in Thomas Holmes and James Schmitz's analysis of the "Characteristics of Business Owners Survey." Holmes and Schmitz focus on estimating "the probability that an individual discontinues ownership of his or her business."[25] Data are based on a sample of tax returns from 1982. The survey was mailed to respondents in 1986. We would expect that those who still owned their business in 1986 would be more likely to return the survey than those who did not currently own their business. Since we want to estimate the probability that a person has "terminated his or her ownership share over the 1982–1986 period"[26] we would expect that the data would underestimate the probability of termination and, in fact, Holmes and Schmitz's analysis shows this to be the case.

Increasing Response

If nonresponse bias is a problem, then what can we do about it? Increasing response is not a guarantee of low bias but a high nonresponse rate raises the possibility of nonresponse bias. Let's look at some ways in which we can increase response.

Groves suggests that there are five factors that affect *survey participation*.[27]

- Societal factors such as the frequency of surveys in the society and public opinion regarding the legitimacy and worth of surveys.
- The survey design itself including such factors as the length of the survey and how respondents are chosen.
- The respondents including such things as gender, income, and age.
- The interviewers including their experience and expectations regarding the interview.
- The interaction between the respondent and the interviewer.

We can't do much about some of these factors. For example, we can't do much about the increase in surveys in our society and the fact that

some people may have recently been asked to do a survey. We can't do much about the growing trend for people to express doubts about the worth of surveys. But we can do something about the survey itself. Dillman has written extensively about reducing the burden on respondents.[28] This is a logical consequence of social exchange theory. If we can reduce the costs of doing surveys, then we will increase the likelihood of people to respond. We can make the survey as easy to take as possible. We can create a survey that flows naturally from question to question. We can avoid asking unnecessary questions which will reduce the length of the survey.

There are psychological principles that we can use to try to increase survey participation. When we ask someone to agree to be interviewed we hope that they will comply with our request. Robert Cialdini suggests that there are certain rules of behavior that can be used to increase compliance.[29] Here are some of these rules summarized by Groves.

- The *rule of reciprocity* suggests that "one should be more willing to comply with a request to the extent that the compliance constitutes the repayment of a perceived gift, favor, or concession."[30] For example, if respondents ask how long the survey will take and we respond 20 minutes, they might refuse saying that's too long. If we then ask them to take a shorter version of the survey, perhaps five minutes, this might be seen as a concession and increase the likelihood of complying with our request.
- The *rule of social validation* suggests that "one should be more willing to comply with a request to the degree that one believes that similar others would comply with it."[31] If we tell the respondent that others have been willing to be interviewed and found it interesting, this might increase the likelihood that they will agree to be interviewed.
- The *rule of scarcity* suggests that "one should be more willing to comply with requests to secure opportunities that are scarce."[32] If we tell people that we are only contacting a small proportion of the population, this might increase the likelihood of their participation.

There is considerable evidence that offering a small prepaid cash *incentive* increases the likelihood of a person responding to the survey.[33] You have probably received a request for donations from nonprofits or political candidates. Often the request comes with a small gift such as a pencil, a key chain, or some other small gift. This is similar to the prepaid cash incentive. Incentives given before the individual responds to the survey have been shown to be more effective than postpaid incentives in increasing survey participation.[34] Offering the person the opportunity to be entered into a lucky draw for a large gift such as a computer tablet or cash does not appear to be as effective.

One of the most effective ways of increasing participation is multiple *follow-ups*. Dillman, talking about mailed surveys, says that "multiple contacts are essential for maximizing response."[35] The same thing can be said for any type of survey—face-to-face, mailed, phone, and web surveys. In face-to-face and telephone surveys, multiple contacts can add considerably to your cost but they are essential for increasing response rate.

Measurement Error

Measurement error is the difference between the true value of some variable and the value that the respondent gives you. A simple example is measuring age. Often we ask the respondent how old they were on their last birthday. But if you are young and you order an alcoholic drink in a bar, the bartender will ask you for proof of age. The age given to the bartender could easily be an overestimate of your age. This would be an example of "measurement error due to the respondent."[36] In other words, respondents might not be giving you an accurate answer because of their self-interest in appearing older. Weisberg contrasts this with "measurement error due to the interviewer."[37] We know that the interviewer's gender, race, and age can influence how respondents answer our questions. We're going to talk about both types of measurement error.

We'll start by discussing error that occurs as a result of question wording and question order. It's important to understand that measurement error, like all types of error, cannot be eliminated. But it can be minimized. Minimizing error is only possible if you are first aware of the ways in which error can occur and second take steps to minimize it.

Measurement Error Associated with Question Wording

Measurement error can occur as a result of *question wording*. One of the classic examples is found in Howard Schuman's and Stanley Presser's discussion of the difference between "forbidding" and "not allowing" certain types of behavior such as "making public speeches against democracy." They conclude that "Americans are much more willing to not allow speeches than they are to forbid them, although the two actions appear to be logically equivalent."[38] Numerous studies have replicated this finding. However, Schuman notes that regardless of which wording is used, there is a clear trend over time toward not forbidding or allowing such speeches. Thus, even with questions like the forbid and not allow questions, you can still track changes over time.[39]

Barbara Bickart and her associates studied the accuracy of reports of other people's behavior. She asked couples to "search for information about a vacation they could win."[40] Then they discussed and actually planned the vacation. Afterwards they "were asked to either count or estimate the number of accommodations, restaurants, and activities that they/their partner examined during the information search task."[41] Their analysis showed that questions asking for counts were more accurate than questions asking for estimations.

Still another example is found in a series of questions in the GSS conducted by the National Opinion Research Center. The GSS asks a series of questions about whether the United States should be spending more money, less money, or about the same amount on such things as welfare. They conducted an experiment by randomly asking one-half of the respondents about "welfare" while the other random half was asked about "assistance to the poor." Tom Smith analyzed GSS data and concluded that "'welfare' typically produces much more negative evaluations than 'the poor.'"[42] Gregory Huber and Celia Paris point out that this assumes that these two terms are equivalent to the respondent. Their research suggests that this is not the case. They conclude that "respondents are twice as likely… to believe that programs like soup kitchens, homeless shelters, and food banks are ATP [assistance to the poor] as opposed to welfare."[43] In other words, the questions are not equivalent because the words "welfare" and "assistance to the poor" bring to mind

different things. Huber and Paris's findings point out that we shouldn't be too quick to conclude that question wording is behind the different responses but that we need to look below the surface and consider how respondents interpret different question wording.

Another example is questions that ask for opinions about *global warming* and *climate change*. Jonathon Schuldt and his associates found that "Republicans were less likely to endorse that the phenomenon is real when it was referred to as 'global warming'… rather than 'climate change'… whereas Democrats were unaffected by question wording."[44] They point out that the difference between Republicans and Democrats is much greater when the question is framed in terms of *global warming*. Lorraine Whitmarsh looked at what respondents think these terms mean and discovered that global warming is more likely to be associated with human causes than climate change.[45] Again this suggests that respondents attach different meaning to these terms. Thus, it becomes critical how the question is worded when making comparisons between Republicans and Democrats.

Questions are often asked about people's attitudes toward abortion. Sometimes a single question is used for asking respondents to indicate their attitude toward abortion in general. For example, do you think abortion should be legal or not? However, the GSS includes a series of seven questions asking whether people think abortion should be legal in various scenarios—in the case of rape, in the case of a serious defect to the baby, in the case of a woman who has low income and can't afford more children, and other such situations. The data show that people are much more likely to feel abortion should be legal in the case of rape or a serious defect to the baby than they are in the case of low income women who can't afford more children. Howard Schuman offers the advice of asking "several different questions about any important issue."[46] The abortion example illustrates this point.

Still another example is found in asking about voting. You would think that whether you voted or who you voted for is pretty straight-forward but here again question wording makes a difference. Janet Box-Steffensmeier and her associates report on a change that was made in the American National Election Study's (NES) question on whether and how one voted in House of Representatives contests. Prior to 1978, there

was little difference between the actual House vote and the vote reported in the NES. Since 1978 the NES has reported a much higher vote for the incumbent than the actual vote. Box-Steffensmeier suggests that the following changes in question wording might account for this finding.

- Question used prior to 1978—"How about the vote for Congressman—that is, for the House of Representatives in Washington? Did you vote for a candidate in Congress? [if yes] Who did you vote for? Which party was that?"[47]
- In 1978 and afterwards, a ballot card was given to the respondent listing the candidates and their party and the following question was asked—"Here is a list of candidates for major races in this district. How about the election for House of Representatives in Washington? Did you vote for a candidate in the U.S. House of Representatives? [if yes] Who did you vote for?"[48]

Box-Steffensmeier concludes that "the ballot format evidently exaggerates the incumbent's support because people are far more likely to recognize… the incumbent's name than… the challenger's name."[49] This study also showed that you can reduce the pro-incumbent bias by making the candidates' party stand out by bolding and italicizing it and using a different font. This reduced but did not eliminate the bias.[50]

Measurement Error Associated with Question Order

It's clear that question wording affects what people tell us. *Question order* also makes a difference. Think about a survey in your community that deals with quality of life. One of the questions you might ask is "what is the most pressing problem facing your community today?" You might also want to ask more specific questions about crime, the public schools, and jobs. Would the order of the questions make a difference? If you asked about crime first, then respondents would probably be more likely to mention crime as one of the most pressing problems. Order matters.

David Moore provides us with some interesting examples of order effects using data from a Gallup Poll that was conducted in 1997. The question was "Do you generally think that [Bill Clinton/Al Gore] is honest and trustworthy?"[51] A random half of the respondents was asked the question with Clinton's name first and the other random half was asked with Gore's name first. The data show

> that when respondents (half the sample) were asked about Clinton first, 50 percent said he was honest and trustworthy; when the other half of the sample was asked about Gore first, 68% said the vice president was honest and trustworthy.[52]

In other words, Gore was considered honest and trustworthy by 18 percentage points more than Clinton. But when Moore took into account the order of the questions, he found that when Clinton's name appeared second 57 percent said he was honest and trustworthy and when Gore's name appeared second 60 percent saw him as honest and trustworthy. The 18 percentage point difference is reduced to three percentage points. He concludes that "this is a classic case of people trying to make their ratings of the two men somewhat consistent" and he refers to this as a consistency effect.[53]

On the same poll, respondents were given the following question: "I'm going to read some personal characteristics and qualities. As I read each one, please tell me whether you think it applies to [Newt Gingrich/Bob Dole]…Honest and trustworthy."[54] Again the order of the names was randomly assigned with half the respondents receiving Gingrich's name first and the other half given Dole's name first. Dole was considered more honest and trustworthy by 19 percentage points when Gingrich's and Dole's names appeared first but that increased to 31 percentage points when their names appeared second. Moore calls this a contrast effect because the data show that when "when people think of Dole and Gingrich, they tend to emphasize the differences between the two men rather than the similarities."[55] This is not to say that the order of the questions always affects what people tell us. But we should be aware of this possibility. The examples provided by Moore show us how this might occur.[56]

Measurement Error Associated with Respondents' Characteristics

Satisficing

Answering questions often requires a lot of effort on the part of respondents. Charles Cannell and his associates suggest that respondents go through a process in trying to answer questions that looks like the following.[57]

- First, they have to understand what the question means.
- Then they have to process the information that is necessary to answer the question. This involves determining what information they need, actually retrieving this information from their memory or records, and then organizing this material.
- Next they have to determine whether this information actually answers the interviewer's question as well as evaluating the information in terms of other things that are important to them such as their self-esteem.

In order to reduce the amount of effort required to answer survey questions, respondents sometimes look for ways to reduce this burden. This is called *satisficing* and can take various forms including:

- Answering don't know to questions;
- Skipping questions or saying they have no opinion;
- Choosing answers randomly; and
- Giving one-word answers to open-ended questions.[58]

For example, let's think about the quality-of-life survey that we mentioned earlier that asks "what is the most pressing problem facing your community today?" Some respondents might give you a one-word answer such as crime or education or jobs. This doesn't really tell us much about what respondents are thinking. Other respondents might say they don't know or that they have no opinion. Such answers reduce the work load of respondents.

Some survey questions give respondents a list of possible response categories from which they are asked to select their answer. Sometimes they are limited to one choice while other times they may select multiple

responses. Marta Galesic and her associates used eye-tracking information for a web survey to show that respondents often spend "more time looking at the first few options in a list of response options than those at the end of the list."[59] They also found that "the eye-tracking data reveal that respondents are reluctant to invest effort in reading definitions of survey concepts that are only a mouse click away or paying attention to initially hidden response options."[60] These are also examples of satisficing.

Jon Krosnick suggests that

> the likelihood that a given respondent will satisfice … is a function of three factors: the first is the inherent *difficulty of the task* that the respondent confronts; the second is the respondent's *ability* to perform the required task; and the third is the respondent's *motivation* to perform the task.[61]

Other researchers have suggested that satisficing occurs more frequently in certain types of surveys. Heerwegh and Loosveldt found that satisficing occurred more frequently in web surveys than in face-to-face surveys,[62] and Holbrook discovered that satisficing occurred more often in telephone surveys than in face-to-face surveys.[63] Krosnick and his associates also found that some respondents are more likely to satisfice than other respondents. For example, low-education respondents were more likely to say that they had no opinion than those with more education.[64]

Social Desirability

Some types of behavior or attitudes are viewed as more *socially desirable* than others. For example, voting is often seen as a responsibility of citizens and as a socially desirable action. On the other hand, cheating on exams is typically viewed as socially undesirable. There is considerable evidence that respondents tend to overreport socially desirable behaviors and attitudes and underreport those that are socially undesirable.

Brian Duff and his associates compared the actual voting turnout in the 2000 and 2002 elections with the turnout reported in the 2000 and 2002 American NES. In 2000, reported turnout exceeded actual turnout by 17.7 percentage points and in the 2002 election by 16.6 percentage points.[65]

Matthew Streb and his associates looked at a different question—whether people would vote for a woman for president if they thought she was qualified. Public opinion data show that the percent of people who say they would vote for a woman increased from slightly over 30 percent in 1945 to slightly over 90 percent in 2005.[66] Clearly the norms of equality and fairness suggest that one ought to be willing to vote for a woman who is qualified. Some people might be giving this answer because they see it as the socially desirable response.

Frauke Kreuter and her associates looked at reports of socially desirable and undesirable behaviors in a survey of university alumni. The types of behavior included dropping a class, receiving a D or F, receiving academic honors, belonging to the Alumni Association, and donating money to the university. Clearly receiving a D or F would be socially undesirable. Using university records, Kreuter found that approximately 61 percent of the respondents who answered this question had received such a grade. Of these respondents, approximately 27 percent failed to report receiving that grade.[67] Kreuter also found that underreporting of the socially undesirable response was less in web surveys than in telephone surveys.

Measurement Error Associated with the Interviewer

Characteristics of the interviewer could refer to physical characteristics such as race, sex, and age or to characteristics such as perceived friendliness. These characteristics can affect what respondents tell us. They can interact with respondent characteristics to produce different effects for males and females or for blacks and whites or for other categories of respondents. We're going to focus on two characteristics of interviewers that have been shown to affect what people tell us—race and sex.

Race of the Interviewer

Two classic studies dealt with questions about race in surveys conducted in Detroit in 1968 and 1971. Howard Schuman and Jean Converse showed that blacks appeared more militant and expressed more hostility toward whites when interviewed by blacks than when interviewed by whites.[68] Shirley Hatchett and Schuman found that whites gave more

"liberal or pro-Black opinions when the interviewer is Black."[69] Both of these studies interviewed respondents face-to-face where the race of both the interviewer and the respondent was generally apparent.

Other studies focused on voting. Barbara Anderson and her associates used five election surveys ranging in time from 1964 to 1984. Their data showed the following.[70]

> … Black nonvoters … who lived in predominately Black neighborhoods and were interviewed by Black interviewers were more likely to report falsely that they voted than Black respondents interviewed by White interviewers. Black respondents in Black neighborhoods who were interviewed by Black interviewers were also more likely *actually* to vote … than Blacks interviewed by Whites.

Steven Finkel and his associates used a 1989 survey in Virginia that looked at voting in a gubernatorial election in which Douglas Wilder who was black ran against Marshall Coleman who was white. Finkel found that "whites are 8–11 percentage points more likely to voice support for the Black candidate to Blacks than to Whites."[71]

Darren Davis and Brian Silver focused on political knowledge in a telephone survey of adults in Michigan. He considered both the actual race of the interviewer and the race perceived by the respondent. For whites, neither the actual race nor the perceived race of the interviewer was related to political knowledge. However, "when Black respondents identify the test-giver as Black, they do much better on the test than when they identify the test-giver as White or when the race of the interviewer is ambiguous."[72] This study is important because it explicitly measured the perceived race of the interviewer and showed perceived race to be an important variable. It also showed that race can be an important factor for some respondents but not for other respondents.

Sex of the Interviewer

Research has also shown that the sex of the interviewer can affect what people tell us. Emily Kane and Laura Macaulay analyzed data from

a national sample of households and found that "male respondents offer significantly different responses to male and female interviewers on questions dealing with gender inequality in employment."[73] Men voiced more equalitarian views to female interviewers than to male interviewers.

Other studies focused on health-related information. Timothy Johnson and Jennifer Parsons reported that the homeless (both male and female) are more likely to report substance abuse to male interviewers than to female interviewers.[74] However, Melvin Pollner found that both male and female respondents were more likely to report substance abuse to female interviewers than to male interviewers suggesting that gender affects respondents differently in various settings.[75]

These studies show that interviewer characteristics such as race and sex can influence what respondents tell us suggesting that we ought to consider the interviewers' race and sex as variables in our analysis of survey data. They also suggest that interviewers ought to be randomly assigned to respondents rather than trying to match the respondents' race and sex.[76]

Recognizing and Minimizing Measurement Error

- Some measurement error is associated with question wording and order.
  - One strategy is to embed an experiment into the survey. Identify two or three different ways to word the question and assign each version to a random half or third of the sample. This will allow you to determine if the different ways to word the question produce similar or different responses. (See the discussion of *global warming* versus *climate change* earlier in this chapter.) The same strategy can be used with question order.
  - Ask your respondents to describe what they think the question means or what they are thinking when they answer the question. George Bishop calls this asking respondents "to think out loud" about how they arrived at their answers.[77]

This can be used with a random part of your sample. Howard Schuman calls this a "random probe."[78]

° Ask people who are survey experts to review your questions and identify questions that might be problematic. Where question wording might be an issue, follow Schuman's advice and ask "several different questions" about that issue.[79] (See discussion of questions on abortion earlier in this chapter.)

• Other measurement error is associated with respondent behavior such as satisficing and social desirability.

° If satisficing is a result of the burden of answering questions, then it follows that reducing this burden might decrease satisficing. For example, instead of asking for the exact total family income in the previous year, we could give respondents a set of categories and ask them to place themselves in one of these categories. We can make sure that the interview is clearly worded and that it flows naturally from question to question. We can avoid asking unnecessary questions thus reducing the length of the survey.

° Research shows that question wording can reduce the tendency to give socially desirable responses. Duff reports that by "providing respondents with socially acceptable excuses for not voting, we [can reduce]… the over-reporting of turnout in the 2002 National Election Study by about 8 percentage points."[80] For example, the question can give respondents the option of saying that they thought about voting but didn't or that they usually voted but didn't vote this time.

° Streb used a *list experiment* to decrease the tendency to offer the socially desirable response to a question about voting for a woman for political office. He selected two random samples from the population. The first sample "asked *how many* of the following four statements make them angry or upset."
1. "The way gasoline prices keep going up."
2. "Professional athletes getting million dollar-plus salaries."
3. "Requiring seat belts to be used when driving."

4. "Large corporations polluting the environment."
The second group was given a fifth statement:

5. "A woman serving as president."[81]

To get the percent that was angry or upset about a woman as president all he had to do was to subtract "the average number of items in the baseline condition [the first group] from the average number of items in the test condition [the second group] and … [multiply] by 100."[82]

Mode Effects

The method or mode of survey delivery might affect what people tell us. This is referred to as mode effects. The four basic modes are face-to-face, telephone, mailed, and web surveys although there are many variations of these four modes. This isn't error but simply differences due to the mode of delivery. We're going to consider several studies that illustrate the nature of mode effects. We'll also discuss mode effects further in Chapter 7.

- Cong Ye and associates reviewed 18 experimental studies that compared telephone surveys to other modes and found that respondents in telephone surveys are more likely "to give extremely positive answers … but are not more likely to give extremely negative responses" compared to other modes.[83]
- Holbrook compared a telephone to a face-to-face survey and found that telephone respondents were more likely to satisfice and to give socially desirable responses than face-to-face respondents.[84]
- Peter Preisendorfer and Felix Wolter compared a face-to-face survey and a mailed survey and found that mailed surveys were somewhat more likely to elicit truthful answers to a question about having been convicted of a criminal offense.[85]
- Exit polls in elections are common. Typically an interviewer approaches the respondent outside the polling area on Election Day and asks the respondent to fill out a paper-and-pencil interview. Since more and more voters are voting before Election Day, the paper-and-pencil survey has been

supplemented by a phone survey of early voters. Michael McDonald and Matthew Thornburg compared the Election Day paper-and-pencil survey with the telephone survey and found that telephone respondents were more likely to have higher item nonresponse to the family income question than the paper-and-pencil respondents.[86]

- Douglas Currivan and his associates compared a telephone survey with a telephone audio computer-assisted self-interview in which an interviewer contacts respondents and gets their consent and then the interview itself is conducted over the phone without the interviewer's presence. Respondents answer questions that are prerecorded by pressing keys on their touch-tone phone. Respondents were youth who were asked about tobacco use. They found that "girls, regardless of race/ ethnicity, seem more likely to report smoking if they can do so by pushing a button on their touch-tone phone rather than by providing answers aloud to a human interviewer."[87]

- Dirk Heerwegh and Geert Loosveldt compared a web survey with a face-to-face interview.[88] The web survey had more don't knows and more item nonresponses than did the face-to-face survey. In other words, the web survey demonstrated more satisficing.

Dealing with Mode Effects

Mode effects are not survey error. Rather, they occur because the mode of survey delivery affects respondents in different ways. Telephone surveys represent a different interview environment than face-to-face interviews, and it's not surprising that this might result in greater satisficing as found by Holbrook and McDonald and Thornburg. How then should we deal with mode effects?

- First, we need to be aware of the possibility of mode effects in our data.
- Second, we need to take the possibility of mode effects into account when we report our findings.

- Third, if we combine different modes of survey delivery in our study, we need to compare our findings across the various modes to try to identify what, if any, mode effects are present.

Postsurvey Error

Error can also occur after the survey data have been collected. Error can occur in the processing of data. If we enter data manually in a spreadsheet or statistical program, there is the possibility of error. If we code open-ended questions such as "what is the most pressing problem facing your community today?" we might make coding errors. The solution here is to check our data entry and our coding to see if there are errors. We can have another person independently code or enter the data and then compare the results to determine if there are discrepancies. These discrepancies can then be corrected.

Error can occur in the analysis of our data. Most quantitative analyses use some type of statistical package such as SPSS, SAS, Stata, or R and many qualitative analyses use some type of computer program such as NVivo or Atlas.ti. A simple type of mistake might occur in writing the data-definition statements that create the variable labels, value labels, and designate the missing values. A much more difficult type of error is using the wrong type of statistical analysis. Our best advice is to talk with a statistical consultant if there is any doubt about the proper method of analysis.

Error can occur in the reporting of data. For example, if we conducted a telephone survey of households in our county and we only sampled landline numbers, it would be an error to claim that our findings apply to all households in the county. This would be an example of over generalization. Rather, we should generalize to all households with landline numbers. We'll discuss reporting further in Chapter 8.

Summary

Here's a brief summary of what we have covered in this chapter.

- Error is inevitable in any survey.
- Error can be either random or systematic. Systematic error is referred to as bias.

- Error is typically categorized as follows.
  - Sampling error occurs whenever we select a sample from a population in order to make inferences about the population from our sample data.
  - Coverage error occurs whenever the sampling frame does not match the population. In other words, sometimes the list from which the sample is selected does not match the population and this produces coverage error.
  - Nonresponse error occurs when the individuals who respond to our survey are different from those who do not respond and these differences are related to what we are asking in our survey.
  - Measurement error is the difference between the true value of some variable and the answer that the respondent gives you. Measurement error can be associated with question wording, question order, respondent behavior such as satisficing and giving the socially desirable response, and with interviewer characteristics such as race and sex.
  - Error can also occur in the processing, analysis, and reporting of data.
  - Mode effects are not error but occur when the mode of survey delivery affects what people tell us.
- How can we minimize survey error?
  - We should be aware of the possibility of error and try to identify possible sources of error in our data.
  - We should carefully inspect the list from which we draw our sample and try to identify elements in the population that are left off the list, elements that are on the list but are not part of the population, and elements that occur more than once on our list.
  - We should take steps to minimize nonresponse. However, it's important to recognize that increasing the response rate "will not necessarily reduce nonresponse bias."[89] Jeffrey Rosen and his associates note that "nonresponse follow-up interventions are successful in reducing nonresponse bias to the extent that they secure participation from (under-

represented) nonrespondents who are unlike cases already interviewed."[90]

° We should try to reduce the burden on the respondent of answering our questions. This might reduce the possibility of satisficing.

° Social desirability can be reduced by considering alternative question wording. For example, Duff gave respondents the option of saying they had thought about voting but didn't.

° We should check and recheck our data to make sure that we didn't make errors in creating our data file.

° We should seek advice from a statistical consultant to make sure that we are using the proper method of analysis.

° We should take the possibility of survey error into account when reporting our findings.

Annotated Bibliography

Total Survey Error

• The best place to start is Herbert Weisberg's *The Total Survey Error Approach*. Weisberg makes the point that we need to focus on all possible types of error.[91]

• Don Dillman's books are the next place to go: *Mail and Telephone Surveys—The Total Design Approach*, *Mail and Internet Surveys—The Tailored Design Method*, and *Internet, Mail, and Mixed-Mode Surveys—The Tailored Design Method*. These books are full of examples of the different types of survey error and how to try to minimize them.[92]

Sampling Error

• Your favorite statistics book probably has a good discussion of sampling error.

• If you don't have a favorite statistics book, take a look at *Social Statistics for a Diverse Society* by Chava Frankfort-Nachmias and Anna Leon-Guerrero[93] and *Working With Sample Data* by

Priscilla Chaffe-Stengel and Donald N. Stengel.[94] Earl Babbie, *The Practice of Social Research*, also has a good discussion of sampling error.[95]

- Another excellent source is Leslie Kish's *Survey Sampling*.[96] But be warned, this is a more difficult book.

Other Types of Survey Error

- Weisberg's and Dillman's works cited previously have a good discussion of the other types of survey error (coverage, non-response, and measurement).

CHAPTER 4

Factors to Consider When Thinking About Surveys

The Groundwork

Do You Really Need a Survey?

Alicia was sitting at her desk early one morning, when her boss Bob came rushing into her cubicle. "Mornin' Alicia", Bob said, slightly out of breath and his face flushed, "We've got a problem." Oh my, not the dreaded, "we've got a problem" problem, Alicia thought to herself. "Candice (Bob's boss) called me into her office the first thing this morning," Bob continued, "Turns out the company plans to launch a new product line aimed at the 18 to 25 demographic this fall in select east coast cities. We know about our customer base in that age range, but we don't really know about the demographics of the six major cities that we're targeting for the launch. You're going to have to get some surveys done, so we can see how those places stack up against our customer base. We'll need that to tweak our marketing. Unfortunately, I'm going to need you really pushing on this because we have to have results in eight weeks at the outside. So give it some thought and we'll talk more tomorrow."

Bob rushed out as fast as he had entered, leaving Alicia's head spinning. How would she get six surveys developed, organized, and into the field and back in eight weeks? Not to mention the time it would take to analyze the information and prepare a report for the VPs. It was going to be a brutal couple of months. With only four researchers in her unit, Alicia wasn't even sure she had enough staff to pull off that kind of effort.

Alreck and Settle make the following observation about why surveys are conducted:

Surveys are often conducted simply because it's the only way to get the information needed. Even when the information is available through other means, survey research may be an easier, quicker, less expensive, or more accurate way to get the required information.[1]

We might suggest that Alreck and Settle's observation can be turned around to serve as a good litmus test as to whether conducting a survey is the best way to answer a research question. Specifically, the sponsor and researcher should ask two questions before undertaking a survey: (1) Is a survey the only way to get the information needed? and (2) Is survey research an easier, quicker, less expensive, or more accurate way to get the information? In our previous example, before Alicia gets too focused on the procedural aspects of doing the surveys, it would be wise for her to consider these questions about the proposed surveys. Of course, in order to answer these questions, the most fundamental issue not only to survey research but also to research in general must be answered, namely: What is the research question? In our earlier example, Alicia needs to get a clear indication from her managers as to what overarching question they are trying to answer. In doing so they will need to tell her what *demographics* they are looking for, if they are interested in the demographics of just the targeted age range or of the target cities in general, and so forth. Only then can she truly answer the two questions about the necessity and practicality of implementing the fairly extensive survey effort which was presented to her. If, for example, Alicia's company managers are look-ing for demographics (including economic data) to compare with the demographics of their customer base, there is a wealth of such data read-ily available through the U.S. Census Bureau, which currently conducts more than 130 well-designed surveys each year.[2] It is a distinct possibility that the information they seek is already available without having to go through the expensive process of conducting surveys.

The Stakeholders

Assuming the question of whether a survey is needed has been answered in the affirmative, it is then important to give thought to who will be involved and why. At this stage and throughout the survey process, it is

important to keep in mind that a survey is actually a social interaction in the same way that social media provides a platform for social exchange. There are multiple stakeholder groups involved in the survey process. Some of these stakeholders are involved in the development, production, and execution of the survey. This group would primarily include the sponsors, researchers, and the survey participants. A second grouping of stakeholders would be the consumers of the survey results. This group would include those who are informed of the results, whether directly by the researchers or indirectly by the media. This group could include the public and other scholars and researchers who use the results of a survey in their own surveys or analyses. A third stakeholder group may occur in settings where individuals are assigned to approve or appraise surveys such as human subjects committees or journal editors when results are formally published.*

Because our discussion in this book centers mainly on the development, production, and execution of surveys, our discussion here will be confined to the stakeholders most directly engaged in these processes: the sponsors, researchers, and participants. Alreck and Settle[3] discuss the roles of two of these stakeholder groups, the sponsors and researchers.

In certain respects, the *sponsor* is the owner of the survey, as the sponsor is the one who provides the financial support and, in the case of proprietary surveys, owns the products derived from it. We would note that while the sponsor generally provides the funding to conduct the study, in some cases a third party will provide the resources while the sponsor's role is taken over by an organization that needs and asks that the survey be conducted. This latter situation is seen most often with nonprofit organizations such as charities that receive funding for the survey from a corporate sponsor, or a university group that receives a foundation grant or government funding to engage in a survey.†

The sponsor identifies the purpose of the survey, the population of interest, the timeline, and the approximate resources available. The sponsor

---

* It is important to note here that members of a particular stakeholder group may also be included within another category of stakeholders. For example, sponsors of the survey could well be involved in the development of the survey, but may also be consumers of the survey's results.
† Ethically, if the funder is different from the sponsor, this fact should be disclosed to the researchers and participants and included in any reports that are released to the public.

should also indicate what the expected deliverables are such as a report, presentation, or data. While not generally writing the specific questions for the survey, the sponsor should provide the researchers guidance as to the areas in which information is needed. It is also important for the sponsor to provide a background context and any technical or specialized background information the researchers will need to construct a relevant survey instrument.

The second stakeholder group, the *researchers*, is comprised of the individuals who have the content and technical expertise to conduct the survey. Their role typically includes the development, design, implementation, analysis, and reporting of the results. Large companies, organizations, universities, and governmental agencies may employ in-house researchers with considerable expertise in designing and conducting survey research. However, even for large organizations, it may not be cost effective to maintain an in-house survey research staff if the organization rarely has a need to do a survey. Similarly, smaller companies and organizations, even those who do have a need to do frequent surveys, may not have sufficient resources to maintain an in-house research capacity. Fortunately, there is a wide assortment of external help available through for-profit companies and individual consultants, university researchers, nonprofit organizations, and governmental units. Unfortunately, however, there is a large variation in the quality of the expertise available. If you type search terms such as "survey expertise" or "help with surveys" into your web browser, you will see a staggering array of websites offering everything from customized one-stop survey design and execution to those whose services do little more than step you through a fill-in-the-blank form supposedly creating an online questionnaire.* It should also be noted that some companies are merely brokers. They have little real in-house capacity, but instead actually subcontract needed survey services such as conducting the survey, data entry, or the expertise required to design questionnaires, develop sampling designs, or run the analyses.

---

* Companies selling survey services and products often provide tutorials on their websites covering topics such as designing a questionnaire, obtaining a sample, analyzing results, and so forth. However, as a note of caution, many of these tutorials only really provide instruction on using the tools that the vendor is selling.

Well-established private companies that have established track records of providing survey expertise are characterized by being able to provide a documented history of work they have done, by not overstating what they are capable of doing, and by being transparent regarding their operations and cost structure. Colleges and universities also typically have individuals well versed in conducting survey research and statisticians who can help with design and analysis issues. University faculty also can often bring disciplinary expertise to your survey effort such as understanding educational survey issues or human resource or marketing concerns. Some universities even have survey research centers such as the ones the authors headed. In short, there is a wealth of information from texts, online services, and professional researchers to help with your survey should you lack in-house staff or expertise. Again, however, it is probably worth repeating that the old adage, "any offer that seems too good to be true probably is," certainly applies here.

To the previously mentioned sponsors and researchers, we add *participants* as the third group of stakeholders in the survey process. The participant group is made up of the individuals who will actually be taking the survey. Such individuals might be patients in a clinical trial, co-workers in a company, clients, customers, potential customers, and so forth. During the survey process, they are commonly referred to as *respondents*. If you accept the argument that we made in earlier chapters regarding the value of probability sampling, then participants should be selected in a random fashion and this group should be reflective of the larger population in which you have an interest. In some cases, participants are not compensated for their participation, and take part in a survey because they feel like they are doing the right thing, or because they are interested in the topic. Sometimes, participants see the survey as a way of voicing their opinions or a way of influencing future practices or events. Other participants are paid to take part in a survey. Recently, online survey panels comprised of individuals paid to take surveys have gained popularity as a way of identifying participants for marketing and consumer surveys used in the business community. It should be noted that such panels are not truly random samples as they use an *opt-in* approach. The key characteristic of *opt-in panels* is that the participant pool is not constructed with random selection. Rather, the group of participants is comprised of

self-selected individuals who choose to sign up with a panel, participating at will. A probability-based sample, in contrast, is comprised of subjects who are randomly selected by a researcher or survey company, in which everyone in the target population theoretically has a nonzero chance of being selected.[4] These panels, which can range in size from 100 to over 1,000,000, are usually recruited by third-party vendors, who vet potential participants based on different background characteristics and willingness to meet the participation requirements of the vendor. For example, American Consumer Opinion, a company that provides online panels advertises for panel participants thusly:

> You will never have to pay any money to be a member. Your participation in our surveys is the only "cost" of membership.
>
> Join our paid online survey panel and help evaluate new products, test new advertising, and tell companies what you think. Make your opinions count.[5]

Another company, QuestionPro, offering online survey panels provides a fairly typical listing of attributes that may be used for selecting individuals for survey panels:

- Basic—age, education, ethnic background
- Household—income, pets
- Employment—status, occupation, revenue
- Financial—credit cards, mortgage, properties
- Technical—computer, cell phone, Internet connection
- Medical—smoker, medical condition
- Physicians—specialities [sic]
- Travel—destinations, business and leisure, hotels
- Vehicle—brand, boat, rv[6]

To illustrate the rapid expansion of companies creating survey panels, there are ironically even companies that provide rankings of different survey panel opportunities for potential participants.[7]

Clearly the advantage to businesses of having such participant ready survey panels is their immediate availability; the risk is whether or not the panel truly represents the target population. The rapid expansion of online survey vendors attests to the popularity (and likely profitability) of these approaches, but also provides concerns about quality.[8]

Ethical Considerations

As we mentioned earlier in this discussion, there is a social interaction during the survey process. Irrespective of how brief the contact, during the survey process a relationship exists between the three major stakeholder groups—sponsors, researchers, and participants. And like relationships in general, survey relationships work better when those involved approach the other parties with openness, honesty, and respect. This is important not just because it's the ethical thing to do, but because the future of survey research depends on it. Talking about market research, Ian Brace notes,

> The ability of the market research industry to continue to use sample surveys as sources of primary data depends upon the willingness of members of the public to give their time and cooperation to answer our questions…. To be able to continue, market research needs to maintain this goodwill.[9]

We believe that Brace's conclusion is applicable to all types of survey research.

The end game, when it comes to survey research, is maintaining ethical relationships among the three stakeholder groups. To achieve this result, we must begin with the relationship between the researchers and sponsors. The sponsor can consciously or unconsciously put pressure on the researchers to produce favorable results. This latter situation frequently occurs when the individuals within the sponsor group carry the attitude that they already *know* how the respondents think or feel, but simply want to have a survey done to *verify* this. When both sponsors and researchers are in the same organization, the researchers may consciously or unconsciously feel a need to give the sponsors, who may be

managers or executives in the company, the results the organization is hoping for or would like to see. Even with external researchers, there may be tendency to produce results that please the sponsor so as to secure future business. Even more egregious is the situation when supposed survey research becomes direct marketing or database building. Research by Ian Brace and his colleagues[10] indicates that in today's environment where surveys have become commonplace with almost any kind of customer transaction, those being surveyed often times cannot distinguish between surveys and marketing ploys. The use of such ploys to direct market services or products or collect data from consumers usually requires some type of complicity between sponsors and researchers and thrives in organizations that turn a blind eye to such survey manipulation. Unfortunately, this subterfuge diminishes the value of legitimate surveys and quality of the information that organizations should be able to obtain from surveys.

Professional associations recognize this issue and have developed a code of ethics for their members. The Council of American Survey Research Organizations (CASRO) in the United States, for example, provides a comprehensive code of standards and ethics for survey research organizations that specifies the responsibilities that survey organizations have to respondents, clients, and outside contractors and in reporting study results.[11] The American Association of Public Opinion Researchers (AAPOR) similarly lays out principles of ethical research which emphasize ethical issues such as:

- The responsibility of minimizing risk or harm to the participants;
- The protection of the privacy and confidentiality of information for participants and clients;
- Eliminating misrepresentation of research or when conducting other activities (such as sales, fundraising, or political campaigning) under the guise of conducting survey and public opinion research;
- Adherence to best practices in survey design; and
- Conducting and reporting of surveys.[12]

Scope of the Survey

What resources are needed to do a survey? Like the response to most very general questions, the answer is, "It depends." Every year on college campuses around the country, hordes of students in research methodology classes descend on their fellow students, professors, and the world beyond, armed with clipboards, armloads of paper surveys, and coveted e-mail and other electronic contact lists to conduct a survey. They have been schooled in the basics of survey design, questionnaire construction, methodology, and data collection. Their efforts are modest, generally requiring a small number of hours, perhaps a few supplies like paper forms, and a notebook or laptop computer. Yet the essentials of these students' survey efforts are the same as those of the vast survey efforts that require enormous investments of time such as the Center for Disease Control and Prevention's Behavioral Risk Factor Surveillance System (BRFSS), a 30-year survey effort to monitor state-level prevalence of the major behavioral risks among adults associated with premature morbidity and mortality. This survey, which currently relies principally on Computer-Assisted Telephone Interviewing (CATI), has a total sample size of more than 100,000 participants and in 2011 became the world's largest telephone surveys with slightly over 500,000 interviews.[*] In essence, however, the real difference between small, student-directed exercises and this huge, comprehensive, and complex one might best be characterized as one of scope. What then determines the scope of a survey?

Snijkers and his colleagues argue that we should conceptualize survey research as a project with "a clear purpose (to collect survey data) [and] a clear beginning and end."[13] When thought of as a project, surveys can be broken down into sequential stages: (1) the initiation, development, and planning stage; (2) the data collection stage; and (3) the data analysis and report production stage. In the language of survey researchers, these phases are often called: (1) the prefield stage, (2) the field stage, and the

---

[*] To learn more about the CDC BRFSS survey go to http://www.cdc.gov/brfss/about/index.htm

(3) postfield stage.* Because these stages are sequential, the third stage is highly dependent on the successful completion of the second. Similarly, for the second and third stages to be undertaken successfully, the initiation, development, and planning stage must be well done. Thus, while surveys differ across the three stages it is the first stage where the scope for the study is set. This includes the identification of the survey purpose and objectives, determination of time and resource parameters, and the design and construction of the data collection and analysis methodologies, which are assembled into the research plan.[14]

Complexity and Number of Responding Participants

There are certain *fixed costs* required to conduct a survey. These items remain relatively static as the scope of the survey changes around them. For example, if an organization decides to equip a CATI lab to do phone surveys, there will be costs for purchasing computers, software, work stations, headphones, and so forth, and these costs will remain the same whether the organization does interviews with 5,000 or 50,000 participants. There are a number of these fixed costs such as office space, salaries and benefits, supplies, and so forth, which will remain fairly static even if the scope of the survey changes. One of the decision points for an organization then is whether it engages in enough survey work to warrant the outlay of fixed cost expenditures to create and maintain survey capacity. The alternative is for the organization to contract out all or part of its survey work. Doing so, however, doesn't mean these costs disappear, they are simply *overhead* (frequently called *indirect costs* in academic or governmental research) that will be embedded within the pricing structure of the vendor. The difference is that they are only incurred if and when survey work is actually done. For the company or organization there is a tradeoff then between the costs to maintain in-house survey capacity and the loss of some control over survey process, data, and cost structure. Once base or fixed costs have been factored in, the scope of a survey determines the actual resources required by the project.

---

* Traditionally surveys are said to be *in the field* during the time when data collection is occurring.

Because there are so many variables that may be included in survey cost, it is best to find a way to standardize costs so as to allow oranges-to-oranges, apples-to-apples contrasts. This standardization can be accomplished by determining *cost per response,* which is the total cost of completing one survey response, whether the information is collected with a mail-out paper survey, a telephone or personal interview, or electronic questionnaire. The cost per survey then is driven by two major factors: *survey complexity* and the *number of responding participants.* These two factors differentially impact different phases of the survey. In this regard, Alreck and Settle offer an insightful guideline,

> Survey research strategy tends to lean in one of two directions: Obtain a large amount of data from a small sample, or obtain a small amount of data from a large sample. Usually the resources are limited while the information needs are insatiable.[15]

The complexity of a survey is based on what information is needed, who we need to get that information from, and how we're going to get that information. As Kennedy and associates point out: "Planning and development efforts rise with survey complexity."[16]

The more complex the survey, the more *up-front time and cost* go into developing a blueprint for it, designing our survey instruments such as questionnaires or interview forms, creating sampling and data collection methodologies, and developing the methods of analyzing the information we receive from the survey's respondents. Similarly, the greater complexity, the more *back-end time and cost* are required for activities once the data collection has been completed, such as setting up the survey database structure, developing and running the analyses programs, writing up the findings, and presenting the report.

The second factor, number of responding participants, impacts cost and time resources most during the *data collection phase* of the survey. Please note here that we place emphasis on responding participants rather than just participants, although the two are correlated. If you recall from the discussion in Chapter 2, required sample size is based on the responding participants. Thus, the *rate of response* becomes an important variable when considering survey cost and time. A survey in which one out of five

participants that are contacted respond results in a much more *efficient survey* than one in which one out of 10 participants respond. There are many issues that affect this rate of response on a survey including, for example, the length and design of a survey instrument such as a question-naire, the accuracy of the participants' contact information, and the use of incentives.[17]

To summarize, while the size of our sample could remain fairly sta-ble, the resource utilization would increase as the complexity increased. Conversely, we could have a well-designed survey, whose resource needs would change if we required a bigger survey sample. It is because up-front and back-end tasks and costs remain fairly constant even as the sample size grows that researchers seek to maximize the precision of their results by increasing the sample size when possible. It also provides a solid ratio-nale for repeating surveys using the same designs, instruments, and so forth since much of the cost of such efforts has already been accounted for in the first administration.

Now that we've reviewed some of the fundamental considerations in conducting surveys, we will turn our attention to the delivery modes used for surveys.

Summary

- Do you need to do a survey?
  - Can a survey answer your research question(s)?
  - Is a survey the only way to get the information?
  - Is the survey easier, quicker, a less expensive, or a more accurate way to get the information needed?
- There are three distinct stakeholder groups involved in a survey effort.
  - Sponsors are essentially the owners of the survey. They create the basic research question(s), help define survey objectives, provide the financial support to conduct the survey, and own the survey products (in proprietary surveys).
  - Researchers have the content and technical expertise to conduct the survey. Researchers are responsible for the

survey's development, design, data collection, data analysis, and reporting of the results.

- ◊ In large or mid-sized organizations that do a lot of surveys, researchers may be in-house employees. Caution must be exercised to keep surveys independent and objective when using in-house research staff.
- ◊ For larger or smaller organizations that only occasionally need to do surveys, or for smaller organizations without resources to maintain full-time researchers, there are a variety of outside consultants, outside research companies, and university units and faculty who are able to provide expertise and capacity.

° Participants or respondents are the individuals who will actually be taking the survey.

- ◊ Participants should be randomly selected from the population to ensure they accurately represent the larger population of interest.
- ◊ Individuals take part in surveys for a variety of reasons. Some participate because they feel they are doing the *right thing*, others because they want to make their opinions known or want to influence future events, and yet others take part simply because they are paid.
- ◊ Recently, online survey panels have gained in popularity as a way of identifying participants for marketing and consumer surveys. Online survey panels are comprised of individuals who self-select or *opt-in* to take part in surveys and therefore are not randomly selected.

- Ethical considerations
  - ° The survey process involves a relationship between sponsors, researchers, and participants, and works best when there is openness, honesty, and respect among the parties.
  - ° Sponsors should be aware of the problem of applying pressure, consciously or unconsciously, on the researchers to produce favorable results.
  - ° Researchers should be equally aware of a tendency to provide results that the sponsor is looking for.

- ° Using surveys as a pretext for marketing or advertising is not only unethical, but it creates a sense of mistrust for participants and can ultimately impact the willingness of individuals to take part in survey research.
- ° Protection of participants from risk or harm should be the foremost consideration for both sponsors and researchers. Confidentiality and privacy concerns are two of the major risk areas that need to be addressed.
- Scope of the survey
  - ° Surveys, whether small, simple, low-cost efforts, or huge, complex, resource intensive projects, contain the same fundamental elements and are comprised of sequential stages: (1) the initiation stage, (2) the data collection stage, and (3) the data analysis and report production stage.
  - ° There are certain fixed costs, such as equipment, office space, personal, supplies, and so forth involved in conducting surveys, which are relatively static even when the scope of the survey changes.
    - ◊ Once the fixed-cost expenses are figured, the scope of the survey determines the actual resources required by the project. Survey costs can be standardized to cost per survey response, which permits apples-to-apples, oranges-to-oranges type comparisons.
    - ◊ Cost per survey response is driven by two factors: survey complexity and number of responding participants.
      - —More complex surveys require more up-front time and cost for development and design, and more back-end time and cost for cleaning and coding data, setting up survey databases, analysis and preparing reports.
      - —The number of respondents impacts resources most during the data collection phase of the survey.
      - —The rate of response, or ratio of participants contacted to the number who return surveys is a critical factor and affects the efficiency of the survey.

Annotated Bibliography

Need for Surveys

- There are many books on survey research which take different disciplinary perspectives as to where survey research fits in the larger frame of data collection and its usefulness to address research questions. In Chapter 6 of their book on marketing research, Hague and his associates talk about "desk research" and the amazing amount of information available from different sources.[18]
- Ger Snijkers and his associates present survey research with a business focus. In Chapter 1 of their book, *Designing and Conducting Business Surveys*, they provide a nice overview of the data available for business decision-making, and how surveys fit into this larger realm.[19]

Survey Stakeholders and Roles

- A nice introduction to survey research stakeholders is provided by Alreck and Settle.[20]
- Another perspective is provided by Ian Brace,[21] who discusses stakeholder roles within the context of the questionnaire design.

Ethical Considerations

- Robert Oldendick[22] provides an in-depth review of ethical issues in survey research. The protection of confidential information in the electronic data collection age is becoming a major focal area in research.
- Hundepool and his associates[23] provide a comprehensive look at what is termed as statistical disclosure control.
- A good source of information on the protection of human subjects and the requirements of institutional review boards (IRBs) is provided by AAPOR, which just announced the release of "updated and expanded resources for researchers who conduct surveys subject to institutional review board (IRB) review."[24]

CHAPTER 5

Modes of Survey Delivery

All research starts with questions. A survey is one way to get the information we need to answer these questions. In order to carry out a survey, we have to deliver the survey to our potential respondents. In this chapter, we're going to discuss four basic *modes of survey delivery*. We'll discuss approaches that combine features of two or more of these modes later in this chapter when we discuss mixed-mode surveys.

- *Face-to-face delivery*—We mentioned in Chapter 1 that Don Dillman and his associates have an excellent history of surveys.[1] According to Dillman, face-to-face or in-person interviewing dominated during the first two-thirds of the 20th century.
- *Mailed delivery*—Dillman goes on to say that by the "early 1980s the dominant mode of surveying for government surveys approved by the Office of Management and Budget was mail."[2]
- *Telephone delivery*—Then he notes that "in the early 1980s the telephone … almost completely replaced in-person interviews for surveys of the general public."[3]
- *Web delivery*—In the 1990s web surveys became still another mode of delivery and continued to grow in popularity during the first decade of the 21st century.
- New technological advances led to many variations and off-shoots of these four basic modes of survey delivery.[4] Here are some examples.
  - CATI—*Computer-Assisted Telephone Interviewing*, where telephone interviewers sit at a computer workstation and read the questions off the monitor and enter the respondent's answers on their keyboard.

- ° IVR—*Interactive Voice Response*, where computers are pro-grammed to administer the survey and respondents enter their answers on their touchtone phones.
- ° CAPI—*Computer-Assisted Personal Interviewing*, where interviewers use a computing device such as a laptop or a tablet and enter the respondent's answers directly on the device.
- ° CASI—*Computer-Assisted Self-Interviewing*, where respon-dents enter the information themselves directly on the computing device.
- ° Address-based sampling, which uses the U.S. Postal Ser-vice's Delivery Sequence File to select a sample, has led to an increased use of mailed surveys.[5]

All these advances have led to a veritable explosion of different forms of survey delivery.

Another important change has been the increase in surveys that use more than one mode of delivery, which are often called *mixed-mode surveys*. As Dillman notes, this has allowed researchers to take advantage of certain features of each mode and to "compensate for the inadequacies" of the different modes.[6]

Survey delivery varies along several important dimensions.

- Some surveys are *interviewer-administered* while others are *self-administered*. Face-to-face and telephone interviews are administered by interviewers, while mail and web surveys are typically self-administered. The presence of an interviewer can influence what people tell us or even if they'll respond.
- These modes of delivery provide the researcher with different amounts of control over the administration of the survey. Mailed surveys provide virtually no control over the adminis-tration of the survey, while web, phone, and face-to-face inter-viewing provide more control. *Control over the administration of surveys* allows researchers to more easily do things such as randomize the order of questions and response categories.

- Some modes of delivery provide greater *personal contact between the interviewer and the respondent.* Web and mailed surveys allow no personal contact, while phone surveys provide some contact and face-to-face interviewing allows the most contact. Personal contact has both advantages and disadvantages.

Our focus in this chapter is going to be on the four basic modes of survey delivery—face-to-face, mail, telephone, and web—and on combinations of these called mixed-mode surveys.

Face-to-Face Survey Delivery

Face-to-face interviewing is an interviewer-administered survey, which provides maximum personal contact between the interviewer and the respondent and allows the interviewer to control the actual administration of the survey. Face-to-face interviewing dominated during the development of interviewing methods. Many of the early developments of interviewing techniques can be applied to other modes of survey delivery but we will discuss them under face-to-face interviewing in this section.

Survey research is based on the assumption that all respondents understand each question in the same way. If that is not the case, then respondents would in effect be answering different questions. While we realize that this goal is probably not completely attainable, we need to make every effort to approximate this ideal.

The question, of course, is how to do this. Stephen Richardson and his associates suggest two very different approaches.[7] In a scheduled or as it is typically called *standardized interview,* the interviewer asks the same question using the same question wording and the same question order for all respondents. Think of a medical experiment where we are trying to determine the effect of a particular drug. The drug is the stimulus that has certain effects. In order to determine the effect of the drug, subjects in a particular experimental group must receive the same stimulus. In an interview, the questions and the order in which the questions are asked are the stimuli. So all respondents must receive the same stimuli,

which means that the question wording and order must be identical for all respondents.

In nonscheduled interviewing or as it is typically called *nonstandardized or conversational interviewing*, the wording and order of the questions vary from respondent to respondent in order to make sure that all respondents infer the same meaning to these questions. As Richardson puts it, if we want the questions to have the same meaning for each respondent, then "they must be formulated in wording that is *appropriate for each respondent*."[8] In other words, the interviewer must vary the wording and sequence of the questions in order to keep the meaning constant.

These two different types of interviewing can best be thought of as a continuum from the completely standardized interview at one end to the completely nonstandardized interview at the other end. Most surveys fall somewhere between these two end points mixing elements of both the standardized and nonstandardized interviews.[9]

Patricia Gwartney provides us with "guidelines for asking questions in the standardized interview." These include the following.

- "read all questions exactly as written"
- "read all questions in order prescribed"
- "never skip a question"
- "read questions in a deliberate manner"
- "remain neutral"
- "keep respondents on task"
- "use positive feedback to guide and reward participants"
- ask sensitive questions "in a normal tone of voice at a normal pace"[10]

Frederick Conrad and Michael Schober make the argument for nonstandardized or conversational interviewing.

Conversational interviewing… is designed to make sure that all respondents understand the question the same way—to standardize the *meaning* of that question—irrespective of who reads it to the respondent. It thus embodies the assumption that simply speaking words does not guarantee that the listener will grasp their intended

meaning; speakers and addressees may engage in further dialogue in order to understand each other as well as they need to.[11]

Conrad and Schober describe their study that compared standardized and nonstandardized or conversational interviewing. Respondents were first interviewed using a standardized interview. Then they were reinterviewed one week later. A random half was reinterviewed using the same standardized interview. The other random half was reinterviewed using conversational interviewing. When Conrad and Schober compared the first and second interviews, they found that more respondents changed their answers in the second conversational interview than in the second standardized interview and that the changes "conformed more closely to official definitions."[12] In other words, they found that the changes were in the direction of what the researchers meant to be asking.

Let's look at some examples. Respondents were asked three different types of questions. One set of questions required a numerical answer. For example, some questions dealt with housing such as "How many bedrooms are there in your home?"[13] Other questions required only yes or no responses. For example, there were questions about purchases such as "Have you purchased or had expenses for household furniture?"[14] The other set of questions asked them to list the purchases they made.

When respondents in the standardized interview were confused or unsure how to answer the questions, interviewers responded in a neutral or nondirective manner. For example, they would reread the questions or repeat the response categories. If they were asked what a question meant, they would respond neutrally by saying "whatever it means to you."[15]

Now let's look at the conversational interviewing approach. One question asked if respondents "purchased or had expenses for inside home maintenance or repair services."[16] This was not supposed to include repairs that they did themselves but was meant to refer to repairs that they paid someone else to do. Some respondents misunderstood this. In the conversational approach, the interviewer reminded the respondents that this only referred to repairs that they paid someone else to do.

Another question asked "How many other rooms are there other than bedrooms and bathrooms?"[17] In the conversational approach, the interviewer and the respondent talked over the respondent's answer to see if it

met the official definition of a room. For example, one issue that was discussed was whether a living room and a kitchen that were not separated by a partition counted as one room or two rooms. Another issue that came up was how to count basements and attics.

When Conrad and Schober compared the first standardized interview with the second standardized interview, they found that "57% of respondent's purchases were consistent with the official definitions" in both interviews.[18] In other words, a little more than half of the time respondents answered the questions about purchases in the way intended by the researchers. As an example of an answer that didn't conform to the official definitions, respondents often listed paying phone bills as a purchase but according to the study's definitions, this was not a purchase. However, when Conrad and Schober compared the first standardized interview with the second conversational reinterview, they found that 95 percent of the answers in the conversational reinterview met the official definitions of a purchase compared to only 57 percent of the answers in the standardized reinterview.

Regardless of whether we are using a standardized or a nonstandardized interview, face-to-face interviewing has several important characteristics.

- Consider the interaction between the interviewer and the respondent during the introductory phase of a survey where the focus is on getting the respondent's consent to be interviewed. Since there is more physical contact when we interview someone face-to-face, the interviewer can use the respondent's body language and voice inflections as cues to help convince the respondent to be interviewed. It's easier to make a personal appeal to a respondent when we interact face-to-face. Using these cues, the interviewer has more opportunity to tailor the request for an interview to particular respondents.
- Interviewer-administered surveys like face-to-face interviewing can rely on the interviewer to *probe* when the respondent's answers are unclear or inadequate. For example, if a respondent says that the most pressing problem facing their community is crime, the interviewer can ask the respondent to *tell*

me a little more about that. Interviewers can also help clarify questions when respondents tell us that they are confused.

- The interviewer can make use of *visual materials.* For example, the interviewer can hand the respondent a card that contains the response categories and ask the respondent to choose the best response. Or the interviewer can hand the respondent a map to confirm that the respondent lives in the area of interest.

- Because the interviewer is present, the possibility of measurement error might be increased. It's possible that the respondent might give the socially desirable response or be less likely to truthfully answer sensitive questions such as questions about illegal drug use because of the interviewer's presence. Respondents might answer differently to white or black, male or female, or younger or older interviewers.

- Face-to-face interviewing is typically the most expensive of all modes of survey delivery because interviewers often have to go to the location of the respondent and because of the cost associated with paying the interviewers.

Mailed Survey Delivery

A mailed survey is a self-administered survey with no personal contact between the interviewer and the respondent and the interviewer has little control over the administration of the survey.* There are a number of good references that will help you learn how to do a mailed survey. We listed some of them in the annotated bibliography at the end of this chapter.

Dillman has written extensively on the different modes of survey delivery. In 1978, he proposed *the total design method* for doing mailed surveys.[19] This approach consisted of two parts.

---

* Mailed surveys can be combined with other modes of survey delivery. For example, those who don't return the mailed survey can be contacted personally and encouraged to return it. Respondents can also be given a web address and have the option of either completing and returning the survey through mail or completing it on the web. However, the survey then becomes more of a mixed-mode survey, which we will discuss later in this chapter.

The first is to identify each aspect of the survey process that may affect either the quality or quantity of response and to shape each of them in such a way that the best possible responses are obtained. The second is to organize the survey efforts so that the design intentions are carried out in complete detail.[20]

Using social exchange theory, he discusses how to maximize the rewards and minimize the costs to the respondents and how to ensure trust between respondents and interviewers.

Dillman takes the reader through the steps in designing a mailed survey from writing questions to putting the survey together to maximize the response rate. In other words, the total design method gives you step-by-step instructions on putting together a good mailed survey. We'll talk about writing good questions in the next chapter of this book and about carrying out the survey in Chapter 7.

In Dillman's 2000 and 2009 books, the *total design method* morphed into the *tailored design method*.[21] The focus shifted to reducing all types of *survey error*—sampling, coverage, nonresponse, and measurement. We discussed these different types of survey error in Chapter 3. Dillman's emphasis was on tailoring the survey to fit the population to be surveyed and the information that the survey is designed to obtain.

Mailed surveys have certain important characteristics.

- The motivation to respond must be contained in the *cover letter* that is included with the survey. You need to carefully consider what *incentives* you could provide to encourage the respondents to complete and return your survey. Incentives could include a prepaid cash incentive or a small gift or you could appeal to the respondents' desire to be helpful. You might contact *nonrespondents* by phone and e-mail if that information is available to encourage them to respond but often you don't have their phone number or e-mail address.
- All directions for completing the survey must be contained in the written information. For example, if you want respondents to skip certain questions based on their answers to previous questions, then that has to be made clear.

- Your strategy for maximizing response to the survey has to be carefully planned out and implemented through a series of mailings that make different appeals and catch the respondent's attention.

- There's virtually no personal contact between the researcher and the respondents in a mailed survey. This means that the opportunity to probe and clarify the respondent's answer is limited. Probes will have to be built into the mailed survey. For example, if we ask whether a person is employed and what his or her job title is, we can include a follow-up question that asks respondents to describe what they do in a typical week. This is basically a probe question since we are asking the respondent to *tell us a little more* about what they do on the job.

- Because the respondent is actually looking at the survey you have the opportunity to use visuals. You can include graphics, charts, maps, and other types of visual information.

- Since there is no interviewer present, some types of measurement error might be reduced. Respondents might be more willing to truthfully answer *sensitive questions* and might be less likely to give the socially desirable response. Without an interviewer, we no longer have to be concerned about the effect of the interviewer's race or gender or age on what respondents tell us.

- Mailed surveys along with web surveys are usually the least expensive of the modes of survey delivery because the survey can be delivered to the respondents without an interviewer's presence.

Telephone Survey Delivery

A telephone interview is an interviewer-administered survey with some personal contact between the interviewer and the respondent that allows the interviewer to control the actual administration of the survey. There are a number of good references that will show you how to do a telephone survey. We listed some of them in the annotated bibliography at the end of this chapter.

Let's look at some of the important characteristics of a telephone survey like we did for face-to-face and mailed surveys.

- Because telephone interviews are interviewer-administered the interviewer has a lot of flexibility to use what the respondent says as cues to try to convince the respondent to consent to the interview request.
- Telephone surveys often have significant *coverage* issues. About 2 percent of households in the United States lack any type of telephone (either cell or landline). While this is certainly a coverage issue, it's rather small. A much larger coverage issue is raised by the expanded use of *cell phones*. About 40 percent of Americans rely exclusively on a cell phone,[22] and some segments are even more likely to have only a cell phone such as "Hispanics, African Americans, younger adults and the poor."[23] Other potential respondents have both a landline and a cell phone but it's difficult to reach some of them on their landlines since they rely heavily on their cell phones. Some segments of the population are particularly hard to reach even on their cell phones. Aria Nilson and Ronald Cossman point out that "young adults may move to attend college out-of-state, retaining their old telephone number. If they reside in a wireless-only household, this creates an 'unreachable' segment of the population for geographically-based random digit dialing."[24] As a result of these coverage issues most telephone surveys now include cell phone numbers. The Pew Center has included cell phones in its surveys for some time. Recently they announced that "60% of interviews in our national polls [will be conducted] via cellphones and 40% on landline phones."[25]
- As the phone survey is interviewer-administered, the interviewer has the opportunity to probe to clarify the respondent's answers or to get more information.
- Because telephone surveys are interviewer-administered, this might increase some types of measurement error such as the tendency of respondents to offer the socially desirable response and to answer sensitive questions in a less than truthful

manner. Respondents might also respond differently based on the perceived race or age or gender of the interviewer.

- Telephone surveys can't rely on visuals unless you are able to send copies of the survey to the respondent, which is unlikely. That means that you can't make use of visual materials as you can in a face-to-face, mailed, and web survey.

- Both face-to-face and telephone surveys are interviewer-administered but there is a critical difference. Because you are interacting over the phone certain constraints on a phone survey are introduced. For example, you have to avoid questions with many response categories. By the time you read the fifth or sixth response category the respondent may have forgotten the first couple of categories. That may result in a tendency to choose categories that occur at the end of the list.

- It's possible to capture *audio recordings* of the telephone interview. That's particularly important for open-ended questions such as "what's the most pressing problem facing your community today?" With the audio recording you can transcribe and enter the respondent's answer into a computer file and analyze it using specialized software such as NVivo.[26]

- Computer-assisted telephone interviewing software has been available for many years. It allows you to write computer code that will display the questions and response categories on the interviewer's computer screen. The software allows you to randomize the order of questions and the order of response categories. You can also randomly assign one form of the question to a random portion of respondents and another form to the remaining random portion.

- Phone surveys are usually less expensive than face-to-face interviewing but more expensive than mailed or web surveys. Interviewers can also contact a larger number of respondents in a shorter period of time than in a face-to-face survey.

Web Survey Delivery

Web surveys are self-administered and often provide either no personal contact or very limited contact between the interviewer and the

respondent* but allow the interviewer to control the administration of the survey. As we said earlier about mailed and telephone surveys, there are a number of good references that will show you how to do a web survey. We listed some of them in the annotated bibliography at the end of this chapter.

Let's look at some of the important characteristics of a web survey like we did for face-to-face, phone, and mailed surveys.

- Since web surveys are self-administered, that means that you won't be able to use the respondent's body language and voice inflections to tailor an approach that would encourage them to comply with your request for an interview. However, you might have other information about the respondents such as the department they work in or their staff position in a business that you can use in your appeal for their cooperation.

- Coverage error depends on the population that you want to study. If your population is faculty at a university or employees in a business, you probably can deliver the survey to your sample with a minimum of coverage error. But for populations such as all adults in your community you have no way to deliver the web survey to those without Internet access. Internet access varies by country[27] and by demographic variables such as education,[28] although Anja Mohorko and associates report that the digital gap between those with less and more education is decreasing over time.[29]

- Even though the interview is self-administered, it's still possible to probe for more information to clarify the respondent's answer. *General probes* can be included to inquire why respondents answer the question as they do. Pamela Alreck and Robert Settle discuss *interactive or dynamic probes* in which specific questions are asked depending on key words that occur in the respondent's answer.[30] You can also provide links to information that might help answer questions the respondents have when they try to answer your questions.

---

* For example, there are some approaches that allow video interface using web cams.

- Web surveys might reduce some types of measurement error. For example, research has shown that respondents are more likely to truthfully answer sensitive questions if an interviewer is not present.[31] Respondents might also be less likely to give the *socially desirable response* to certain questions. Without the interviewer's presence, you don't have to worry about the effect of the interviewer's race or age or gender.
- Web surveys allow you to randomize questions and response categories and to use visuals such as maps or tables or charts in your questions. These are powerful tools that can be used because of advances in technology.
- Web surveys along with mailed surveys are usually the least expensive of the modes of survey delivery because the survey can be delivered to the respondents without the cost of interviewers.

Mixed-Mode Surveys

Dillman reports that one of the changes in survey delivery is the increased use of multiple modes of survey delivery to take into consideration the advantages of each mode.[32] This has occurred for a number of reasons.[33]

- Some delivery modes are less expensive. Mailed and web surveys have lower delivery costs than face-to-face and telephone surveys.
- Different modes have different types of coverage error. Phone surveys are limited to those you can reach by phone. Web surveys are limited to those that you can reach over the Internet. By combining various modes you can take advantage of the different coverages of each method.
- Nonresponse is always a concern with surveys. By using different modes of delivery you are often able to increase response to your survey.
- Measurement error is another concern. As we noted earlier, research has shown that the presence of interviewers often affects what people are willing to say. By combining mailed

and web surveys with other modes you can decrease this type
of measurement error particularly for sensitive questions and
questions dealing with socially desirable or undesirable topics.

There are a number of ways in which face-to-face, mailed, telephone, and web surveys can be combined. For example, in a mailed survey you can provide the respondents with a phone number or a web address that they can use to get answers to their questions or express their concerns about issues such as confidentiality. In a telephone survey, you can mail information to respondents that they might request. In a web survey, you can provide a telephone number that respondents can call to indicate their concern about issues raised in the survey.

Let's look at a couple of actual examples of a mixed-mode survey. Marilyn Worthy and Dannielle Mayclin describe the "Residential Energy Consumption Study (RECS) [which] is a survey of housing units conducted by the U.S. Energy Information Administration to measure energy-related characteristics, consumption, and expenditures in U.S. homes."[34] The survey starts with a face-to-face interview of households, which is "followed by the Energy Supplier Survey (ESS), a mandatory survey of the energy suppliers for each household."[35] In the past, a mailed survey was used to collect the information from the suppliers but recently multiple modes of delivery have been used. Suppliers are initially contacted by phone and then by mail to provide the link to the U.S. Energy Information Administration's website. Once the supplier accesses the website, they are given three different ways of providing the information—"paper forms, online forms, or Excel template."[36] After the information is received, it is screened for "missing data, anomalous data (such as outliers or inconsistent patterns), or respondent comments."[37] Data quality was determined by the number of companies that did not respond at all (unit nonresponse), did not answer specific questions (item nonresponse), and by the number of times the Energy Administration had to edit the data.

Data quality varied by the way the energy suppliers chose to provide the information.[38]

- Large suppliers were more likely to use the Excel spreadsheet and smaller suppliers were more likely to use paper or online forms.

- Some suppliers used other modes of responding such as "other electronic files and nonstandard printouts." Item nonresponse was lower "among companies that used the standard modes (paper form, online form and Excel template) than those using nonstandard modes."[39]
- Data quality was higher for the paper and online forms and lower for the Excel spreadsheet. Data changes were made to 31 percent of the cases using the Excel spreadsheet but only 14 percent of those using online forms.

This example shows the utility of mixed-mode surveys but it also shows the importance of thorough editing of the data and the importance of following up to resolve problems.

Another example of a mixed-mode survey is the American Community Survey (ACS). Prior to 2010, the U.S. Census used two forms—a short form and a long form. The ACS which started in 2005 gradually replaced the long form so that only the short form was used in the 2010 U.S. Census. Currently, the ACS is conducted monthly to provide sample data to make estimates for the U.S. population. With the ASC, small areas require five years of data to produce estimates while larger areas require one- or three-year samples. For most households the initial survey was mailed to households. After repeated mailings, telephone calls were made to the households where phone numbers were available. The final phase of the survey was in-person interviews. This series of different survey delivery modes produced very high response rates and careful editing of the survey data produced very high quality data.[40]

Meaning of Survey Questions and Answers

Regardless of the mode by which the survey is delivered, what is essential is determining how respondents interpret the questions and what respondents mean by their answers. Robert Groves and his associates suggested several different approaches.[41]

- *Focus groups*—Form a focus group to talk about the issues you are asking about in your survey. A focus group is a small number of individuals who are part of the population for your survey.

- *Cognitive interviews*—Select some individuals from your population and administer the survey. As they are taking the survey, ask them to tell you how they interpreted the questions and what they meant by their answers.
- *Randomized or split-ballot experiments*—Assign different versions of a question to random parts of the sample and compare the results to see the effect of question wording.

Howard Schuman also proposed an easy way to determine what respondents mean by their answers. Ask them why. For example, if respondents are asked what they consider to be the most pressing problem facing their community, follow up their answers with probes such as "Could you tell me a little more about that" or "Would you explain what you mean by _____?"[42] The same type of probe could also be used in questions that give respondents a choice of several responses. For example, Schuman cites a question asked of respondents during the Vietnam War. The question was "in view of the developments since we entered the fighting, do you think the United States made a mistake in sending troops to fight in Vietnam?"[43] Possible answers were yes, we made a mistake and no, we didn't make a mistake. But what did respondents mean by *mistake*?[44] An easy way to find out is to ask them what they meant by their answer. Another possibility is to randomly select respondents for probing, what Schuman calls the *random probe*.[45] A similar approach is suggested by George Bishop, who proposed asking respondents "to think out loud" or to "talk about" how they arrived at their answers.[46]

These suggestions illustrate several ways to find out what respondents mean by their answers and how they interpret the question. They can be incorporated into most surveys without much difficulty and will help the researcher interpret their data.

Summary

- Modes of survey delivery include:
  - Face-to-face
  - Mailed
  - Telephone

- ° Web
- ° Mixed-mode
- Dimensions of survey delivery
  - ° Interviewer-administered versus self-administered
  - ° Degree of control over the administration of the survey
  - ° Degree of personal contact between interviewers and the respondents
- These different modes of survey delivery all have advantages and disadvantages.
  - ° The presence of an interviewer provides greater opportunity for convincing the respondent to comply with our request for an interview by allowing the interviewer to observe body language and vocal inflections.
  - ° There is a greater tendency to give the socially desirable answer in an interviewer-administered interview. Respondents are less likely to respond accurately to sensitive questions with an interviewer present. And the interviewer's race or gender or age might affect what respondents tell us.
  - ° Visual materials such as maps or charts or tables can be used in face-to-face, mailed, and web surveys but typically not in telephone surveys. The interactive nature of web surveys allows display of visual materials tailored to particular respondents.
  - ° Audio-recordings can be made of face-to-face and phone surveys to allow for better capture of the respondent's answers to open-ended questions.
  - ° Technological advances in the administration of phone and web surveys make it easy to randomize the order of questions and response categories.
  - ° Coverage error remains a concern in all types of surveys depending on the mode of survey delivery and the nature of the population being surveyed.
  - ° Mailed surveys along with web surveys are usually the least expensive of the modes of survey delivery because the survey can be delivered to the respondents without the cost of having an interviewer.

- ○ Mixed-mode surveys are used to decrease cost, increase response and coverage, and reduce measurement error by taking advantage of the different survey delivery modes.
- ○ There are several ways to try to understand how respondents interpret survey questions and what they mean by their answers. These include approaches such as asking why and thinking out loud.

Annotated Bibliography

Face-to-Face Survey Delivery

- Stephen Richardson and his associates' *Interviewing—Its Forms and Functions* is a good place to start.[47] They discuss both the standardized and nonstandardized interviewing approaches, which they refer to as scheduled and nonscheduled interviewing.
- For further reading, look at Raymond Gorden's two books—*Interviewing: Strategy, Techniques and Tactics*[48] and *Basic Interviewing Skills.*[49]

Mailed Survey Delivery

- Don Dillman's three books: *Mailed and Telephone Surveys—The Total Design Method,*[50] *Mail and Internet Surveys—The Tailored Design Method,*[51] and *Internet, Mail, and Mixed-Mode Surveys—The Tailored Design Method*[52] are the places to go to learn about mailed surveys.

Telephone Survey Delivery

- Start with Dillman's book on telephone interviewing mentioned previously.
- Also look at Patricia Gwartney's *The Telephone Interviewers' Handbook.*[53]

Web Survey Delivery

- Mick Couper's *Designing Effective Web Surveys* is an excellent discussion of all aspects of web surveys.[54]

Mixed-Mode Surveys

- Dillman and his associates' 2009 book mentioned previously is a great source of information about conducting mixed-mode surveys.

Determining the Meaning of Survey Questions and Answers

- Robert Groves and his associates' *Survey Methodology*[55] and Howard Schuman's *Method and Meaning in Polls and Surveys*[56] will help you figure out ways to better understand what respondents mean when they answer your questions.

CHAPTER 6

Writing Good Questions

Nearly four decades ago, Warwick and Lininger indicated that

> Survey research is marked by an unevenness of development in its various subfields. On the one hand, the science of survey sampling is so advanced that discussion of error often deals with fractions of percentage points. By contrast, the principles of questionnaire design and interviewing are much less precise. Experiments suggest that the potential of error involved in sensitive or vague opinion questions may be twenty or thirty rather than two or three percentage points.[1]

While both sampling methodology and survey question development have advanced significantly since Warwick and Lininger made that observation, the development of question methodology continues to be important because of its relevance to measurement error (which is an important component of the total survey error). Continuing research across different disciplines is expanding our horizons as to subtle problems regarding survey question construction. For example, a recent study by Fisher[2] examined the effect of survey question wording in a sensitive topic area involving estimates of completed and attempted rape and verbal threats of rape. The results of the study show significant differences between the two sets of rape estimates from two national surveys: the National Violence Against College Women study and the National College Women Sexual Victimization study, with the latter study's estimates ranging from 4.4 to 10.4 percent lower than that of the former. While Fisher attributes the difference between the two surveys to four interrelated reasons, "the use of behaviorally specific questions cannot be overemphasized, not necessarily because they produce larger estimates of rape but because they use words and phrases that

describe to the respondent exactly what behavior is being measured."[3]
Essentially, Fisher is pointing out that estimates coming from surveys
that use more specific language describing victim's experiences in behavioral terms, such as "he put his penis into my vagina," produce more accuracy in the responses and thus improve the overall quality of the survey.

In a different discipline, potential response bias resulting from racial or ethnic cultural experience was found in research on health behavior by Warnecke and associates.[4] Among the researchers' findings was evidence in support of differences in question interpretation related to respondent ethnicity. They suggest that providing cues in the question that help respondents better understand what is needed may address these problems.

Finally, in a third discipline, an economic study examining household surveys asking individuals about their economic circumstances, financial decisions, and expectations for the future conducted by Bruine de Bruin and associates[5] found that even slight changes in question wording can affect how respondents interpret a question and generate their answer. Specifically, the authors' concluded that questions about "prices in general" and "prices you pay" focused respondents more on personal price experiences than did questions about "inflation." They hypothesized that thoughts about personal price experiences tend to be biased towards extremes, such as large changes in gas prices, leading respondents to over-estimate overall inflation. Essentially what would be considered irrelevant changes in question wording affected responses to survey questions.

These three examples from different disciplines serve as important illustrations of the sensitivities continuing to be explored related to the structure and format of survey questions. The goal of this chapter is to acquaint you with some structural issues on questions design and to provide general guidelines on writing good survey questions. Before we plunge into the topic, however, some context on survey questions will help clarify some of the points we make later.

We begin by highlighting the distinction between the survey mode, the survey instrument, and the survey questions. The mode, discussed in Chapter 5, is the method of delivery for the survey. The survey instruments often referred to as questionnaires can range from a traditional paper and

pencil mail-out form, to a self-administered electronic online survey with embedded audio and video and to the survey interview screen seen by CATI telephone interviewers when they call interview participants and record the responses into a computerized database. Essentially, the survey instrument is the platform for the questions, while the questions are the expressions—a word, phrase, sentence, or even image—used to solicit information from a respondent. When talking about survey elements, the mode, the survey instrument, and the survey questions are interwoven and frequently discussed simultaneously (see for example Snijkers et al.[6]). However, in our discussion here we have decided to put our focus on the questions rather than the questionnaire for two reasons. First, as Dillman notes in his discussion of the evolution from his Total Design Method (introduced in 1978)[7] to his current Tailored Design Method,[8] advances in survey methodology, changes in culture, rapidly changing technology, and greater emphasis on self-administered surveys have created a need to move away from a one-size-fits-all approach. This more formulaic approach has been replaced by a more customized one that can be adapted to the situation and participants.[9] Gone is the emphasis on the rigorous standardization of questionnaire design and administration down to the order and positioning of questions, the number and timing of mailings, the size and folding of the questionnaires, and so forth.[10] The survey field now embraces the idea of mixed-mode surveys, in which different participants may be surveyed using different modalities as were discussed in the last chapter. As a result, the ability to create questions that can be formatted to fit into different modes is very critical to these efforts. Second, while the modes of survey delivery continue to change, with a greater emphasis on self-administration and rapid electronic delivery and response, the fundamentals of good questions remain the same. In summary, it is essential to understand the most basic element in survey design and administration—the question—as we try to assure compatibility across a range of delivery modes and formats if we are to produce valid and reliable surveys. Even a well-designed and formatted questionnaire cannot undo the damage to a survey caused by questions that are poorly conceived, badly constructed, or offensive to the respondent.

Begin at the Beginning

One of the themes you will see throughout this book is the importance of having a clear idea of what information you need to get in order to answer your research question(s). Before you ever begin to tackle the development of your questions for the survey, you should have these research questions and objectives written down, and a clear understanding and agreement between the sponsor and the researchers as to what they are. At the end of the survey project, the test of whether a survey was successful will be whether those original research questions were answered.

Dillman notes three goals for writing good questions for self-administered surveys so that every potential respondent will: (1) interpret the question the same way, (2) be able to respond accurately, and (3) be willing to answer.[11] Let's briefly take a look at Dillman's three goals, which will help frame the actual design of questions.

Validity and Reliability in Survey Questions

The first two question qualities that Dillman's goals highlight center on two important concepts that must be addressed in our questions: *reliability* and *validity*. Reliability and validity are two of the most important concepts in research, generally, and much of the effort we put into survey research is directed toward maximizing both to the greatest extent possible. In surveys, reliability refers to the consistency in responses across different respondents in the same situations.* Essentially, we should see consistency of the measurement, either across similar respondents or across different administrations of the survey. In a questionnaire, this means that the same question elicits the same type of response across similar respondents. To illustrate, the question "In what city do you currently live?" is an extremely reliable question. If we asked

---

* It is important for researchers to recognize changes in the context of the situation, which might affect the consistency of responses. For example, if a survey on school safety was administered to children in a particular school districted before and after a major school shooting was reported in another part of the country, the situation of the survey might appear to the same, but the situation context would be substantially different.

100 people this question, we would expect to see a very high percentage responding to *city of residence* in a similar fashion, by naming the city in which they are currently living. Likewise, if we were to ask similar groups this same question in three successive decades, we would again expect the kind of responses we would get to be very parallel across those years. By contrast, if we asked 100 people, "How much money in dollars does it take to be happy?" we would find a great deal of inconsistency in their responses. Further, if we posed the same question to groups across three decades, we would likely find a great deal of variation in their responses. One of the major differences between the two questions is the degree to which perceptions of key concepts are shared among participants. In the first question, the concept of city of residence has a generally shared definition. By contrast, in the second question, the definition of concept of happiness is vague and not as widely shared. Many people, for example, would probably have great difficulty in putting a monetary value on what they view as essential to being happy.

While perceptions of the question by survey participants affect reliability on written survey instruments, when the questions are presented in an interview format, we must also deal with the differences between interviewers. Reliability can be impacted if the question is asked differently by different interviewers, or if a single interviewer varies the way the question is presented to different participants. Once the question has been asked, reliability can also be impacted by how the response is recorded by the interviewer. Interviewer issues, including the importance of interviewer training, are discussed later in Chapter 7.

Validity,* on the other hand, refers to the extent that the measure we are using accurately reflects the concept we are interested in, or as Maxfield and Babbie note, "Put another way, are you really measuring what you say you are measuring?"[12] Let's revisit our first question again, "In what city do you currently live?" If we asked that question to a group of high school students, we would likely get accurate responses. On the

---

* There are four major areas of validity: face, content, criterion, and construct, which we will not discuss. The interested reader can go to any introductory statistics or social research methodology text to learn more about these different types of validity.
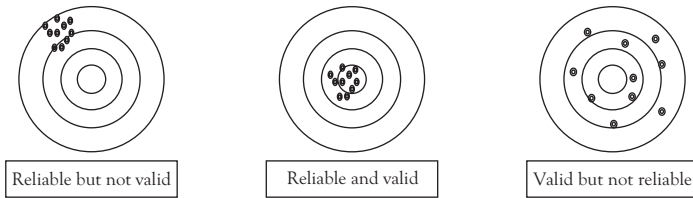
Figure 6.1 Relationship of reliability and validity in question design

other hand, if we were trying to assess the intelligence of high school students and asked them what their grade point average was, and then based on those grade point averages concluded that the percentage of students with an A average was the percentage of very intelligent students, the percentage of those with a B average was the percentage of students with above-average intelligence, the percentage of those with C averages was the percentage of average intelligence students, and so forth, we would be drawing an invalid conclusion because grade point average isn't an accurate measure of intelligence.

There is interplay between reliability and validity in survey research, and when creating survey questions, we must pay attention to both. One of the common analogies used to help understand the relationship between reliability and validity is the target shown in Figure 6.1. The objective in our questions is to hit the bull's-eye on the target.

Willingness to Answer Questions

Being asked to complete a survey, whether it's conducted in person or online, by mail or on the phone, probably isn't on the top of most people's list of things they most enjoy. The issue surrounding potential participants' willingness to take part in a survey has become very visible in the past couple of decades because of substantially declining survey participation response rates.[13] As a result, there has been considerable attention paid to the use of incentives as ways to improve participation, as well as the design factors that intrinsically make participation more likely. Research has consistently shown that incentives to reduce the burden on respondents as well as rewarding them for their help have a significant impact on improving responsiveness.[14]

The willingness of individuals to participate in surveys has been explained by two major theoretical perspectives. As discussed in Chapter 2, Robert Groves and his colleagues have presented what they call *leverage-saliency theory* to describe the decision to participate in a survey.[15] They theorize that the individual's participation decision is due to multiple factors—some survey-specific (topic and sponsorship), others person-specific (such as concerns about privacy), and still others specific to the respondent's social and physical environment. Groves and associates believe that each factor may move a particular person toward or away from cooperation with a specific survey, and these factors carry different weights for different people. They become relevant or salient when, for example, an interviewer introduces the survey and requests participation.

Don Dillman and his colleagues offer a different theoretical perspective, based on social exchange theory, which maintains that respondents will be more likely to participate in a survey if their costs are minimized, the benefits to them are maximized, and they trust that the benefits will actually occur.[17] Although both of these perspectives refer to the totality of participation, we would argue that the rationales they offer regarding survey participation also apply to the willingness of participants to be engaged with the survey questions, which are the fundamental building blocks of the survey. Participants must see a value to their involvement that outweighs the effort they need to expend by participating. If participants are not motivated to answer each question, if they see no benefit from their effort, if a question is offensive or demeaning, if they don't understand a question, or if they believe that answering a question will result in harm to them (such as a violation of their privacy), it is likely they simply won't answer the question.

Key Elements of Good Questions

So what are the essential elements of good questions? A reading of the literature from general textbooks to highly specialized journal articles will provide a vast, almost alarming assortment of recommendations, cautions, and directions—as Dillman puts it, "a mind-boggling array of generally good, but often confusing and conflicting directions about how

to do it."[17] So as to avoid falling into this trap ourselves, we will use a slightly modified version of attributes suggested by Alreck and Settle to attempt to distill these recommendations down into three major areas: specificity, clarity, and brevity.[18]

Specificity, Clarity, and Brevity

By *question specificity*, we are referring to the notion that the question addresses the content of the information sought as precisely as possible. Does the information targeted by the question match the target of the needed information? If a question does not, the results it produces will have low validity in terms of addressing the research objective. A rough analogy might be using a pair of binoculars to look at a distant object. If you close first your left eye, then your right, each eye will see the object independently. However, if you open both eyes and instead of seeing one image with both, you see two images, then you know some adjustment to the binoculars is needed. Similarly, there should be a high level of congruence between the research objectives and the question(s) asked. If there is not, some tweaking of the question(s) will be needed. For the question to accurately address the research question(s), it must also be relevant to the survey respondent. If you ask survey respondents about a topic with which they are unfamiliar, the question may have high congruity between its topic and the information needed, but would do a poor job of getting that information from respondents.

The second area, *question clarity*, is one of the biggest problems in survey research, particularly when it's used with self-administered survey instruments. Lack of clarity has a large impact on both question validity and reliability because the question must be equally understandable to all respondents. The core vocabulary of the survey question should be attuned to the level of understanding of the participants. There is often-times a disparity between what the survey sponsors or the researchers know about the question content and the respondents' level of understanding. Frequently, this happens when technical terms or professional jargon, very familiar to sponsors or researchers but totally unknown to respondents, is used in survey questions. To illustrate, consider the following question that was actually found on a consumer satisfaction

survey sent to one of this book's authors: "How satisfied are you with your ability to engage the safety lock-out mechanism?" The response categories ranged from *very satisfied* to *not satisfied at all*. The problem was the author wasn't aware there was a safety lock-out mechanism on this product or how it was used! The reverse of this problem can also hamper question clarity. This happens when sponsors or researchers have such a superficial knowledge of the topic that they fail to understand the intent of their question. For example, if a research firm asked small businesses if they supported higher taxes for improved city services, they might find respondents asking "Which taxes?" "Which services?"

The third area, *brevity*, has to do with the length of the question. The length and complexity of questions affects the response rate of participants as well as impacts the validity and reliability of the responses. Basically, questions should be stated in as straightforward and uncomplicated a manner as possible, using simple words rather than specialized ones and using as few words as possible to pose the question[19] (although this last caution may be more applicable to self-administered questionnaires than for interview formats).[20] More complex sentence structures should be avoided. For example, compound sentences (two simple sentences joined by a conjunction such as *and* or *or*) or compound–complex sentences (those combining an independent and a dependent clause) should be broken down into two simpler questions.[21]

Avoiding Common Question Pitfalls

Before moving on to question types, let's look at some common question pitfalls that apply equally to different question types and formats.

- *Double-barrel questions*: These are created when two different topics are specified in the question, essentially asking the respondent two questions in one sentence. This leaves the respondent puzzled as to which part of the question to answer.
  - Example. How would you assess the success of the Chamber of Commerce in creating a favorable business climate and an awareness of the negative impact of over-taxation on businesses?

Correction: The question should be split into two separate questions:

How would you assess the success of the Chamber of Commerce in creating a favorable business climate?

How would you assess the success of the Chamber of Commerce in creating an awareness of the negative impact of over-taxation on businesses?

- *Loaded or leading questions*: These originate when question wording directs a respondent to a particular answer or position. As a result, the responses are biased and create false results. Political *push polls*, which are sometimes unethically used in political campaigns, illustrate extreme use of loaded questions. They create the illusion of asking legitimate questions but really use the question to spread negative information by typically using leading questions (see the second example).
  - ° Example. Don't you see some problem in letting your children consume sports drinks?

    Correction: The question should be reworded to a neutral statement.

    Is letting your children consume sports drinks a problem?

  - ° Example. Are you upset by Senator _____'s wasteful spending of your tax dollars on programs for illegal immigrants?

    Correction: All negative references in the *question* should be removed.

    Should programs for undocumented immigrants be supported with tax dollars?

- Questions with built-in assumptions: Some questions contain assumptions that must first be considered either true or false in order to answer the second element of the question. These pose a considerable problem as the respondent may feel disqualified from answering the second part of the question, which is the real topic focus.

- ° Example. In comparison to your last driving vacation, was your new car more comfortable to ride in?
- ° Correction: Potential respondents may hesitate to answer this question because of an assumption contained in it: that the individual has taken a driving vacation. This question could be split into two separate questions. (1) Have you previously taken a driving vacation? (2) If yes, in comparison to your last driving vacation, was your new car more comfortable to ride in?
- Double-negative questions: Questions that include two negatives not only confuse the respondent, but they may also create a level of frustration resulting in nonresponse.
  - ° Example. Please indicate whether you agree or disagree with the following statement. A financial advisor should not be required to disclose whether or not the advisor gets any compensation for clients who purchase any of the products that the financial advisor recommends.
  - ° Correction: The *not be required* in the question adds a layer of unnecessary complexity. The question should be worded, "A financial advisor should be required to disclose whether or not he or she gets any compensation for clients who purchase any of the products that the advisor recommends."

Question Types and Formats

The formatting of the survey question takes into account the research objectives, the characteristics of the respondent, the survey instrument and mode of delivery, and the type of analysis that will be needed to synthesize and explain the survey's results. Broadly speaking, there are two principal types of survey questions: unstructured and structured. *Unstructured questions* are sometimes called *open-ended* because they do not restrict the possible answers that survey respondent may give. The second general question type, *structured*, is commonly referred to as *closed-ended* because the survey participant is limited to responses or

response categories (pre)identified by the researchers. For example, if we are doing a telephone survey of community needs, we might use either of the following two survey questions.

1. What do you like best about living in this city?
2. What do you like best about living in this city?
   a. A good transportation system
   b. A low crime rate
   c. A lot of entertainment and recreational opportunities
   d. A good school system
   e. Good employment opportunities

The first question is open-ended, while the second is a closed-ended question. Let's briefly take a look at the characteristics of both types.

Open-Ended Questions

In responding to the first question, which is open-ended, a respondent's answer could obviously cover many different areas, including some the researchers had not previously considered. In this respect, open-ended questions are particularly well-suited to exploring a topic or to gathering information in an area that is not well known. With open-ended questions, the scope of the response is very wide, so they generally work best when the researcher wants to provide, in essence, a *blank canvas* to the respondent. On the downside, with the open-ended question the participant's response may be well outside the question's focus, with an answer such as "My best friend lives here" or "I get to live rent-free with my parents," neither of which really addresses community needs issue, which was the intent of the question. Thus, the question response would have low validity.

When self-administered surveys use open-ended questions, the ability to interactively engage in follow-up questions or *probe* to clarify answers or get greater detail is limited. For this reason, considering the aspects of specificity, clarity, and brevity in question design is especially important. Wordy, ambiguous, or complex open-ended question formats not only create difficulty in terms of the respondent's understanding of the questions, but also may present a visual image format that suggests to the respondent

that this question will be difficult and time-consuming to answer. For this reason, an open-ended question should never cross over more than one page on a written survey or require respondents to read across different screens in a computerized format. Similarly, if the question is provided in a paper format, the space provided to answer the question should directly follow the question rather than being placed on a separate page or after additional questions. Simple formatting elements such as providing sufficient space to allow the respondent to write or type in a narrative-type response in a paper or online survey are very important.[22]

When open-ended survey questions are administered to the survey participant in an interview, the response is often recorded verbatim or with extensive notes, which is useful to later pick up on the nuances of the response such as how the person responding phrases an answer or the strength of a feeling they express in their response. Similarly with in-person, telephone, or interactive online interviews, the interviewer can use follow-up questions to obtain more specific information or probe for more detail or explanation of the open response. In some cases, these *probes* are anticipated and preprogrammed into the interview questionnaire, but in others, they are spontaneously developed by the interviewer based on answers that are not clear or lack detail. Obviously, the experience and qualifications of the interviewers have a major impact on the ability to follow up with conditional probes.

Open-ended questions can be time consuming for both the respondent and researcher. For the respondent, it requires the individual to not only recall past experiences but also to make a judgment as to how best and with how much detail to answer. For the researcher, open-ended questions often yield many different responses, which may require additional coding and can complicate or even prevent the analysis. Also, since open-ended question responses are typically in narrative format, a qualitative rather than quantitative analysis of the information must be anticipated. Such analysis, even when aided by computerized qualitative analysis programs,[23] typically requires more time and effort. Due to the increased effort on the part of the respondents and researchers with these types of questions, from a practical perspective, the number of open-ended questions must be held to a reasonably low number on a particular survey instrument. It also prompts researchers to avoid using open-ended questions when they are working with large samples.

Closed-Ended Questions

The format of closed-ended questions is more defined, and has been standardized to a greater extent than the open-ended style. While both open- and closed-ended questions require researchers to craft questions carefully, closed-ended questions place an additional burden on researchers to carefully consider what responses are needed and are appropriate. As can be seen in the second question, the closed-ended responses are restrictive, and therefore must be precisely targeted to research questions. The closed-ended format does allow researchers to provide greater uniformity to the responses and to easily determine the consensus on certain items, but only on those items that were specified by the answers provided. This, in turn, can lead to another problem as pointed out by Krosnick and Fabrigar;[24] because of researcher specification with closed-ended questions, open-ended questions are less subject to the effect of the researcher.

With closed-ended questions, the response choices should be both *exhaustive* and *mutually exclusive*. This means that all potential responses are listed within answer choices, and that no answer choice is contained within more than one response category. We should point out a distinction here between single- and multiple-response category questions. In single-response questions, only one of the choices can be selected and therefore the response choices must be exhaustive and mutually exclusive. However, some questions are worded in a way that a respondent may choose more than one answer from the choices following the question. In this situation, the response choices are each still unique, but the person responding can select more than one choice (see Example 4).

Consider the following closed-ended question:

Example 1

In which of the following categories does your annual family income fall?

a. Less than $20,000
b. $21,000–$40,000
c. $40,000–$60,000
d. $61,000–$80,000

Can you see a problem with the response set? If you said that it has both nonmutually exclusive categories and does not provide an exhaustive listing of possible family income levels you are right. Both problems are seen with the current response categories. If you look carefully, for a respondent whose family income is between $20,000 and $20,999, there is no response category from which to choose. Similarly, if a respondent's family income level is $105,000 a year, the individual would be in a similar quandary, as again there is no appropriate answer category. A problem also arises if the respondent has an annual family income of $40,000 a year. Which category would the individual choose—(b) or (c)? Fortunately, these two problems are easy to fix. To solve the issue of nonmutually exclusive categories, we would change response (c) to "$41,000, to $60,000." To correct the problem of transforming the response set of answers to be exhaustive, we could change the first response option (a) to "less than $21,000" and add another response option at the end of the current group, (e) "more than $80,000."

One of the first considerations with closed-ended questions is the level of precision needed in the response categories. Take for example the following three questions that are essentially looking for the same type of information.

Example 2

When you go clothes shopping, which of the following colors do you prefer?
   (Please check your preference)

_____Bright colors   _____Dark colors

Example 3

When you go clothes shopping, which of the following colors do you prefer?
   (Please check your preference)

_____Bright colors   _____Dark colors   _____No preference

Example 4

When you go clothes shopping, which of the following colors do you prefer?

(Please check each of your preferences)

___Yellows   ___Browns   ___Reds   ___Greens   ___Blues   ___Pinks
___Blacks   ___Whites   ___Oranges   ___Purples   ___Lavenders
___No preference

Each of these questions provides a different level of precision. In the first question, the choices are limited to two broad, distinct color categories, which if researchers were only looking for general impressions would be fine. However, if the question was on a fashion survey, this level of detail wouldn't be sufficient. The second question also opens up another answer possibility, that is, a response that indicates the individual doesn't have a color preference. The third question, which could be expanded to any number of response categories, not only provides an indication of specific color choices, but also allows the individual to select specific colors from both bright and dark areas. This question could be further enhanced by providing a visual image, such as a color wheel, that would let the respondents mark or check precise colors, thus insuring greater reliability in the answers provided across different respondents.

Unfortunately, there is always a delicate balance in trying to get to the greatest level of precision in question responses on one hand, while not sacrificing the respondent's ability to answer the question accurately on the other. With closed-ended questions, the formatting of the response categories can impact important response dimensions such as the ability to accurately recall past events. More than 40 years ago, Seymour Sudman and Norman Bradburn described the problem of recall on memory:

There are two kinds of memory error that sometimes operate in opposite directions. The first is forgetting an episode entirely, whether it is a purchase of a product, a trip to the doctor, a law

violation, or any other act. The second kind of error is compression of time (telescoping) where the event is remembered as occurring more recently than it did. Thus, a respondent who reports a trip to the doctor during the past seven days when the doctor's records (show) how that it took place three weeks ago has made a compression-of-time error.[25]

The second problem in trying to make the response categories too precise occurs when the response categories become impossible to differentiate in the respondents mind. It's a little bit like the average wine drinker trying to distinguish if the sauvignon blanc wine promoted by the local wine shop actually has *a fruit forward taste with plum and cherry notes and a subtle flowery finish*. In a survey question directed to office workers, the problem might look like this:

Example 5

If you spend more time responding to business e-mails and text messages this year compared to last year, please choose the category below that *best* describes the difference in the average amount of time per day you spend responding to business e-mails and text messages this year compared to last year?

a. _____Less than 8 minutes
b. _____8–16 minutes more
c. _____17–25 minutes more
d. _____26–34 minutes more
e. _____35–43 minutes more
f. _____More than 43 minutes

As you can imagine, the average worker would likely have great difficulty in trying to estimate time differences with this degree of specification. In essence, the researchers are trying to create too finite a distinction in the categories. This concept is often referred to in terms of the *granularity* of the response category, which essentially refers to the level of detail in the response categories.

There are two basic formats for closed-ended questions: (a) unordered or unscalar and (b) ordered or scalar.[26] The first of these, the *unordered or unscalar* response category, is generally used to obtain information or to select items from simple dichotomous or multiple choice lists. The data obtained from this type of question is usually categorical, measured at the nominal level, which means there are discrete categories but no value is given to the categories. Here are some examples.

Example 6

Will your company be doing a significant amount of hiring in the next year?

_____Yes   _____No

Example 7

If your company plans to expand its workforce in the coming year, which of the following *best* explains that expansion?

_____Rehiring from previous downsizing
_____New markets have created greater product demand
_____Expansion in existing markets has created greater product demand
_____New products coming to market

Such unordered response categories are sometimes referred to as forced-choice categories because the respondent can only choose one answer. Some forced-choice questions are used to help determine choices between areas which, on the surface, appear to have an equal likelihood of being selected.

Example 8

When buying a new home, which of the following do you consider most important?

_____Cost   _____Location   _____Size   _____Age

Unordered response categories may be partially *opened* by providing an alternative to the choices listed by including an *other* category with a blank line that allows a respondent to insert another response in addition to those listed.

Example 9

When buying a new home, which of the following do you consider *most* important?

_____Cost
_____Location
_____Size
_____Age
_____Other (please explain)_____

Ordered or scalar response category, as the name implies, arranges responses in an order by requiring the respondent to select a response that conveys some order of magnitude among the possible choices. These response choices are measured by ranking or rating the response on a scale at the ordinal, interval, or ratio level. With ordinal ranking, the response categories are sorted by relative size, but the actual degree of difference between the items cannot be determined. For example, consider commonly seen scales that ask respondents to indicate whether they *strongly agree*, *agree*, *neither agree nor disagree*, *disagree*, or *strongly disagree*. Thus, each identified response category becomes a point along a continuum. One of the first and most commonly used rating scales is the *Likert Scale* that was first published by psychologist Rensis Likert in 1932.[27] The Likert Scale presents respondents with a series of (attitude) dimensions, which fall along a continuum. For each of the attitude dimensions, respondents are asked whether, and how strongly, they agree or disagree, using one of a number of positions on a five-point scale. Today, Likert and Likert-type scales are used commonly in surveys to measure opinions or attitudes. The following example shows a Likert scale question and response set.

Question: Please rate the employees of this company for each of the areas listed below.					
For each item below, please check the answer that best applies, from strongly disagree to strongly agree	Strongly disagree	Dis- agree	Neither agree nor disagree	Agree	Strongly agree
1. The employees in this company are hard working					
2. The employees in this company have good job skills					
3. The employees in this company are dependable					
4. The employees in this company are loyal to the company					
5. The employees of this company produce high quality work					

A variant of the Likert Scale is the *Semantic Differential Scale*, another closed-ended format, which is used to gather data and interpret it based on the connotative meaning of the respondent's answer. It uses a pair of clearly opposite words that fall at the ends of a continuum and can be either marked or unmarked. Here are some examples of a semantic differential scale.

Marked Semantic Differential Scale

Please answer based on your opinion regarding the product:

	Very	Slightly	Neither	Slightly	Very	
Inexpensive	[]	[]	[]	[]	[]	Expensive
Effective	[]	[]	[]	[]	[]	Ineffective
Useful	[]	[]	[]	[]	[]	Useless
Reliable	[]	[]	[]	[]	[]	Unreliable

Unmarked Semantic Differential Scale

The central line serves as the neutral point:

Inexpensive _____		_____	Expensive
Effective _____		_____	Ineffective
Useful _____		_____	Useless
Reliable _____		_____	Unreliable

Source: Sincero (2012)

With *interval scales*, by contrast, the difference between the categories is of equal distance and can be measured, but there is no true zero point. A common example of an interval scale is calendar years. For example, there is a specific distance between 1776 and 1876, 100 years, yet it makes no sense to say 1776 is 95 percent of the later year. With *ratio scales*, which do have a true zero point, you can calculate the ratios between the amounts on the scale. For example, salaries measured on a dollar scale can be compared in terms of true magnitude. A person who makes $200,000 a year makes twice as much as someone whose salary is $100,000.

In summary, question content, design, and format serve as the fundamental elements in building and executing good surveys. Research has provided some guidance on best practices. For example, some general design and question order recommendations have emerged from research: (a) order questions from easy to difficult, (b) place general questions before specific questions, (c) do not place sensitive questions at the beginning of the survey, and (d) place demographics at the end of the questionnaire to prevent boredom and to engage the participant early in the survey.[28]

It is also recognized that survey responses can be affected by how the question and response categories are presented, particularly ordinal scale questions.[29] There is considerable debate regarding many of the facets of response sets and scales. For example, there is little agreement as to the optimum number of points on a scale. The only agreement is that between 5 and 10 points is good, with 7 considered the optimal number by many researchers.[30] But there is a range of opinions on this issue and on whether extending the number of points to 10 or more increases the validity of the data.[31]

Similarly, while it seems there is general agreement about using an odd number of categories, so as to have a defined midpoint in the scale, other issues such as the inclusion of *don't know* categories[32] remain controversial, as some contend they are used when the respondent means no or doesn't want to make a choice.

Because the primary mode of delivery continues to undergo changes as technology drives our ways of communicating and interacting, questions need to be adaptable over multiple platforms and retain their validity and reliability in mixed-mode designs. While we have a solid research base, we are struggling to see how new technologies such as web-based surveys and smartphone applications (apps) change the dynamics of survey design and administration.

Summary

- The design, format, and wording of questions are extremely important in surveys.
  - Questions form the basic building blocks of surveys
  - Questions have a major impact on measurement error (which is an important component of the total survey error)
  - Question construction has a major impact on how individuals respond
  - Questions must be targeted to answer the research questions
- The relationship between survey mode, survey instrument, and survey questions is important to consider.
  - The method of delivery of the survey (mode), the platform for the questions (survey instrument or questionnaire), and the expressions of words, phrases, images, and so forth used to solicit information (questions) are all interrelated and must be developed in concert.
  - Because mixed-mode surveys are becoming increasingly popular, questions must be designed to be adaptable across different instruments and modes.

- Addressing question reliability and validity
  - ° If a question is reliable we see consistency in responses across different respondents in similar situations.
  - ° If a question is valid it accurately measures what we say we are measuring.
  - ° Questions must have both reliability and validity.
- Addressing participants' willingness to answer questions
  - ° Declining participation rates have focused more attention on ways of improving respondent's willingness to answer questions.
  - ° Motivation to answer questions can be increased or decreased by several factors.
    - ◊ Value of participation to the respondents
      - —Balance of effort needed to respond against benefit of responding
      - —Respect and courtesy shown to participants
      - —Providing incentives
    - ◊ Avoiding questions that are offensive or demeaning
    - ◊ Making questions understandable
    - ◊ Assuring participants that they will not be put at risk of harm (such as violating privacy) by responding
- Key elements of good questions
  - ° Specificity—Addressing the content of information as precisely as possible.
  - ° Clarity—Ensuring that question wording and concepts are understandable to the respondent.
  - ° Brevity—Making the question as short, straightforward, and simply worded as possible.
- Common question pitfalls
  - ° Double-barrel questions
  - ° Loaded or leading questions
  - ° Questions with built-in assumptions
  - ° Double-negative questions
- Open-ended questions
  - ° Can be administered through interviews, other interactive formats, or in self-administered forms;

- ° Are good for exploring topics or gathering information in areas that are not well known;
- ° Allow participants a *blank canvas* to respond, usually in narrative format;
- ° Have some problems with validity because responses may miss question intent;
- ° Are frequently used with follow-up questions or *probes* to get more detail or further information; and
- ° Require more effort on the part of both respondents and researchers.
- Closed-ended questions
  - ° More defined and standardized than open-ended.
  - ° Generally require less effort on the part of respondents and researchers.
  - ° Response categories are restricted and predetermined by researchers.
  - ° Wording and format of response categories must be carefully constructed to ensure that the information required to answer research questions is obtained.
    - ◊ Response categories may be (a) ordered or scalar or (b) unordered or unscalar.
    - ◊ Response categories may be measured at the nominal, ordinal, interval, or ratio level.
    - ◊ Response categories must be mutually exclusive and exhaustive.

Annotated Bibliography

General

- See Chapters 7 and 8 in Robert Groves and associates, *Survey Methodology*, 2nd ed.[33]

Characteristics of Good Questions

- Dillman and his associates' text on surveys[34] provides a great deal of information on the characteristics of good questions including Dillman's 19 principles for good question design.

Reliability and Validity

- For a conversational and easy to understand overview of reliability and validity in survey research, see "Understanding Evidence-Based Research Methods: Reliability and Validity Considerations in Survey Research" by Etchegaray and Fischer.[35]
- For an in-depth review of reliability in surveys, see Alwin's *Margins of Error: A Study of Reliability in Survey Measurement.*[36]

Question Type and Structuring Questions on a Survey

- Ian Brace[37] provides a good overview of question types and the importance of how questions are structured on survey instruments.

CHAPTER 7

Carrying Out the Survey

So far we have talked about a number of different aspects of doing a survey including:

- How to select the cases for your survey (Chapter 2);
- The different types of error that can occur in a survey (Chapter 3);
- Things you need to think about when planning a survey (Chapter 4);
- Different ways of delivering the survey to your sample (Chapter 5); and
- Writing good questions (Chapter 6).

In this chapter, we're going to talk about how you carry out the survey. We're not going to get into the nuts and bolts of doing a survey. There are lots of good books that will do this and we'll mention them in the annotated bibliography at the end of this chapter. Rather we're going to describe the steps that every researcher must go through in carrying out a survey.

Developing the Survey

Let's assume that you want to do a survey of adults in your county to determine their perception of the quality of life. You know that there are certain areas that you want to explore including perceptions of crime and the economy. You want to develop a survey that can be repeated on an annual or biannual basis to track how perceived quality of life varies over time. You're aware of other quality-of-life surveys to which you would like to compare your survey results. What should you do to begin developing your survey?

Looking at Other Surveys

It's often helpful to look at the types of questions that other researchers have used. One place to search is Google (http://google.com) and Google Scholar (http://scholar.google.com). If you happen to be on a college campus that subscribes to the Roper Center for Public Opinion Research (http://www.ropercenter.uconn.edu), consider using iPOLL, which is a database of over 600,000 survey questions. You can search all these search engines by keywords. Entering the words *quality* and *life* will search for all questions containing both words in the question. Often what others have asked will give you ideas of what you might ask.

Focus Groups

Focus groups are another tool that you can use in developing your survey. A focus group is a small group of individuals from your study population who meet and discuss topics relevant to the survey.[1] Typically, they are volunteers who are paid to take part in the focus group. For example, if your study deals with quality of life, you might explore with the focus group what they think quality of life means and which issues such as crime and jobs are critical to quality of life. A focus group gives you the opportunity to discuss the types of information you want to get from your survey with a group of people who are similar to those you will sample.

Cognitive Interviews

A *cognitive interview* is a survey administered to volunteers from your study population that asks them to "think out loud"[2] as they answer the questions.[3] Cognitive interviews give you the opportunity to try out the questions and discover how respondents interpret them and what they mean by their answers. Let's say that one of the questions you want to ask in your survey is "what is the most pressing problem facing the community in which you live?" In a cognitive interview you can ask respondents how they interpret this question. What does "most pressing problem" mean to them? And you can ask them to take you through their thought processes as they think through the question and formulate an answer.

Asking Experts to Review the Survey

When you have a draft of the survey completed, ask *survey experts* to review it and point out questions that might be confusing to respondents as well as other types of problems. Most colleges and universities will have someone who is trained in survey research and willing to review your draft.

Pretesting the Survey

When you think you are ready to try out your survey, select a small number (25 to 40) of respondents from your study population and have them take the survey using the same procedures you will use in the actual survey. In other words, if you are using a telephone survey, then do your *pretest* over the phone. If it's a web survey, then your pretest should be over the web. You probably won't be using these responses as part of your data since you are likely to make changes in the survey based on the pretest results.

Here are some of the things that you ought to look for in your pretest[4]

- How much variation is there in the answers to each question? Questions that don't have much variation will not be very useful when you analyze your data. For example, if you want to explore why some people are concerned about being a crime victim and others aren't and if almost everyone is concerned, then this question doesn't have much variation and there isn't anything to explain. Of course, you can point out that there is near universal concern about being a victim of crime but that's about all you will be able to say. You won't be able to explore why some are more concerned about being a victim than others since there is little variation in how respondents answer this question.
- How many respondents skip certain questions or say they don't know how to respond? No answers and don't knows could be an indication of a problem with the way the question is worded or it could indicate that the question asks for information that respondents can't or don't want to provide.

- Is there evidence of *satisficing*? Some questions require a lot of effort to answer and sometimes respondents look for ways to reduce the burden of answering certain questions. This is what is called satisficing. For example, giving one-word answers to open-ended questions can indicate satisficing. Asking people what is the most pressing problem facing their community requires a lot of effort to answer. Answering with one word such as "crime" or "education" is one way to reduce the burden. We discussed satisficing in Chapter 3. You might want to refer back to that chapter.

- If you are asking respondents to skip particular questions based on their answers to previous questions, did the *skip patterns* work as you intended? For example, you could ask respondents if they are very satisfied, somewhat satisfied, somewhat dissatisfied, or very dissatisfied with their life in general. You might want to ask only those who are dissatisfied to tell you why they are dissatisfied. This requires a skip pattern in the questions. If you're using a telephone or web survey, you can program that skip into the software you are using. If you are using a face-to-face survey, the interviewer will have to be instructed when to skip to the next question. If you are using a mailed survey, the instructions will have to be written in the survey. However you build the skip pattern into your survey did it work as you intended? It's important to check to make sure that the skip patterns are working properly before you begin the actual survey. The pretest is the place to check it out.

- How long did it take for the respondents to complete the survey? Do you think respondents will be willing to spend that much time on your survey? You can ask respondents in the pretest whether the survey took too long to complete.

- If you are using an interviewer-administered survey, did the interviewers report any problems during the survey? Be sure to debrief your interviewers after the pretest.

Pretesting is an essential step in preparing your survey so it is ready for delivery to your sample. Here are some other suggestions for the pretest.

- There are two questions that are always important to ask when preparing a survey. How are respondents interpreting the questions? What do respondents mean by their answers? We talked about the usefulness of cognitive interviews when you are developing your survey. They are just as useful during the pretest. Howard Schuman suggests the following probes. "Could you tell me why you say that?" "Would you explain what you meant by _____?"[5] George Bishop suggests asking respondents to "think out loud" while answering the question.[6]

- Ask respondents to tell you about the problems they encountered while doing the pretest. Were there questions they had difficulty in answering? Were there questions that were confusing?

- If it's possible, record the pretests so you can go back over them with the interviewers and talk about particular questions. These could be audio or video recordings. Remember that you will need to get the respondent's permission to record the interviews.

Administering the Survey—Using Probe Questions

Administering the survey depends in part on your *mode of survey delivery*. In Chapter 5, we talked about the four basic modes of survey delivery—face-to-face, mailed, telephone, and web—and mixed-mode surveys, which combine two or more of these delivery modes. You might want to go back and look at this chapter again and at some of the references mentioned in the annotated bibliography.

One of the most important tasks of survey administration is to clarify the answers of respondents through follow-up questions. These types of questions are referred to as *probes*. There are a number of different types of probes. For example, we could ask respondents to "tell us more" or what they meant by a particular answer. Patricia Gwartney suggests some other probes.[7]

- Silence—Don't be afraid of not saying anything for a few seconds. This can encourage respondents to expand on what they told you.

- Repetition—We could repeat what respondents tell us in their own words to encourage them to expand on their answers.

- Repeating the question—Another type of probe is to simply repeat the question and the response categories.
- Asking for help—Saying that you don't understand the respondent's answer and asking for help is a useful probe. Asking for help can encourage respondents to work with you to clarify an answer.

Some questions are particularly likely to require a follow-up question in order to clarify what respondents tell us. Here are some examples.

- Suppose we ask a respondent "what is the most pressing problem facing your community today?" and the respondent says "crime." We could probe by saying "could you tell me a little more about that?"
- Researchers often want to know a person's race and ethnicity. Often we start with a question such as "would you describe yourself as being Hispanic or Latino?" This could be followed by "what race do you consider yourself to be?" But what do you do if the person says that he or she is German or Italian? One approach is to probe by rereading the question but this time asking them to select their answer from among a set of categories such as White, American Indian, African American or Black, Asian or Pacific Islander, and other. Many surveys allow respondents to select more than one category. There also needs to be a category for "refusal."[8]
- Sometimes we want to know what respondents do for a living. We might start by asking them if they are currently employed and, if they are, by asking "what is your current occupation (or job)?" Some respondents may not give you the information you need. Gwartney suggests the following probes: "What kind of work do you do?" "What is your job title?" "What are your usual activities or duties at your job?"[9]

Probing in Web Surveys

The way in which we probe depends in large part on the mode of survey delivery. Surveys that are interviewer-administered such as face-to-face

and telephone surveys provide the interviewer with considerable control over the use of probe questions. Web surveys are not interviewer-administered but technological advances give the researcher considerable control here as well.

There are some questions that you know will require a probe question. For example, if you ask someone their job title, you will need to follow that up with a question asking about the duties and activities of their job. If you ask people what they think is the most pressing problem facing their community, you might want to follow that up with a probe asking "why do you feel that way?" This type of probe can easily be built into any survey including web surveys.

There are other types of probe questions that depend on what respondents tell you. Pamela Alreck and Robert Settle call these *interactive or dynamic probes*.[10] For example, if respondents give you a one-word answer such as "crime" or "drugs" to the most pressing problem question you would want to ask them to "tell me a little more about that." That's more difficult to carry out in a web survey unless you can identify the specific keywords for which you want to ask a probe question. In addition, you need to be using web survey software that allows you to use this type of probe question.

Probing in Mailed Surveys

Probing is more difficult in a mailed survey. Mailed surveys are not interactive. There is no contact between the interviewer and the respondent unless one provides the respondent with a telephone number or web address that they can use to contact you. Consequently all instructions and questions have to be written out in the survey. This limits you to probes that can be anticipated in advance. If you are asking about the respondent's occupations or jobs, you can include a probe question asking the respondents to tell you about their job's duties and activities. If you are asking about attitudes or opinions on some issue, you can ask them to tell you "why they feel that way." But there is no opportunity for following up on respondents' specific answers. If they tell you that their race is Swedish, you can't follow that up. You have to make your instructions clear and specific enough to make sure that respondents know what you are asking.

Administering the Survey—Record Keeping

Another important part of survey administration is *record keeping*. It's essential to keep good records regardless of the survey delivery mode. But the information that is available for your records will vary by the mode of survey delivery. In an interviewer-administered survey, you might have information about individuals you are unable to contact or who refuse to be interviewed. Each time you attempt to reach a potential respondent, a record must be kept of the result. These are often referred to as disposition codes. You should be sure to record the following information.

- Was the respondent eligible to be part of the survey or ineligible based on whom you were trying to contact or you don't know? If the respondent was ineligible or you don't know, why?
- Were you able to make contact with the respondent? If not, why?
- Was the interview completed? If not, why?

Patricia Gwartney has a detailed list of *disposition codes* for telephone interviews, which could be adapted for face-to-face surveys.[11] You can also look at the disposition codes published by the American Association for Public Opinion Research.[12]

Often respondents are unable to do the interview at the time you reach them and the interview needs to be scheduled for a *callback*. This should be recorded on a callback form. You should attach a *call record* to each survey, which records each contact, the outcome, the date and time of the contact, the interviewer's name, and when to call back along with any other information that the interviewer wants to convey to the next interviewer. If you are doing a phone survey and are using Computer-assisted Telephone Interviewing (CATI) software, the program will create this record for you.

In a self-administered survey, you probably won't have much information about nonrespondents. You may only know that they didn't respond. However, sometimes the person will contact you and indicate why they aren't completing your survey. This could be because they have moved and aren't part of your study population or because they don't have the time or aren't interested or because they have a problem about survey

confidentiality. Be sure to record this information. But at the very least, you need to be able to report the *response rate*[13] for your survey.

Another reason that good record keeping is so important is that it provides a record of the way in which you carried out your survey. For example, when you create a data file, you make decisions about how to name your questions and how you record the responses to these questions. An example is a person's age. You would probably name this question as *age* and record the person's age as a number. But what will you do if a person refuses to answer this question? You might decide to use 98 for any person who is 98 years of age or older and use 99 for refusals. You should record this decision in a permanent file so that you will remember what you did when you come back to this data file after several years. Or you might give someone else permission to use your data sometime in the future and they will need to know how you recorded age. There needs to be a permanent record of the way in which the survey was carried out to enable future use of this survey.

Processing the Data

Coding

If your survey includes open-ended questions, you will probably want to code the responses into categories. Let's consider the question we have been using as an example—"What is the most pressing problem facing your community today?" Responses to this question could be coded into categories such as the economy, crime, education, traffic and transportation, and so on. You will probably want to divide each of these categories into more specific categories such as lack of jobs, violent crime, and property crime. Once you have developed the categories, have two or more people code the data independently so you can see if the *coding* done by different individuals is consistent.

Editing the Data

In addition to coding answers to open-ended questions you will want to review all the answers. For example, let's say that you're doing a mailed survey and you ask an agree–disagree question with the following categories—strongly agree, agree, disagree, strongly disagree. What are you going

to do if someone selects more than one answer? With other survey delivery modes you have more control over the types of answers that respondents give so you would be able to avoid this type of problem. But you still need to edit the data to check for completeness and consistency. You may need to have a category for uncodable and you will definitely need categories for people who say they don't know or refuse to answer questions.

Data Entry

There are several options for *data entry*. You could enter your data directly into a program such as Excel or into a statistical package such as SPSS. If you are using CATI software or web survey software such as Survey Monkey or Qualtrics, the data can be exported into a number of statistical packages such as SPSS or SAS or into an Excel file.

Data Analysis

Data analysis is beyond the scope of this book. There are many good books on statistical analysis and we'll mention some of them in the annotated bibliography at the end of this chapter.

Writing the Report

Writing reports will be one of the topics covered in Chapter 8.

Listening

In an interviewer-administered survey, it's important for the interviewer to be a good listener. Raymond Gorden talks about *active listening* and suggests that interviewers ask themselves several questions as they are listening to the respondent.

- Is it clear what that means?
- Is that really relevant to the question?
- Is the answer complete?
- What does that tone of voice mean?
- Should I interrupt now to probe or should I wait till later?[14]

Gorden also suggests a number of keys to being a good listener.[15]

- "Know your objective." Interviewers should understand what each question is attempting to find out about the respondent and what the purpose of the question is.
- "Pay attention from the beginning." Don't get distracted and miss what the respondent is telling you.
- "Control your urge for self-expression." Don't interject your own thoughts into the interview. Remember it's not about what you think; it's about what the respondent thinks.
- "Listen actively." As the respondents are talking, pay attention to what they are saying. Think about possible probes that you may want to ask.
- "Be patient." Don't rush. Let the respondents tell you in their own words.

Interviewer Training

In interviewer-administered surveys interviewers need to be trained. It's unreasonable to expect them to pick up what they need to know through on-the-job training. Here are some different *training techniques*. A good training program will combine several of these approaches.

Providing Documentation

You will need to provide documentation for interviewers to study and to have available for reference during interviews. These should include:

- Copies of the survey questions with skip patterns.
- List of questions that respondents might ask and suggestions for answering these questions. Questions might include, for example: How did you get my name and address or phone number? How long will it take? Do I have to do it? What's the survey about? What's the purpose of the survey? Who is the survey for? Is what I tell you confidential? There are some excellent examples of handouts on answering respondent's

questions in Don Dillman's and Patricia Gwartney's books on survey research.[16]

- Why people refuse and how you might respond to these refusals. For example, people might respond by saying:
  - I don't have the time to do it.
  - I never do surveys.
  - I'm sick now.
  - I'm not interested.
  - It's nobody's business what I think or do.

  For some of these reasons there's an easy response. For example, if someone doesn't have time to do it now or is sick, you should offer to call back at a more convenient time. If someone says they never do surveys, you should explain why this survey is important and worth their time. Don Dillman and Patricia Gwartney also have examples of handouts on how to handle *refusals*.[17]

- *Interviewer manual* including information on the following:
  - Getting the respondents to participate
  - The structure of the interview
  - How to ask questions
  - How and when to probe
  - What to say when the respondent doesn't understand a question
  - Disposition codes that indicate the result of the contact
  - Scheduling callbacks
  - Time sheets to record hours worked
  - Getting paid*

Practice Interviews

Interviewers should have the opportunity to *practice* the interview before actually starting data collection. A good place to start is to practice interviewing themselves. Have them read through the questions and think about how they would answer and what they might find confusing. Then interviewers could pair off with another interviewer and take turns interviewing each other. They could also interview friends and family.

---

* This list is not meant to be exhaustive. It's only meant to give examples.

Role playing is often a useful training device. Have experienced interviewers play the role of respondents and simulate the types of problems interviewers might encounter. For example, problems often arise when asking questions about race and ethnicity. Respondents often give one-word answers to open-ended questions. These types of difficulties could be simulated in a practice session.

Another useful training tool is to have experienced interviewers work with new interviewers and *coach* them on how to handle difficult problems that arise. Experienced interviewers could listen to practice interviews and then discuss with the new interviewers how they might improve their interviewing technique. If it's possible, record the practice interviews so you can review them and use them as teaching tools.

Survey Participation

One of the concerns of survey researchers is the declining *response rates* that all modes of survey delivery have experienced during the last 35 to 40 years.[18] This has been one of the factors that have led to the increased cost of doing surveys. But the concern is not just over cost. The concern is also that this will lead to increased *nonresponse bias*. Bias occurs when the people who do not respond to the survey are systematically different from those who do respond, and these differences are related to the questions we ask. Increasing response does not necessarily decrease bias. Jeffrey Rosen and his associates note that increasing response rates among those who are underrepresented is what is necessary to reduce nonresponse bias.[19]

We discussed survey participation in Chapter 3 so we're not going to repeat the discussion here. Rather we want to emphasize that declining response to surveys is a serious potential problem since it increases the possibility of nonresponse bias. Take a look at our discussion in Chapter 3 of various theories of survey participation and how you might increase response rates.

Robert Groves and Katherine McGonagle describe what they call a "theory-guided interviewer training protocol regarding survey participation."[20] It starts with listing the types of concerns that respondents have about participating in the survey and then organizing these concerns into a smaller set of "themes." Training consists of:

- "Learning the themes;"
- "Learning to classify sample person's actual wording into these themes;"
- "Learning desirable behavior to address these concerns;"
- "Learning to deliver… a set of statements relevant to their concerns;" and
- "Increasing the speed of performance" so this process can be done quickly.[21]

For example, if the respondent says "I'm really busy right now!" the interviewer might respond "This will only take a few minutes of your time." Basically what the interviewer is doing is *tailoring* his or her approach and response to the respondent's concerns.[22]

Summary

- Tools for developing the survey
  - Focus groups allow the researcher to get a sense of how people feel about the issues covered in the survey.
  - Cognitive interviewing is a way to find out how respondents interpret the questions and what they mean by their answers. One way to do this is to ask respondents to "think out loud" as they answer the questions.
  - Look at other surveys with a similar focus.
  - Survey experts can review your survey and point out problems.
- Pretesting the survey
  - Try out your survey on a small group of individuals from your survey population.
  - Ask respondents in your pretest to talk about the problems they had taking the survey.
  - In interviewer-administered surveys ask the interviewers about the problems they had while administering the survey.
- Administering the survey
  - Probes are follow-up questions that elicit additional information or clarify what the respondent said.

- ° There are many types of probes including the following:
  - ◊ Silence
  - ◊ Repetition
  - ◊ Asking respondents for help in understanding their response
- ° It's critical to keep good records of each attempt to conduct an interview and to keep an accurate record of the ways in which the survey is carried out.
- Processing the data includes coding open-ended responses, editing the data, data entry, data analysis, and writing reports (covered in Chapter 8).
- When an interview is administered by an interviewer, it's essential for the interviewer to be a good listener. Being a good listener is something people can learn to do.
- There are several approaches to train interviewers for face-to-face and telephone surveys.
  - ° Providing copies of the survey and skip patterns, questions interviewers might be asked, how to respond to refusals, and interviewing manuals.
  - ° Practice interviews
  - ° Coaching
- Survey participation
  - ° Survey response rates have been declining for the last 35 to 40 years.
  - ° This increases the possibility of nonresponse bias.
  - ° Increasing the overall response rate does not necessarily decrease bias unless you increase the response rate for those who are underrepresented in the survey.

Annotated Bibliography

- Developing the survey
  - ° Floyd Fowler's *Survey Research Methods*[23] and Robert Groves and associates's *Survey Methodology*[24] discuss focus groups and cognitive interviews.

- Pretesting the survey
  - ° Earl Babbie's *Survey Research Methods*[25] has a good discussion of pretesting.
  - ° Jean Converse and Stanley Presser's *Survey Questions: Handcrafting the Standardized Questionnaire*[26] is another excellent discussion of pretesting.
- Administering the survey

  There are a number of very good books on how to do various types of surveys. Here are some excellent sources.
  - ° Don Dillman's series of three books on survey research
    - ◊ *Mail and Telephone Surveys—The Total Design Method*[27]
    - ◊ *Mail and Internet Surveys—The Tailored Design Method*[28]
    - ◊ *Internet, Mail, and Mixed-Mode Surveys—The Tailored Design Method*[29]
  - ° Patricia Gwartney—*The Telephone Interviewer's Handbook*[30]
  - ° Mick Couper—*Designing Effective Web Surveys*[31]
- Listening
  - ° An excellent discussion of how to be a good listener is Raymond Gorden's *Basic Interviewing Skills*.[32]
- Interviewer Training

  Here are some good references on training interviewers.
  - ° Floyd Fowler—*Survey Research Methods*[33]
  - ° Patricia Gwartney—*The Telephone Interviewer's Handbook*[34]
  - ° Robert Groves and Katherine McGonagle—"A Theory-guided Interviewer Training Protocol Regarding Survey Participation"[35]
- Nonresponse

  These are excellent discussions of nonresponse, nonresponse bias, and increasing response.
  - ° Herbert Weisberg—*The Total Survey Error Approach*[36]
  - ° Robert Groves and his associates—*Survey Methodology*[37]
- Data Analysis

  Data analysis is beyond the scope of this book but here are some excellent references.
  - ° Earl Babbie—*The Practice of Social Research*[38]

- ° Jane Miller—*The Chicago Guide to Writing About Multivariate Analysis* [39]
- ° Your favorite statistics book. If you don't have a favorite statistics book, take a look at *Social Statistics for a Diverse Society* [40] by Chava Frankfort-Nachmias and Anna Leon-Guerrero and *Working with Sample Data* [41] by Priscilla Chaffe-Stengel and Donald N. Stengel.

CHAPTER 8

Presenting Survey Results

How we present our survey results is one of the most important aspects of the entire survey effort. It doesn't matter how well the survey was conceived, designed, or executed, if it's poorly or inaccurately presented, none of the effort that went into the survey will be recognized. Moreover, important findings may never get the visibility or attention they should.

Paul Hague and his colleagues echo this notion in their book on Marketing Research, "These reports should be the crowning glory of the huge amount of time and money that has been invested in the research and yet so often the results are disastrous."[1]

In this chapter, we try to provide you with some guidelines based on our own experience and the expertise of others on how to ensure that survey results are presented well.

In this discussion, we use the term survey *presentations* generically to refer to all forms of reporting including written materials, verbal presentations, and visual materials such as PowerPoint, web graphics, and so forth. As a memory tool and means of providing organization to survey presentations, we present the discussion of survey presentations following a simple three-part acronym, ACE, to help focus on the three major considerations in developing a survey presentation.

A—*Audience*: Who is the audience of the presentation?
C—*Content*: What are the key points that we need to convey?
E—*Expression*: How do we present the survey in a way that is clear, understandable, and complete?

Too often, the expression or appearances of the presentations become the main focus in their development. And too often the dominant themes in constructing presentations become centered on things such as formatting questions, *Is the blue background in the PowerPoint too dark?*

or *Do you think that the color graph will copy well in black and white?* Because these concerns should not be the drivers in the development of presentations and because your material may be presented in a number of different formats and media types, we recommend that the first two key components—the audience and the content—become the first areas to get your attention. Once these areas are addressed, the expression of the information will be a much easier task.

The Audience

You may recall that in Chapter 4 we discussed the three primary stake-holder groups in the survey process: the sponsors, the researchers, and the participants. The sponsor, in addition to identifying the purpose of the survey, the population of interest, the timeline, and the approximate resources available, should specifically indicate what project *deliverables* are expected, such as a report, a presentation to a specific audience, or the submission of data files. There should be agreement on these deliver-ables with regard to the presentation. Should there be a written report? If so, is there a specification of the topics that are to be covered in the report? Is there a requirement to make a live presentation to an executive or administrative group, or perhaps in a more public forum such as a public hearing or a shareholders meeting? Is there to be an online presen-tation, possibly being posted to the organization's website, or a Twitter or Facebook posting? Thus, how the sponsor wants those findings presented should be explicitly stated in the deliverables the sponsor provides in the initial stages of the project. It is important to remember that each differ-ent type of presentation not only requires a different presentation format, but brings different audience considerations into play.

Beyond the explicit conditions for the presentation detailed in the deliverables, there are also implicit specifications for the presentation cen-tered on the sponsor's stated or unstated expectations of what is the most important information and how that information should be presented. Sometimes these implicit expectations closely match the formal specifica-tion of deliverables, but in other situations, the sponsor's real expectations may be very different. For example, the deliverables may call for a detailed report covering all aspects of the survey project, but in conversations with

the individuals in the sponsoring organization they may indicate they're most interested in an executive summary and a 30-minute presentation with an opportunity to ask questions. In this case, if the researchers put in a great deal of effort producing a massive tome on the survey, but fail to deliver a concise, understandable executive summary, or if they created a presentation that was essentially just images grabbed from the report and pasted into a PowerPoint, the sponsors would be very dissatisfied with the presentation and see little value in the survey or its findings. Therefore, in addition to being very familiar with the project's stated deliverables, it is critical for researchers to get to know their audience and what that audience expects regarding the presentation of the survey and its findings.

Knowing Your audience

Too often research presentations, especially reports and PowerPoint type presentations are created with little thought about who will be reading or viewing their content. Like possible compromising photos that ultimately end up posted on the Internet, what they reveal may be expected by one audience, misinterpreted by another, and totally misunderstood by a third. Therefore, it is very important to know and understand your audience. Conversations with sponsors are an essential part of the process not to just understand their expectations, but to gauge how best to present material to them.

The Hidden Lack of Understanding by the Audience

The increasing use of surveys to gather information on different populations in the academic, political, and business realms has created a constant stream of survey results. The fact that surveys have become such a common part of the landscape in information gathering has created a familiarity with surveys to the point of sometimes promoting a false sense of understanding them. Lynn McAlevey and Charles Sullivan aptly note, "The news media almost daily quote from them, yet they are widely misused."[2]

In a study focusing on the understanding of surveys, McAlevey and Sullivan looked at students with prior managerial experience embarking

on an MBA program. What they found was that common sample survey results are misunderstood even by those managers who have previous coursework in statistics. In general, those managers with some statistics background fared no better than managers who had never studied statistics. McAlevey and Sullivan's succinct conclusion was, "In general, we find no difference. Both groups misuse the information substantially."[3] McAlevey and Sullivan put the implications of this hidden lack of understanding about survey methodology into perspective thusly,

> For example, great statistical care may be used to take account of the effects of complex survey design (e.g., stratification and clustering)* on estimates of sampling error. But what is the practical value of this if the primary users have gross misconceptions and misunderstandings about sampling error?[4]

If a lack of understanding prevails in the audience, then presentation emphasis on methodological and design features important to the researchers may be lost on the audience. Perhaps even more importantly, such a disconnect could have a negative impact. Brunt,[5] for example, notes that sample surveys have counterintuitive properties for nonspecialists. Thus, a lack of understanding by the audience creates mistrust of the survey process and in the end a rejection of the findings, basically resulting in the proverbial, "throwing the baby out with the bathwater."

For these reasons, the conversations with the individuals to whom the presentation is directed should not only focus on their expectations, but also assess their understanding of the survey process. This latter area can be somewhat sensitive, particularly for individuals in upper level management who may not want to appear unknowledgeable, especially if the presentation will include subordinates. One way of approaching this is to point out that every specialty has its own set of technical concepts and jargon, and ask them which of these they think it would be helpful to review in the survey presentation. Another way is to split the audience into homogenous groups, as we discuss in the following section.

---

* These were discussed in Chapter 2 on Sampling.

Types of Audiences

In her discussion of survey presentations, Arlene Fink classifies audiences into three categories: nontechnical, technical, and mixed (both nontechnical and technical).[6] The nontechnical audience could best be characterized as:

- Primarily interested in the survey findings;
- Wanting to know if the findings are important, and if so what makes them important;
- Wanting to know how to use the results;
- Not interested in the methodological details including the data collection processes; and
- Not understanding the details of the analysis, and not wanting to see a lot of statistics or an emphasis on data presentation.

By contrast, the technical audience is interested in the details of the study's design and methodology. This audience is characterized as being interested in:

- Why a particular design was used;
- The way the sample was selected;
- How the data collection was carried out;
- Characteristics of the respondents, information about response rates, and details about nonresponse; and
- Details of the analysis, including a detailed data review, and even a discussion of survey error.

The third audience is the one comprised of both nontechnical and technical audiences. From the perspective of preparing or presenting information on a survey project, the mixed audience can be a major problem, but it can also be an asset. The problem is that nontechnical and technical members of a mixed audience will come in with contrasting and sometimes conflicting expectations. One option to deal with the conflicts between the technical and nontechnical is to prepare separate presentations for each. For example, reports can be separated into components (an idea we'll explore later in the chapter) or two separate meetings can

be held, one for the technical audience and another for the nontechnical audience. This notion can also be extended to other logical groupings. For example, separate presentations might be made to different groups based on positions, such as departments or divisions within the sponsor's organization. As Paul Hague and his associates point out, the different members of an audience can also vary by their positions, which create different needs and expectations.

> Typical audiences for [marketing] research reports consist of product managers, marketing managers, sales managers, market research managers, business development people, technical development managers and of course the "C-suite" of top executives. A researcher needs to balance the needs of these groups within the report.
>
> The job responsibilities of the audience will strongly influence the specific intelligence they look for from the report. Sales people want to know specifics such as what each of their customers, and especially potential customers, is thinking and doing. Communications managers are interested in different things, such as which journals people read, which websites they visit and what messages are effective. Technical staff is likely to be interested in which product features are valued.[7]

The value of being confronted by a mixed audience is that it forces those presenting the results to consider topic areas that would otherwise be missed if only a single audience type was involved. For example, let's say that an online healthcare survey sent to a healthcare organization's members had a very low response rate for individuals in the 65 to 75-year-old age range; yet the overall survey results found that members indicated they would benefit from online healthcare information. If a presentation only included a broad overview of the results without a more detailed view of the methodology and response rates, the decision makers might erroneously decide that going forward with an online health awareness campaign would be a great idea, when in reality, such a program would be of little value to this very important demographic in the health organization's membership.

The Content

The principal focus of the presentation will be to answer the research questions and address the study's objectives, which were first identified at the beginning of the study. The content should lead the audience to those results by addressing: (1) why the study was undertaken (introduction and background), (2) how it was designed and structured (methodology), (3) how the information was gathered (data collection), (4) how the data were examined (the analysis), (5) what the findings (results) were, and (6) what the findings mean (summary and recommendations). If this sounds familiar, it's because it follows the traditional structure for structuring and reporting out research—a standard format in academic research. It is commonly presented as:

- Statement of the problem
- Review of the relevant literature
- Methodology
- Data collection
- Data analysis and findings
- Conclusions
- Summary and recommendations

To illustrate, the APA report style, developed by the American Psychological Association and widely used throughout the social sciences, typically divides research reports into seven sections:

- *Title page*—The title of the paper, the names of authors, and the affiliations of the authors.
- *Abstract*—A brief overview of the entire project of about 150 to 250 words.
- *Introduction*—The background and logic of the study, including previous research that led to this project.
- *Method*—Minute details of how the study proceeded, including descriptions of participants, apparatus, and materials, and what researchers and participants actually did during the study.

- *Results*—A detailed statement of the statistics and other results of the study.
- *Discussion*—What the results tell us about thought and behavior.
- *References*—Where to find the work cited in the paper that relates to the presentation.[8]

Of course, not every research report follows this formal format. Some commonly found components may be combined or embedded within another section; for example, a statement of the hypothesis may be included as part of a statement of the problem section. Further, while some survey research reports are structured in a very formal style, particularly when they are presented in academic journals or in formal conference settings, survey reports using a more informal structure are more common, especially when the reports are primarily intended for the sponsor's use. Some of the differences include the fact that in the more informal presentations, a literature review will likely be omitted (unless specifically requested), or if presented, it will be in an abbreviated format, such as quickly reviewing recent surveys similar to the current effort. More informal survey reports and PowerPoint type presentations also tend to have a lower *level of granularity* in the information in the Methodology and Results sections. By a lower level of granularity we mean the level of detail is less. For example, in more informal presentations, the Methodology section may contain only the key points of the design, sampling approach, and data collection. It is also common, particularly with surveys that employ sophisticated methodologies or use consistent methodological approaches across similar surveys, to break the methodology section out into an entirely separate document or to place it in a report appendix.[9] Similarly, the Results section will contain less detail in the body of the report, again placing detailed information such as a comprehensive breakdown of the survey population in an appendix at the end of the report. Finally, in the more informal survey format, an abstract is often not included. However, one component not typically included in more formal presentation formats, but commonplace in today's organizationally sponsored survey presentations is an executive summary.

Executive Summaries

As the title might suggest, *executive summaries* are usually directed at *executives* or decision makers primarily because they might not have time to attend a full presentation or read an entire survey report. Because the executive summary may be the only exposure that some people may get to the survey content and findings, it is important that it presents the survey approach and results information as accurately as possible, and that it captures all the important content of a larger report or presentation. Like the report or live audience presentation, an executive summary should be targeted to the audience. Executive summaries are typically one to three pages long and frequently use a bulleted-type format rather than lengthy narrative discussions. The executive summary usually contains the following elements:

- Brief overview of the survey: Why the survey was conducted, who or what organization (or department within a large organization) initiated the survey, when and where it was conducted, and any other key background points.
- Goals and objectives of the survey: These are lifted from the goals and objectives that were stated at the beginning of the survey. The discussion should include which goals were met (e.g., which research questions were answered and which were not). For those goals and objectives that were not met a brief explanation should be provided as to why not.
- Survey methodology: Again, this section of the executive report will be a much truncated version of the material presented in a full report or presentation to an audience. It should, however, contain all the essential elements of the methodology including who the participants of the survey were, how they were selected (what type of sampling was used), what the sample size and response rate were, what survey mode (e.g., interviewing, online, mail-out questionnaire, mixed) and data collection procedure were used, and what type of survey instrument was used including a brief description of the kinds of questions used. Any particular difficulties or unusual circumstances that might affect the results should also be mentioned here, for example,

if a major blizzard occurred which affected the survey
administration.

- Survey results: This section will briefly explain the major
findings of the survey. The research questions should specifi-
cally be addressed here. Any significant unusual or unexpected
findings should be highlighted.

- Recommendations: This is an optional section. Some spon-
sors want recommendations stemming from the survey to
be included. Others may prefer that recommendations be
omitted.

Executive summaries are typically attached to the full report or are
disseminated in conjunction with the presentation or meeting with a live
audience. However, this isn't always the case. In preparing an executive
summary, it is wise to consider the possibility that this may serve as a
standalone summary and may be the only document about the survey
that some may see. For this reason, it is good to have someone not famil-
iar with the survey read it to make sure it covers the important points.
In this regard, we have sometimes found it very useful to reference more
detailed materials in the executive summary that are available outside the
document, such as sections in a full report including page numbers, so
those reading the executive summary will have access to more detailed
backup material, if it is desired.

Privacy, Confidentiality, and Proprietary Information

As we are constantly reminded by the news media, there is no such thing
as privacy. When conducting a survey, it is important to keep in mind
that the materials produced may have a much longer shelf-life than was
originally intended. Results of a consumer satisfaction survey conducted
years before may end up in a later product liability court case. Similarly,
the release of personal data about survey respondents may violate not only
ethical (and sometimes legal) standards, but could ultimately cause harm
to the individuals who were willing to take part in the survey. Thus, it is
crucial to keep in mind that certain types of information collected during
the course of a survey may be confidential, anonymous, or proprietary.

There are federal regulations that protect individuals from harm during the course of research—these regulations are referred to as *human subject protections*. These protection regulations were developed, in part, due to past abuses of individuals during research, such as the now infamous Tuskegee syphilis clinical research.[10] Federal regulations pertaining to the protection of human subjects can be found online[11] and a review of these regulations related to survey research can be found on the American Association for Public Opinion Researchers (AAPOR) website.[12]

There are three points in the survey process and its presentation where the privacy, confidentiality, and data proprietary issues become particularly relevant. The first occurs in terms of who has access to the data collected. If personal data, such as a respondent's name, birth date, contact information (address and phone number), social security number, and so forth, is obtained, then there is a duty by the researchers to notify the respondents that the data is being collected and how it will be used, and to make a diligent effort to keep that information from becoming public. If the survey participants are told their information will be kept in strictest confidence, then both the researchers and the sponsor have an obligation to uphold this promise. In this regard, it is also important to make sure the researchers and those employed by research firms have signed *data confidentiality agreements*, which detail the requirements for maintaining data confidentiality and the obligations of those who have access to the data.

Second, when survey data is provided to sponsors as a deliverable it should be *de-identified*, meaning that personal or confidential information should be removed. That way, if the data somehow becomes public at a later time, no one will be able to connect the specific respondent's personal information with that individual's responses on the survey. This safeguard is particularly important with surveys that focus on sensitive topics such as employee job satisfaction. Similarly, it is common practice to aggregate individual survey responses in reporting results, so data are only reported at the group rather than individual level. For example, in reporting survey results, only the percentage of respondents who checked a particular response category on a close-end question is reported.

Third, there is also the issue of the proprietary nature of the data collected, the techniques used for data collection, and the methods of

analysis. Both the sponsor and the researchers have vested proprietary interests at stake. The sponsors may have provided the researchers with proprietary information such as contact information for clients taking part in a social program, a customer database, or confidential information about marketing a new product. If the researchers are in-house, this is less of a problem than when the survey is being provided through contract. In Chapter 4, we indicated that the sponsor *owns* the survey, but the extent and terms of this ownership need to be agreed upon at the beginning of the process. If not addressed beforehand, these problems can become major challenges in the presentations of the survey. For example, a sponsor may want a full disclosure of the sampling methodology including the algorithms used to weight the survey responses in the final report. The researchers may decline to provide this because they developed it by extracting and statistically manipulating population data obtained from other sources. Similarly, if researchers were to include proprietary information about a healthcare provider's intent to expand services in a particular geographic location when making a live audience survey presentation, without first obtaining the sponsor's permission, a serious problem could emerge. Again, good communication between the researchers and sponsors about the materials to be included in written reports, executive summaries, live audience meetings, and online postings of materials is essential to ensure that there is agreement about proprietary content, timing of the presentation, and ultimate ownership of materials.

The Expression

We use the term expression to refer to the way in which the content of the survey is presented or delivered. Just as the mode of delivery is the platform for getting the survey instrument to participants (you may want to review the discussion on mode of survey delivery in Chapter 5), the expression of the survey findings is the platform upon which the content is delivered. Before we turn to specifics of expression, there are three important reminders that should be kept in mind when preparing a survey presentation. First, the presentation should focus on substance not style. If the content is not well-defined or organized, no matter how well it is presented, the audience will not get the intended information.

Second, it is important to remember that the content of the survey remains the same irrespective of whether it's presented in a formal report, an executive summary, a PowerPoint presentation, or in other venues. While each of these different types of presentations shapes how the information is formatted, the level of detail presented, and the length of the presentation, the fundamental elements to be presented are the same. In a good *checks-and-balances* review of a survey report, PowerPoint, and so forth, it is good to ask yourself, what would I change if I were completing this in a different format? A third point to keep in mind is that a survey presentation is a reflection of the researchers and sponsors. If it is poorly constructed, has grammatical or spelling errors, has inappropriate language, or sets an improper tone, those problems will be attributed not just to the survey but also to the organizations and individuals who commissioned, created, and carried it out.

Presenting Data

Presenting survey data is one of the most difficult parts of presentation. Nothing will cause an audience's eyes to glaze over more quickly than PowerPoint type slides with row after row of numbers presented on large tables. Similarly, a report with page after page of tables or figures displaying survey statistics is one of the best tonics around for sleeplessness. As Hague and his colleagues note, "Data are the problem. Often there is so much information it is difficult to work out what to include and exclude, and making sense of it is not so easy."[13] Unfortunately, sometimes researchers present massive tables filled with data simply because it's easier than spending the time and effort to distill the information down to a summary level. Summary presentation, of course, does not mean that the detailed data should not be available (with a reference in the report body or at a PowerPoint presentation to the appropriate report appendix or separate document where it can be found) for those who want to dig deeper or verify the accuracy of the summary. However, if you are confronted with statements in reports or meetings such as, "As you can see from the percentages in the third column of the second page of Table 2…," a red flag should go up in your mind. If the information was important enough to point out in raw form, then why did those

presenting the material not take the time to synthesize and explain this important information?

Text, Graphs, and Tables

Synthesizing information and providing commentary on statistics brings meaning to the data collected in surveys. "[W]riting commentary to accompany statistics should be approached as 'statistical story-telling' in that the commentary needs to be engaging to the reader as it provides an overview of the key messages that the statistics show."[14] It is this synthesis and commentary that provides both context for the data presented and a connection between different pieces of the results. What then is the best way to present data? Basically, information can be presented in text, table, or graphic form, and will "generally include all three approaches as this assists in ensuring that the wider user base is catered for."[15] Again, selecting the approach to presenting the information should be based on the needs of the audience.

Text is used for commentary around and in summary of numerical or quantitative results. It is a good way to point out particularly significant statistical findings, which might be lost in the numbers reported in a table or depicted in a graph. It might be considered the train that carries the *statistical story* mentioned earlier. For example, a conclusion might point out that "More than 95 percent of the respondents said they were 'very satisfied' with the service they received," or "There were statistically significant differences between men and women as to how satisfied they were with the service they received." However, using only text to express large numbers is both time-consuming and requires a great deal of effort by the audience to decipher. For example, text saying, "The initial survey was mailed to six-thousand, five-hundred individuals," is much harder to grasp than saying, "The initial survey was mailed to 6,500 individuals." The APA style manual, mentioned earlier in this chapter, uses the convention of spelling out numbers under 10, but using Arabic numerals for values greater than nine.[16] Similarly, trying to relate a lot of numeric data in a sentence will cause the point to become muddied and vague. Consider the following text description,

According to the survey findings, almost no respondents (.9%) were "very dissatisfied" with the service they received, but approximately 35% of the respondents were "dissatisfied," while 32% were "satisfied," and another 32% reported they were "very satisfied."

It may take a couple of readings to pick up the idea that a negligible percent of respondents said they were *very dissatisfied* with the program, and that there was little difference between the respondents on the remaining response categories. In this case, a graph might be a simpler and more straightforward way to point out the similarities between responses given in three categories, as illustrated in Figure 8.1.

It is important here that you not get the impression that all text in survey presentations must be in the briefest form possible. Trying to be too brief in text can have an equally bad effect as trying to provide too much information. Essentially, while the old expression, "Brevity is the soul of wit," may be true, we would argue that "Too much brevity is the soul of a half-wit." The point we are trying to make with this rather bad joke is that if wording becomes too general or leaves out too much detail, it can become as meaningless and unintelligible as trying to describe a table worth of data in a sentence. Take the following sentence, which might be found in a report, as a bullet point in an executive summary or on a PowerPoint slide. "The change in customer satisfaction was tremendous!"
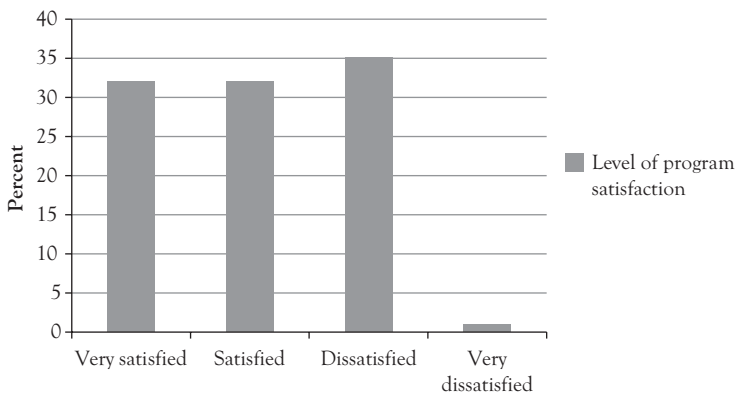
Figure 8.1 Responses to survey question "Are you satisfied with the program?"

The sentence is meaningless because the terms used are vague and abstract. Due to a lack of detail, the audience wouldn't know whether the *change* in satisfaction was up or down—a small detail that could have major consequences. Similarly, the word *tremendous*, while having an emotional connotation, doesn't convey any real sense of magnitude. What do you think should be done to improve the wording?

As the example presented earlier demonstrates, graphic representations of data can be used to present or relate complex findings as an alternative to descriptions using text, or to further illustrate summary points made in the text. A simple visual chart or graph is an easy way not only to provide precise data values, but also to show relationships between different pieces of data. This feature is particularly useful in PowerPoint presentations, where reading extensive material on a slide may be difficult and unproductive. While an extensive review of designing and creating is beyond the scope of our discussion here, there are a few basic ideas you should keep in mind when considering graphs and charts in your presentation. (There is a considerable amount of research available on this, and we provide a couple of good references in the Annotated Bibliography at the end of the chapter.)

One of the problems that we face in creating good graphics for presentations is that our graphic software has grown so sophisticated that it is easy to create both an attractive and a bad graphic at the same time. The first fundamental rule of creating graphs and charts is that they should be appropriate to the type of data they are displaying. For example, categorical information is best displayed in graphics that capture distinct units, such as a bar chart. (Figure 8.2 provides a simple illustration.) A common mistake is connecting categorical data in a trend graph, which is more appropriate for showing change over time (see Figure 8.3). Similarly, presenting data that represents percentages should be displayed in a format that shows portions of a total whole (usually 100 percent). Pie charts and stacked bar charts are effective graphs for this purpose (see Figures 8.4 and 8.5).

Graphs should be simplified as much as possible. Years ago, Edward Tufte, one of the premier researchers in the field of the visual presentation of data, coined the term "chartjunk"[17] to describe elements that appear in graphs that distract the reader from understanding the content in the graph. He suggested that, if you create a figure, you should maximize the
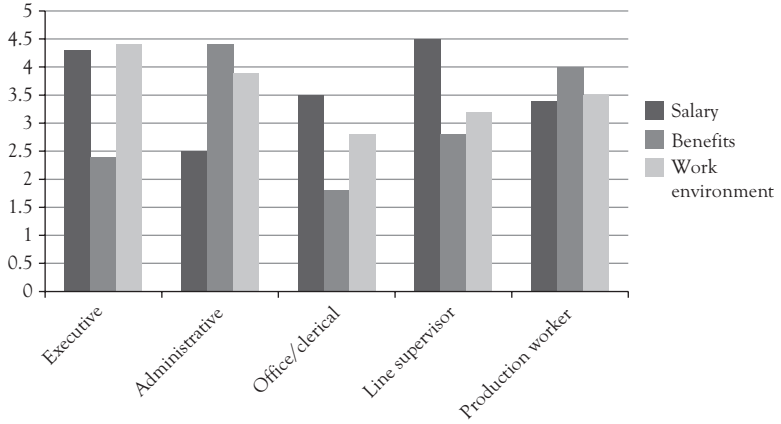
Figure 8.2 Employee satisfaction by position level (bar chart representation)
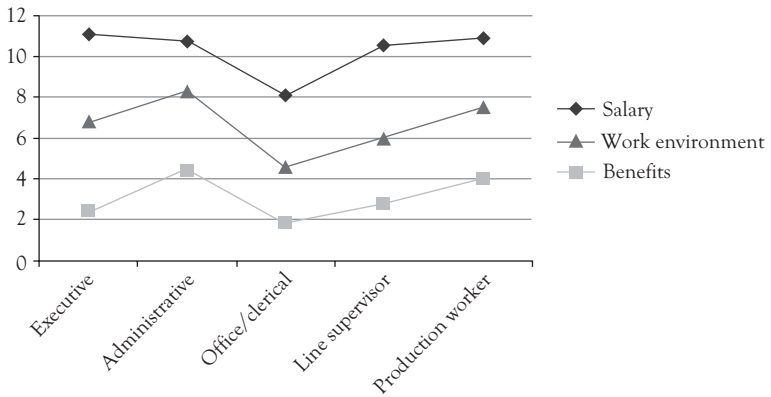
Figure 8.3 Employee satisfaction by position level (line graph representation)
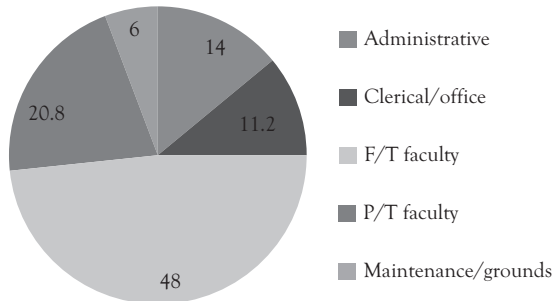
Figure 8.4 Percent of college workforce
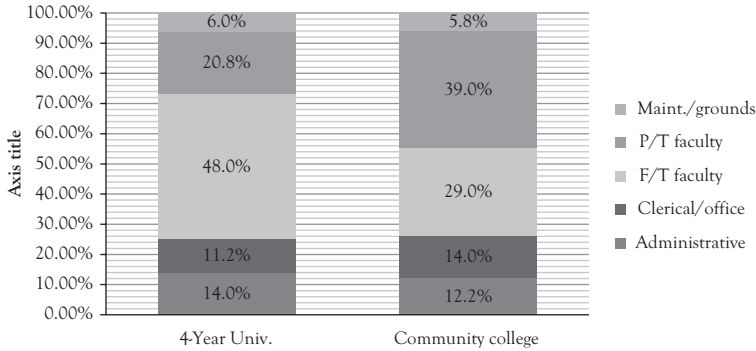
Figure 8.5 Comparison of four-year university and community colleges by position type
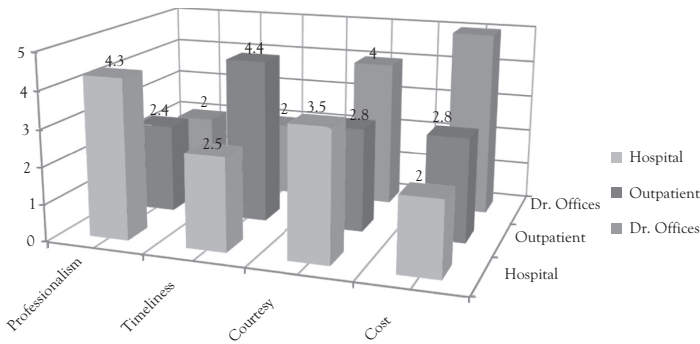
Figure 8.6 Premier healthcare—facility ratings

amount of relevant information and minimize anything that would distract the reader. Tufte identified a principle he referred to as maximization of the data-to-ink ratio—graphs become more useful when the amount of data grows and the amount of ink decreases.[18] Tufte's criticisms of extraneous graphic objects or tools inserted into graphs included the use of unnecessary and distracting lines and background grids. He also objected to patterns and visual features, such as three-dimensional depictions of data that was only two-dimensional.[19] Figure 8.6 shows a hypothetical *worst-case* scenario in terms of the clutter that can make a graph unreadable, while Figure 8.7 displays the same information in a much cleaner graphic.

Julian Champkin perhaps captured the essence of creating good graphics best when he said, "The mark of a good graphic is that the user reads the diagram, not the caption. Information should be in the picture, not in the words underneath."[20]
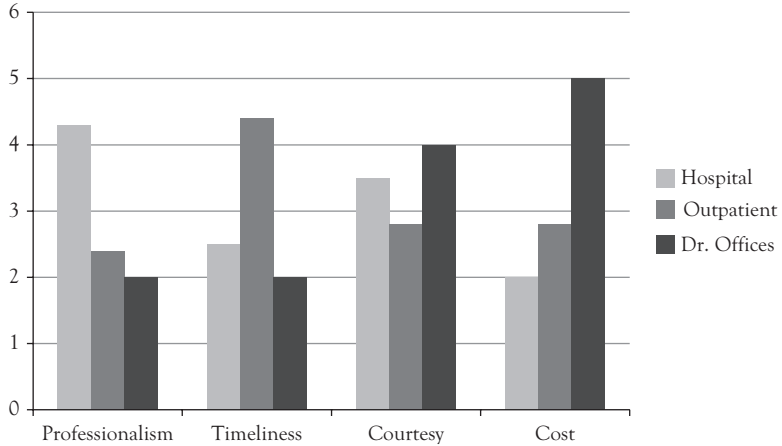
Figure 8.7 Premier healthcare—facility ratings

Presenting tables and presenting figures share much in common. Tables are a convenient way to report lots of data in an organized and succinct manner. They are particularly helpful in presenting the details of the data. For example, numbers containing decimal points are difficult to read in text, particularly if you are presenting a series of numeric values. In Western countries, tables are commonly read top to bottom, left to right. The fact that there are columns and rows in a table makes it convenient to provide data both across and within different number sets. For example, if you want to compare survey responses across three different cities, then it is relatively easy to construct a table with three columns, one for the responses from each city. By then looking across the rows of the three columns, you could compare the responses to a particular survey question across the three cities. However, within the same table, you could also look down a particular column and see how responses to different question areas compared within a particular city. It is this matrix quality that allows easy access to specific pieces of data. In the illustration, the intersection cell of the table would contain a specific value on a certain survey item for a specific city. We typically only create tables in two dimensions for presentations, but hypothetically, if you had a good way to display a third dimension (think Rubik's cube), you could find a particular piece of data using the intersection of three different variables, say the response (1) to a specific question, (2) by males, and (3) within a specific city. However, because we typically only present two-dimensional tables, what we usually do is create a *step-down* table to display this

relationship. By this we mean that we would take our original table that provided the responses to a particular question by cities, then break it down into two sub-tables, in this case, one for males and one for females. While this works fairly well for tables containing only a small number of categories, imagine the difficulty of creating such tables if you were looking at data across some dimension such as income level or race and ethnicity. It is for this reason that we typically move the presentation of such data into a summary form using statistics. Those summary statistics can then be put into the narrative in a report, or presented with a PowerPoint, with an explanation. In our illustration, for example, our summary could then read something along the lines, "Males in Chicago had (statistically) higher satisfaction levels with regard to service than males in either Atlanta or Phoenix. However, there were no significant differences in satisfaction levels among females in the three cities."

As it was with graphs and charts, the use of text in conjunction with tables is a very important part of the presentation of survey research. The emphasis in tables must be on clarity and simplicity. Granularity, in terms of the detail presented, is an important consideration in the inclusion of tables in presentations. Neither PowerPoint presentations nor executive summaries lend themselves to great detail, and for this reason, it is wise to leave tables to the written report, or in the report appendixes or to a separate methodology and data report.

In summary, it is the integration of the presentation narrative, graphics, and tables that tells the story of the survey results. The fundamental content is the foundation of the presentation, but if it is delivered in a way that doesn't meet the audience or members' needs and expectations, if it doesn't interest them, and if it doesn't create an understanding of the survey, then all of the work and effort of the survey will likely just gather electronic dust on some hard drive or network server.

Summary

- Presenting survey findings is one of the most critical aspects of the survey process, yet it is often given only passing attention.
  - Presentations reflect the amount of work that went into the survey.

- ° If the presentation is poorly done, lacks clarity, or is not well organized, important information from the survey isn't conveyed to the sponsors and other intended audiences.
- The sponsor should identify the deliverables at the onset of the survey project, which will determine what form the presentation takes and how it is to be delivered.
  - ° Commonly, survey results are provided in the form of reports, executive summaries, PowerPoint presentations to audiences, materials posted to website, or through feeds on social media. Details of the methodology, statistical analysis, and data are frequency placed in appendices to the full report, or are presented in a separate methodology report.
- The acronym ACE can be a helpful tool in focusing on the three major components to be considered when preparing a presentation.
  - ° A—Audience: Who is the audience of the presentation?
  - ° C—Content: What are the key points that we need to convey?
  - ° E—Expression: How do we present the survey in a way that is clear, understandable, and complete?
- Knowing your audience is a critical part of developing an appropriate presentation. Maintaining contact with the sponsors and designated audiences helps you tailor the presentation both in content and form.
  - ° What are the audience's needs and expectations?
  - ° Is the audience technical, nontechnical, or mixed?
  - ° What level of understanding does the audience have of the survey process, methods, analysis, and so forth?
- The content serves as the foundation for the presentation. The content should address the following:
  1. Why the study was undertaken (introduction and background)
  2. How it was designed and structured (methodology)
  3. How the information was gathered (data collection)
  4. How the data were examined (the analysis)
  5. What the findings (results) were

6. What the findings mean (summary and recommendations)
- The content can be presented in a formal fashion, or may be less formally presented depending on the audience's needs and expectations.
  - ° Today, presentation of survey results tends to be done in a less regimented structure, especially when surveys are commissioned by organizations for internal purposes.
    - ◊ More informal styles usually mean omitting certain elements characteristically seen in formal reports such as an abstract, literature review, and detailed presentation of the methodology. However, today's survey presentations typically contain an executive summary, which was not part of the traditional format.
    - ◊ Executive summaries may be written in a narrative style, but typically make extensive use of bulleted formats.
  - ° Contents presented as PowerPoint presentations or executive summaries have a lower level of granularity, meaning the level of detail is less.
- Privacy, confidentiality, and proprietary information
  - ° It is important to keep in mind that certain types of information collected during the course of a survey may be confidential, anonymous, or proprietary.
    - ◊ There are federal regulations that protect individuals from harm during the course of research, including both the collection and presentation of survey data. These regulations are referred to as human subject protections.
    - ◊ Researchers and sponsors have an obligation to survey participants to ensure that commitments regarding the protection of their private and confidential information are kept, and that participants are not exposed to risk because of the disclosure of such information.
  - ° Sponsors are considered the *owner* of the survey, but sponsors and researchers should have clear agreements made at the beginning of the project as to the extent and terms of this ownership, particularly with regard to data and proprietary information.

- The expression component of the presentation refers to how the survey findings are disseminated.
  - ° The focus in the presentation should be on substance not style.
  - ° Content remains the same irrespective of the way it's delivered.
  - ° The survey presentation is a reflection of the sponsors and researchers.
- Presenting survey data is one of the most difficult aspects of survey presentation.
  - ° Presenting survey data without commentary and summary is not good practice.
  - ° Synthesizing and providing commentary on survey statistics brings meaning to the collected data; it is essentially statistics story telling.
  - ° Data can be presented in text, graphic, or table formats, but presentations usually include all three because it meets the needs and expectations of a broader audience.
- The strength of text or narrative presentation is in providing explanation, commentary, or summary of numerical or quantitative data, not for actually presenting the data. Text should be simple, clear, and brief, but not so brief that the wording becomes meaningless or unintelligible in terms of describing the data.
- Graphics can be used to present or relate complex findings as an alternative to lengthy descriptions in text. Graphics should be appropriate to the kind of data they're displaying, and should be simplified as much as possible.
- Tables can be used to display a lot of data in an organized and succinct manner. Tables are particularly helpful in presenting detailed numeric results, but are not typically suitable for executive summaries and PowerPoint presentations.

Annotated Bibliography

General

- Arlene Fink as part of *The Survey Kit* series provides a short book, *How to Report on Surveys*[21], that provides an easy to understand overview of important elements in the survey report.

- A very helpful ancillary resource is *The Facts on File Guide to Research*[22] by Jeff Lenburg. This book covers a wealth of research sources as well as information on different writing formats.

Audience Considerations

- Patrick Forsyth provides a good discussion of the interface between the audience and the written research report.[23]

Presentation Elements

- For a very good perspective on the elements of report technical writing see Lutz Hering and Heike Hering's *How to Write Technical Reports*.[24]

Narrative, Charts, Graphs, and Tables

- See J. Miller's *The Chicago Guide to Writing About Numbers*.[25]
- Pamela Alreck and Robert Settle provide information about formatting different types of tables and graphs to display survey questions responses, and show how to write around the visual presentation of data.[26]
- Edward Tufte's classic works on the visual presentation provide excellent insights into some of the *dos* and *don'ts* of presenting data in visual form.[27]

Notes

Chapter 1

1. Monette, Sullivan, and DeJong (1998), pp. 22–24.
2. Deutscher (1973), p. 12.
3. Riley (1963), p. 184.
4. Liebow (1967).
5. Liebow (1967), p. 137.
6. Liebow (1967), p. 140.
7. Liebow (1967), pp. 144–5.
8. Liebow (1967), p. 145.
9. Liebow (1967), p. 145.
10. Deutscher (1973), pp. 13–22.
11. LaPiere (2010), p. 9. We cited the most recent reprint of this article. The original article appeared in 1934 in *Social Forces*. See LaPiere (1934).
12. LaPiere (2010), p. 9.
13. Schuman (2008), pp. 132–6.
14. Schuman (2008), p. 132.
15. Schuman (2008), p. 133.
16. Schuman (2008), p. 136.
17. Schuman (2008), p. 136.
18. Singer, Van Hoewyk, and Neugebauer (2003), p. 368.
19. Singer, Van Hoewyk, and Neugebauer (2003), p. 379.
20. Webb et al. (1981), p. 35.
21. Whyte (1975).
22. Whyte (1975), p. 375.
23. Whyte (1975), p. 390.
24. Gorden (1987), p. 63.
25. Gorden (1987), p. 64.
26. Gorden (1987), p. 64.
27. Dillman, Smyth, and Christian (2009), p. 3.
28. de Leeuw (2005), p. 233.
29. de Leeuw (2005), p. 233.
30. Dillman, Smyth, and Christian (2009), pp. 4–6.
31. Dillman, Smyth, and Christian (2009), p. 5.
32. Tourangeau and Plewes (2013), p. 24.
33. Babbie (2013), pp. 187–8; Bradburn and Sudman (1988), pp. 19, 22.

34. Babbie (2013), pp. 188–9; Bradburn and Sudman (1988), pp. 28–30.

35. Blumberg et al. (2011), p. 1. This report indicates for the time period from July 2009 through June 2011 approximately 24 percent of adults were in cell-phone-only households. When this is broken down by state, the estimates vary from a low of approximately 13 percent in Rhode Island and New Jersey to a high of approximately 46 percent in Arkansas. By 2012, the percent had reached almost 40 percent nationally. See Pew Research Center (2014).

36. Riley (1963).

37. Babbie (2013).

38. Miller and Salkind (2002).

39. Deutscher (1973).

40. This article was reprinted in 2010. See LaPiere (2010). The original article appeared in *Social Forces* in 1934. See LaPiere (1934).

41. Whyte (1943).

42. Whyte (1975).

43. Gorden (1987).

44. Webb (1981).

Chapter 2

1. Research focusing on the relationship between religiosity and substance use is extensive and has been summarized by Geppert and her colleagues (Geppert, Bogenschutz, and Miller 2007).

2. Frankel and Frankel (1987).

3. Neyman (1934).

4. Tippett (1927).

5. Sudman (1983), p. 165.

6. Such research dealing with the confidentiality of healthcare data and patient privacy is strictly regulated by the U.S. Dept. of Health and Human Services under the Health Insurance Portability and Accountability Act of 1996 (HIPAA), which covers privacy, security, and notification rules for patient health information including medical records. See the U.S. government website (U.S. Department of Health and Human Services 2014a).

7. See for example Earl Babbie's classic text *The Practice of Social Research* (2013).

8. Fowler (1993).

9. Fink (1995), p. 44.

10. Brick (2011).

11. Baker et al. (2013), p. 91.

12. Berman (2013).

13. Hartford (2014).
14. Babbie (2013).
15. Frankel (1983).
16. Sudman (1983).
17. Stat Trek (2014).
18. Daniel (2012)
19. Baker et al. (2013).
20. Engel et al. (2014).

Chapter 3

1. Weisberg (2005), p. 18.
2. Weisberg (2005), pp. 22–23.
3. For those that like history, Weisberg (2005, 13) notes that the term, total survey error, "first appeared in a book title involving health surveys" written by Anderson, Kasper, and Frankel (1979).
4. Biemer (2010), p. 817.
5. Dillman, Smyth, and Christian (2009); Salant and Dillman (1994); Weisberg (2005).
6. Weisberg (2005, 19) also considers comparability effects which he defines as "the differences between survey results obtained by different survey organizations or in different nations or at different points in time." We're not going to consider comparability effects in this book.
7. Babbie (2013), pp. 212–13.
8. Babbie (2013), pp. 196–7.
9. Babbie (2013), p. 206.
10. Random-digit dialing is more complicated but this captures the essence of the approach. For a more detailed description of random-digit dialing see Pew Research Center (2014).
11. Roth, Han, and Montaquila (2013), p. 1.
12. McNabb (2014), pp. 92–94.
13. See Valliant, Mercer, and Hicks (2014) for some interesting examples of address-based sampling.
14. Marsden (2012), pp. 370–1.
15. Dillman, Smyth, and Christian (2009), pp. 49–52.
16. Lists of the population are not always available. You may have to construct the list or use creative solutions such as this one suggested by Dillman, Smyth, and Christian (2009), p. 52.
17. Dillman, Smyth, and Christian (2009), p. 22.
18. Dillman, Smyth, and Christian (2009), p. 22.
19. Groves et al (2004), pp. 176–7.

20. Tourangeau and Plewes (2013), p. 24.
21. See Tourangeau and Plewes (2013), pp. 9–12 for various ways of calculating response rate suggested by the American Association for Public Opinion Research (AAPOR). APPOR defines response rate as "the number of complete interviews with reporting units divided by the number of eligible reporting units in the sample" (2011, 5).
22. de Leeuw (2008), p. 128.
23. Peytchev, Peytcheva, and Groves (2010), p. 319.
24. Tourangeau, Groves, and Redline (2010), p. 429.
25. Holmes and Schmitz (1996), p. 231.
26. Holmes and Schmitz (1996), p. 231.
27. Groves, Cialdini, and Couper (1992), pp. 477–9.
28. Dillman, Smyth, and Christian (2009).
29. Cialdini (2001).
30. Groves, Cialdini, and Couper (1992), p. 480.
31. Groves, Cialdini, and Couper (1992), p. 482.
32. Groves, Cialdini, and Couper (1992), p. 483.
33. See, for example, Groves, Singer, and Corning (2000); Petrolia and Bhattacharjee (2009); Porter (2004).
34. See Petrolia and Bhattacharjee (2009); Singer, Van Hoewyk, and Maher (2000); Tourangeau and Plewes (2013),
35. Dillman, Smyth, and Christian (2009), p. 242. See also Dillman (2000).
36. Weisberg (2005), p. 18.
37. Weisberg (2005), p. 19.
38. Schuman and Presser (1981), pp. 276–8. See also Schuman (2008), pp. 110–12.
39. Schuman (2002), pp. 44–45.
40. Bickart, Phillips, and Blair (2006), p. 170.
41. Bickart, Phillips, and Blair (2006), p. 171.
42. Smith (1987), p. 75.
43. Huber and Paris (2013), p. 386.
44. Schuldt, Konrath, and Schwarz (2011), p. 115.
45. Whitmarsh (2009), p. 410. Human causes include such things as "pollution, carbon dioxide and other greenhouse gases, CFCs, fossilfuel consumption, cars/traffic fumes, and overuse or misuse of earth's resources."
46. Schuman (2002), pp. 43–45.
47. Box-Steffensmeier, Jacobson, and Grant (2000), p. 260.
48. Box-Steffensmeier, Jacobson, and Grant (2000), p. 261.
49. Box-Steffensmeier, Jacobson, and Grant (2000), p. 265.
50. Box-Steffensmeier, Jacobson, and Grant (2000), pp. 268–9.
51. Moore (2002), p. 82.

52. Moore (2002), p. 83.

53. Moore (2002), p. 83.

54. Moore (2002), p. 84.

55. Moore (2002), pp. 83–84.

56. See also Schuman (2008) for a discussion of order effects.

57. Cannell, Miller, and Oksenberg (1981).

58. See Krosnick (1991) for an excellent discussion of various forms of satisficing.

59. Galesic et al. (2008), p. 892.

60. Galesic et al. (2008), p. 892.

61. Krosnick (1991), p. 221.

62. Heerwegh and Loosveldt (2008).

63. Holbrook, Green, and Krosnick (2003).

64. Krosnick et al. (2002), p. 396.

65. Duff et al. (2007).

66. Streb et al. (2008), p. 78.

67. Kreuter, Presser, and Tourangeau (2008), p. 859.

68. Schuman and Converse (1971).

69. Hatchett and Schuman (1975), p. 525.

70. Anderson, Silver, and Abramson (1988), pp. 53–54.

71. Finkel, Guterbock, and Borg (1991), p. 326.

72. Davis and Silver (2003), p. 40.

73. Kane and Macaulay (1993), p. 1.

74. Johnson and Parsons (1994), p. 90.

75. Pollner (1998).

76. See Anderson, Silver, and Abramson (1988), p. 54.

77. Bishop (1992).

78. Schuman (2008), pp. 81–88.

79. Schuman (2002), pp. 43–45.

80. Duff et al. (2007), p. 67.

81. Streb et al. (2008), p. 80.

82. Streb et al. (2008), p. 81.

83. Ye, Fulton, and Tourangeau (2011), p. 349.

84. Holbrook, Green, and Krosnick (2003).

85. Preisendorfer and Wolter (2014).

86. McDonald and Thornburg (2012).

87. Currivan et al. (2004), p. 559.

88. Heerwegh and Loosveldt (2008).

89. Rosen et al. (2014), p. 2.

90. Rosen et al. (2014), p. 2. Rosen and his associates discuss how they targeted *low-propensity* individuals in their sample.

91. Weisberg (2005).

92. Dillman (1978, 2000); Dillman, Smyth, and Christian (2009).

93. Frankfort-Nachmias and Leon-Guerrero (2011).

94. Chaffe-Stengel and Stengel (2011).

95. Babbie (2013).

96. Kish (1967).

Chapter 4

1. Alreck and Settle (2004), p. 1.

2. U.S. Census Bureau (2014).

3. Alreck and Settle (2004).

4. Groves et al. (2009).

5. American Consumer Opinion (2014).

6. QuestionPro (2014).

7. Survey Police (2014).

8. In 2006, the ESOMAR Research Organization, a group focusing on the quality and standards of marketing research, published a paper entitled "26 Questions to Help Research Buyers of Online Samples" that provides a helpful guide to those considering investing in an online survey panel.

9. Brace (2013), p. 200.

10. Brace, Nancarrow, and McCloskey (1999).

11. CASRO (2014).

12. AAPOR (2014b).

13. Snijkers, Haraldsen, and Jones (2013), p. 128.

14. Snijkers, Haraldsen, and Jones (2013).

15. Alreck and Settle (2004), p. 39.

16. Kennedy, Tarnai, and Wolf (2010), p. 575.

17. Over the years Don Dillman has conducted extensive research on the factors that impact survey response rates, including considerable focus on the use of incentives to improve participation rates. Certain types of monetary and nonmonetary incentives consistently improve the participation rates (Dillman 2009).

18. Hague, Hague, and Morgan (2013), pp. 47–56.

19. Snijkers, Haraldsen, and Jones (2013).

20. Alreck and Settle (2004).

21. Brace (2013).

22. Oldendick (2012).

23. Hundepool et al. (2012).

24. AAPOR (2014a).

Chapter 5

1. Dillman, Smyth, and Christian (2009), pp. 2–10.
2. Dillman, Smyth, and Christian (2009), p. 6.
3. Dillman, Smyth, and Christian (2009), p. 6.
4. Dillman, Smyth, and Christian (2009), pp. 7–8.
5. Tourangeau and Plewes (2013), pp. 68–69.
6. Dillman, Smyth, and Christian (2009), p. 11.
7. Richardson, Dohrenwend, Klein (1965).
8. Richardson, Dohrenwend, Klein (1965), p. 45.
9. Beatty (1995) has an excellent history of what he calls the "standardized/non-standardized interviewing controversy." He cites Paul Lazarsfeld's (1944) article as a classic attempt to mediate between these two positions.
10. Gwartney (2007), pp. 188–203.
11. Conrad and Schober (2000), p. 4.
12. Conrad and Schober (2000), p. 1.
13. Conrad and Schober (2000), p. 7.
14. Conrad and Schober (2000), p. 7.
15. Conrad and Schober (2000), p. 9.
16. Conrad and Schober (2000), p. 17.
17. Conrad and Schober (2000), p. 18.
18. Conrad and Schober (2000), p. 12.
19. Dillman (1978).
20. Dillman (1978), p. 12.
21. Dillman (2000); Dillman, Smyth, and Christian (2009).
22. Pew Research Center (2014).
23. Pew Research Center (2014).
24. Nilson and Cossman (2014), p. 1.
25. Pew Research Center (2014).
26. QSR International (2014).
27. See Internet World Stats (2014); Mohorko, de Leeuw, and Hox (2011); Mohorko, de Leeuw, and Hox (2013).
28. Couper (2008), p. 86; Couper et al. (2007), pp. 143–6.
29. Mohorko, de Leeuw, and Hox (2011), p. 10; Mohorko, de Leeuw, and Hox (2013), p. 618.
30. Alreck and Settle (2004), p. 185.
31. See Lind et al. (2013); Currivan et al. (2004).
32. Dillman, Smyth, and Christian (2009), pp. 300–2
33. See Dillman, Smyth, and Christian (2009), pp. 302–6 for a further discussion of these reasons.

34. Worthy and Mayclin (2013), p. 2.
35. Worthy and Mayclin (2013), p. 2.
36. Worthy and Mayclin (2013), p. 2.
37. Worthy and Mayclin (2013), p. 3.
38. Worthy and Mayclin (2013), pp. 3–5.
39. Worthy and Mayclin (2013), p. 4.
40. U.S. Census Bureau (2009).
41. Groves et al. (2004), pp. 241–53.
42. Schuman (2008), p. 63.
43. Schuman (2008), p. 65.
44. Schuman (2008), pp. 65–73.
45. Schuman (2008), pp. 81–88.
46. Bishop (1992).
47. Richardson, Dohrenwend, and Klein (1965).
48. Gorden (1987).
49. Gorden (1992).
50. Dillman (1978).
51. Dillman (2000).
52. Dillman, Smyth, and Christian (2009).
53. Gwartney (2007).
54. Couper (2008).
55. Groves et al. (2004).
56. Schuman (2008).

Chapter 6

1. Warwick and Lininger (1975).
2. Fisher (2009).
3. Fisher (2009), p. 143.
4. Warnecke et al. (1997).
5. de Bruin et al. (2012).
6. Snijkers et al. (2013).
7. Dillman (1978).
8. Dillman (2009), pp. 29–31.
9. Dillman (2009), pp. 3–5.
10. Dillman (2009), p. 4.
11. Dillman (2009), p. 32.
12. Maxfield and Babbie (1998), p. 109.
13. See, for example, de Leeuw and de Heer (2002).
14. Over the years Don A Dillman (2009) has been particularly attentive to these issues and has built on the notion of using incentives as a consistent component of survey design. Also see Church's (1993) meta-analysis.

15. Groves, Singer, and Corning (2000).
16. Dillman, Smyth, and Christian (2009).
17. Dillman (2009), p. 50.
18. Alreck and Settle (2004), pp. 89–92.
19. Dillman (2009), p. 33.
20. Bradburn (1983), pp. 304–6.
21. Alreck and Settle (2004), p. 94.
22. Zuell, Menold, and Körber (2014) found that particularly large answer boxes should be avoided, however, because they reduce respondents' willingness to respond.
23. A number of companies market qualitative analysis software, two of the best known are ATLAS (Scientific Software Development GmbH 2014) and NVivo International (QSR International 2014). Even companies traditionally known for their quantitative analysis software such as IBM SPSS provide qualitative software.
24. Krosnick and Fabrigar (1997).
25. Sudman and Bradburn (1973), p. 805.
26. Dillman (2009), pp. 40–47.
27. Likert (1932).
28. Dillman (2009).
29. Christian, Parsons, and Dillman (2009), p. 394.
30. Krosnick and Fabrigar (1997).
31. Brace (2013).
32. See, for example, Groothuis and Whitehead (2002).
33 Groves et al. (2009).
34. Dillman (2009).
35. Etchegaray and Fischer (2010).
36. Alwin (2007).
37. Brace (2013).

Chapter 7

1. See Alreck and Settle (2004); Fowler (2014), pp.100–1; Groves et al. (2004), pp. 243–5.
2. Bishop (1992).
3. See Fowler (2014), pp. 102–4; Groves et al. (2004), pp. 245–7.
4. See Converse and Presser (1986), pp. 54–65.
5. Schuman (2008), p. 81.
6. Bishop (1992).
7. Gwartney (2007), pp. 207–12.
8. Gwartney (2007), pp. 212–16 has a very useful discussion of probing when asking race and ethnicity questions.

9. Gwartney (2007), pp. 216–22.
10. Alreck and Settle (2004), p. 185.
11. Gwartney (2007), pp. 103–15.
12. AAPOR (2011) p. 3.
13. See Tourangeau and Plewes (2013), pp. 9–12 for various ways of calculating response rate suggested by the American Association for Public Opinion Research (AAPOR). APPOR defines response rate as "the number of complete interviews with reporting units divided by the number of eligible reporting units in the sample." American Association for Public Opinion Research (2011), p. 5.
14. Gorden (1992), pp. 82–93. Gorden includes other questions interviewers might ask themselves as respondents are answering questions.
15. Gorden (1992), pp. 92–96. We picked out particular keys to on which to focus.
16. Dillman (1978), pp. 260–1; Gwartney (2007), pp. 86–88.
17. Dillman (1978), p. 262; Gwartney (2007), pp. 163–5.
18. Tourangeau and Plewes (2013), p. 24.
19. Rosen et al. (2014).
20. Groves and McGonagle (2001).
21. Groves and McGonagle (2001), p. 253.
22. See Dillman (2000); Dillman, Smyth, and Christian (2009) for a systematic and thorough discussion of tailoring.
23. Fowler (2014).
24. Groves et al. (2004).
25. Babbie (1990).
26. Converse and Presser (1986).
27. Dillman (1978).
28. Dillman (2000).
29. Dillman, Smyth, and Christian (2009).
30. Gwartney (2007).
31. Couper (2008).
32. Gorden (1992).
33. Fowler (2014).
34. Gwartney (2007).
35. Groves and McGonagle (2001).
36. Weisberg (2005).
37. Groves et al. (2004).
38. Babbie (2013).
39. Miller (2005).
40. Frankfort-Nachmias and Leon-Guerrero (2011).
41. Chaffe-Stengel and Stengel (2011).

Chapter 8

1. Hague, Hague, and Morgan (2013), p. 197.
2. McAlevey and Sullivan (2010), p. 911.
3. McAlevey and Sullivan (2010), p. 912.
4. McAlevey and Sullivan (2010), p. 911.
5. Brunt (2001), p. 179.
6. Fink (2003).
7. Hague, Hague, and Morgan (2013), p. 197.
8. Beins (2012), p. 4.
9. For example, *The Field Poll,* an independent, media-sponsored, and non-partisan survey of California public opinion founded in 1947 by Mervin Field, has established a long-standing methodology which is detailed separately from the reports which it releases on various major social and political issues (The Field Poll 2014).
10. The Tuskegee syphilis experiment was an infamous clinical study conducted between 1932 and 1972 by the U.S. Public Health Service to study the natural progression of untreated syphilis in rural African American men who thought they were receiving free healthcare from the U.S. government (*Wikipedia* 2014).
11. U.S. Department of Health and Human Services (2014a).
12. AAPOR (2014a).
13. Hague, Hague, and Morgan (2013), p. 196.
14. Snijkers, Haraldsen, and Jones (2013), p. 536.
15. Snijkers, Haraldsen, and Jones (2013), p. 536.
16. Beins (2012).
17. Tufte (1983).
18. Tufte (1983).
19. Tufte (2006).
20. Champkin (2011), p. 41.
21. Fink (2003).
22. Lenburg (2010).
23. Forsyth (2013).
24. Hering and Hering (2010).
25. Miller (2004).
26. Alreck and Settle (2004), pp. 341–85.
27. Tufte (1983); Tufte (2006).

References

Alreck, P.L., and R.B. Settle. 2004. *The Survey Research Handbook*. 3rd ed. Boston, MA: McGraw-Hill Irwin.

Alwin, D.F. 2007. Vol. 547 of *Margins of Error: A Study of Reliability in Survey Measurement*. Hoboken, NJ: John Wiley & Sons.

AAPOR (American Association for Public Opinion Research). 2011. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 7th ed. AAPOR.

AAPOR. 2014a. "IRB Faqs for Survey Researchers." http://www.aapor.org/ irb_faqs_for_survey_researchers1.htm#question4 (accessed September 30, 2014).

AAPOR. 2014b. "Code of Ethics." http://www.aapor.org/AM/Template .cfm?Section=AAPOR_Code_of_Ethics&Template=/CM/ContentDisplay .cfm&ContentID=3142 (accessed September 4, 2014).

American Consumer Opinion. 2014. http://www.acop.com/ (accessed September 4, 2014).

Anderson, B.A., B.D. Silver, and P.R. Abramson. 1988. "The Effects of Race of the Interviewer on Measures of Electoral Participation by Blacks in SRC National Election Studies." *Public Opinion Quarterly* 52, no. 1, pp. 53–83.

Anderson, R., J. Kasper, and M.R. Frankel. 1979. *Total Survey Error*. San Francisco, CA: Jossey-Bass.

Babbie, E. 1990. *Survey Research Methods*. 2nd ed. Belmont, CA: Wadsworth.

Babbie, E. 2013. *The Practice of Social Research*. 13th ed. Wadsworth Cengage Learning.

Baker, R., J.M. Brick, N.A. Bates, M. Battaglia, M.P. Couper, J.A. Dever, K.J. Gile, and R. Tourangeau. 2013. "Summary Report of the AAPOR Task Force on Non-Probability Sampling." *Journal of Survey Statistics and Methodology* 1, no. 2, pp. 90–105.

Beatty, P. 1995. "Understanding the Standardized/Non-Standardized Interviewing Controversy." *Journal of Official Statistics* 11, no. 2, pp. 147–60.

Berman, J.J. 2013. *Principles of Big Data Preparing, Sharing, and Analyzing Complex Information*. Waltham, MA: Morgan Kaufmann.

Beins, B.C. 2012. "APA Style Simplified: Writing in Psychology, Education, Nursing, and Sociology." http://csus.eblib.com/patron/FullRecord.aspx?p= 822040

Bickart, B.A., J.M. Phillips, and J. Blair. 2006. "The Effects of Discussion and Question Wording on Self and Proxy Reports of Behavioral Frequencies." *Marketing Letters* 17, no. 3, pp. 167–80.

Biemer, P.P. 2010. "Total Survey Error: Design, Implementation, and Evaluation." *Public Opinion Quarterly* 74, no. 5, pp. 817–48.

Bishop, G.F. 1992. "Qualitative Analysis of Question-Order and Context Effects: The Use of Think-Aloud Responses." In *Context Effects in Social and Psychological Research*, eds. N. Schwarz and S. Sudman, 149–62. New York: Springer-Verlag.

Blumberg, S.J., J.V. Luke, N. Ganesh, M.E. Davern, M.H. Boudreaux, and K. Soderberg. April 20, 2011. "Wireless Substitution: State-Level Estimates from the National Health Survey, January 2007–June 2010." *National Health Statistics Reports* 39.

Box-Steffensmeier, J.M., G.C. Jacobson, and J.T. Grant. 2000. "Question Wording and the House Vote Choice: Some Experimental Evidence." *Public Opinion Quarterly* 64, no. 3, pp. 257–70.

Brace, I. 2013. *Questionnaire Design: How to Plan, Structure and Write Survey Material for Effective Market Research*. http://www.eblib.com

Brace, I., C. Nancarrow, and J. McCloskey. 1999. "MR Confidential: A Help or a Hindrance in the New Marketing Era?" *The Journal of Database Marketing* 7, no. 2, pp. 173–85.

Bradburn, N.M. 1983. "Response Effects." In *Handbook of survey research*, eds. P.H. Rossi, J.D. Wright, and A.B. Anderson, 289–328. Orlando, FL: Academic Press.

Bradburn, N.M., and S. Sudman. 1988. *Polls & Surveys*. San Francisco, CA: Jossey-Bass.

Brick, J.M. 2011. "The Future of Survey Sampling." *Public Opinion Quarterly* 75, no. 5, pp. 872–88.

Brunt, L.T. 2001. "The Advent of the Sample Survey in the Social Sciences." *Journal of the Royal Statistical Society* 50, pp. 179–89.

Callegaro, M., R.P. Baker, J. Bethlehem, A.S. Göritz, J.A. Krosnick, and P.J. Lavrakas. 2014. "Online Panel Research: A Data Quality Perspective." http://csus.eblib.com/patron/FullRecord.aspx?p=1662761

Cannell, C.F., P.V. Miller, and L. Oksenberg. 1981. "Research on Interviewing Techniques." In Vol. 12 of *Sociological Methodology*, ed. S. Leinhardt, 389–437. San Francisco, CA: Jossey-Bass.

Chaffe-Stengel, P., and D.N. Stengel. 2011. *Working with Sample Data-Exploration and Inference*. New York: Business Expert Press.

Champkin, J. 2011. "Making Information Beautiful—and Clear." *Significance* 8, no. 1, pp. 39–41.

Christian, L.M., N.L. Parsons, and D.A. Dillman. 2009. "Designing Scalar Questions for Web Surveys." *Sociological Methods & Research* 37, no. 3, pp. 393–425. doi: 10.1177/0049124108330004

Church, A.H. 1993. "Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta-Analysis." *Public Opinion Quarterly* 57, no. 1, pp. 62–79.

Cialdini, R.B. 2001. *Influence: Science and Practice*. 4th ed. Boston, MA: Allyn and Bacon.

Conrad, F.G., and M.F. Schober. 2000. "Clarifying Question Meaning in a Household Telephone Survey." *Public Opinion Quarterly* 64, no. 1, pp. 1–28.

Converse, J.M., and S. Presser. 1986. Vol. 63 of *Survey Questions: Handcrafting the Standardized Questionnaire*. Beverly Hills, CA: Sage.

CASRO (Council of American Survey Research Organizations). 2014. "Code of Standards and Ethics for Survey Research." http://smari.com/images/uploads/CASRO-Code-of-Ethics.pdf (accessed September 3, 2014).

Couper, M.P. 2008. *Designing Effective Web Surveys*. Cambridge, MA: Cambridge University Press.

Couper, M.P., A. Kapteyn, M. Schonlau, and J. Winter. 2007. "Noncoverage and Nonresponse in an Internet Survey." *Social Science Research* 36, no. 1, pp. 131–48.

Currivan, D.B., A.L. Nyman, C.F. Turner, and L. Biener. 2004. "Does Telephone Audio Computer-Assisted Self-Interviewing Improve the Accuracy of Prevalence Estimates of Youth Smoking? Evidence from the UMASS Tobacco Study." *Public Opinion Quarterly* 68, no. 4, pp. 542–64.

Daniel, J.N. 2012. *Sampling Essentials: Practical Guidelines for Making Sampling Choices*. Thousand Oaks, CA: Sage Publications, Inc.

Davis, D.W., and B.D. Silver. 2003. "Stereotype Threat and Race of Interviewer Effects in a Survey on Political Knowledge." *American Journal of Political Science* 47, no. 1, pp. 33–45.

de Bruin, Bruine, W., W. van der Klaauw, G. Topa, J.S. Downs, B. Fischhoff, and O. Armantier. 2012. "The Effect of Question Wording on Consumers' Reported Inflation Expectations." *Journal of Economic Psychology* 33, no. 4, pp. 749–57. doi: http://dx.doi.org/10.1016/j.joep.2012.02.001

de Leeuw, E.D. 2005. "To Mix or Not to Mix Data Collection Modes in Surveys." *Journal of Official Statistics* 21, no. 2, pp. 233–55.

de Leeuw, E.D. 2008. "Choosing the Method of Data Collection." In *International Handbook of Survey Methodology*, eds. E.D. de Leeuw, J.J. Hox, and D.A. Dillman, 113–35. New York: Psychology Press.

de Leeuw, E.D., and W. de Heer. 2002. "Trends in Household Survey Nonresponse." In *Survey Nonresponse*, eds. R.M. Groves, D.A. Dillman, J.L. Eltinge, and R.J.A. Little, 41–54. New York: John Wiley and Sons.

Deutscher, I. 1973. *What We Say/What We Do*. Glenview, IL: Scott, Foresman and Company.

Dillman, D.A. 1978. *Mail and Telephone Surveys: The Total Design Method*. New York: John Wiley & Sons.

Dillman, D.A. 2000. *Mail and Internet Surveys: The Tailored Design Method*. 2nd ed. New York: John Wiley & Sons.

Dillman, D.A. (2009). *Internet, Mail and mixed-Mode Surveys: The Tailored Design Method*. New York: John Wiley & Sons.

Dillman, D.A., J.D. Smyth, and L.M. Christian. 2009. *Internet, Mail and Mixed-Mode Surveys: The Tailored Design Method*. 3rd ed. New York: John Wiley & Sons.

Duff, B., M.J. Hanmer, W.-H. Park, and I.K. White. 2007. "Good Excuses: Understanding Who Votes with an Improved Turnout Question." *Public Opinion Quarterly* 71, no. 1, pp. 67–90.

Engel, U., B. Jann, P. Lynn, A. Scherpenzeel, and P. Sturgis. 2014. "Improving Survey Methods: Lessons from Recent Research." http://csus.eblib.com/patron/FullRecord.aspx?p=1775347

ESOMAR Research Organization. 2006. "26 Questions to Help Research Buyers of Online Samples." http://www.esomar.org/uploads/public/knowledge-and-standards/codes-and-guidelines/ESOMAR_26-Questions-To-Help-Research-Buyers-Of-Online-Samples.pdf (accessed September 3, 2014).

Etchegaray, J.M., and W.G. Fischer. 2010. "Understanding Evidence-Based Research Methods: Reliability and Validity Considerations in Survey Research." *Health Environments Research & Design Journal* 4, no. 1, pp. 131–35.

The Field Poll. 2014. "The Survey Methodology of the Field Poll." http://www.field.com/fieldpoll/methods.html (accessed September 30, 2014).

Fink, A. 1995. Vol. 6 of *How to Sample in Surveys*. Thousand Oaks, CA: Sage Publications.

Fink, A. 2003. Vol. 10 of *How to Report on Surveys*. 2nd ed. Thousand Oaks, CA: Sage Publications.

Finkel, S.E., T.M. Guterbock, and M.J. Borg. 1991. "Race-of-Interviewer Effects in a Preelection Poll." *Public Opinion Quarterly* 55, no. 3, pp. 313–30.

Fisher, B.S. 2009. "The Effects of Survey Question Wording on Rape Estimates: Evidence from a Quasi-Experimental Design." *Violence against Women* 15, no. 2, pp. 133–47.

Forsyth, P. 2013. "How to Write Reports and Proposals." http://csus.eblib.com/patron/FullRecord.aspx?p=1131620 doi: 10.1177/1077801208329391

Fowler, F.J., Jr. 1993. Vol. 1 of *Survey Research Methods*. 2nd ed. Newbury Park, CA: Sage Publications.

Fowler, F.J., Jr. 2014. *Survey Research Methods*. 5th ed. Los Angeles, CA: Sage Publications.

Frankel, M.R. 1983. "Sampling Theory." In *Handbook of Survey Research*, eds. P.H. Rossi, J.D. Wright, and A.B. Anderson, 21–66. Orlando, FL: Academic Press, Inc.

Frankel, M., and F. Lester. 1987. "Fifty Years of Surveying in the United States." *Public Opinion Quarterly* 51, no. 2, pp. 127–38.

Frankfort-Nachmias, C., and A. Leon-Guerrero. 2011. *Social Statistics for a Diverse Society*. Los Angeles, CA: Pine Forge Press.

Galesic, M., R. Tourangeau, M.P. Couper, and F.G. Conrad. 2008. "Eye-Tracking Data: New Insights on Response Order Effects and Other Cognitive Shortcuts in Survey Responding." *Public Opinion Quarterly* 72, no. 5, pp. 892–913.

Geppert, C., M.P. Bogenschutz, and W.R. Miller. 2007. "Development of a Bibliography on Religion, Spirituality and Addictions." *Drug and Alcohol Review* 26, no. 4, pp. 389–95.

Gorden, R.L. 1987. *Interviewing: Strategy, Techniques, and Tactics*. 4th ed. Chicago, IL: Dorsey Press.

Gorden, R.L. 1992. *Basic Interviewing Skills*. Itasca, IL: F.E. Peacock Publishers.

Groothuis, P.A., and J.C. Whitehead. 2002. "Does Don't Know Mean No? Analysis of 'Don't Know' Responses in Dichotomous Choice Contingent Valuation Questions." *Applied Economics* 34, no. 15, pp. 1935–40. doi: 10.1080/00036840210128717

Groves, R.M., R.B. Cialdini, and M.P. Couper. 1992. "Understanding the Decision to Participate in a Survey." *Public Opinion Quarterly* 56, no. 4, pp. 475–95.

Groves, R.M., F.J. Fowler, M.P. Couper, J.M. Lepkowski, E. Singer, and R. Tourangeau. 2004. *Survey Methodology*. Hoboken, NJ: Wiley-Interscience.

Groves, R.M., F.J. Fowler, M.P. Couper, J.M. Lepkowski, E. Singer, and R. Tourangeau. 2009. *Survey Methodology*. 2nd ed. Hoboken, NJ: John Wiley.

Groves, R.M., and K.A. McGonagle. 2001. "A Theory-Guided Interviewer Training Protocol Regarding Survey Participation." *Journal of Official Statistics* 17, no. 2, pp. 249–65.

Groves, R.M., E. Singer, and A. Corning. 2000. "Leverage-Saliency Theory of Survey Participation: Description and Illustration." *Public Opinion Quarterly* 64, no. 3, pp. 299–308.

Gwartney, P.A. 2007. *The Telephone Interviewer's Handbook: How to Conduct Standardized Conversations*. San Francisco, CA: Josey-Bass.

Hague, P.N., N. Hague, and C.-A. Morgan. 2013. "Market Research in Practice: How to Get Greater Insight from Your Market." http://csus.eblib.com/patron/FullRecord.aspx?p=1412261

Hartford, T. 2014. "Big Data: Are We Making a Big Mistake?" *FT Magazine.* 11/02/2014http://www.ft.com/intl/cms/s/2/21a6e7d8-b479-11e3-a09a-00144feabdc0.html#axzz3HxNUwGpa (accessed November 2, 2014).

Hatchett, S., and H. Schuman. 1975. "White Respondents and Race-of-Interviewer Effects." *Public Opinion Quarterly* 39, no. 4, pp. 523–28.

Heerwegh, D., and G. Loosveldt. 2008. "Face-to-Face Versus Web Surveying in a High-Internet-Coverage Population: Differences in Response Quality." *Public Opinion Quarterly* 72, no. 5, pp. 836–46.

Hering, L., and H. Hering. 2010. *How to Write Technical Reports*. http://csus.eblib.com/patron/FullRecord.aspx?p=645771

Holbrook, A.L., M.C. Green, and J.A. Krosnick. 2003. "Telephone Versus Face-to-Face Interviewing of National Probability Samples with Long Questionnaires: Comparisons of Respondent Satisficing and Social Desirability Response Bias." *Public Opinion Quarterly* 67, no. 1, pp. 79–125.

Holmes, T.J., and J.A. Schmitz Jr. 1996. "Nonresponse Bias and Business Turnover Rates: The Case of the Characteristics of Business Owners Survey." *Journal of Business & Economic Statistics* 14, no. 2, pp. 231–41.

Huber, G.A., and C. Paris. 2013. "Assessing the Programmatic Equivalence Assumption in Question Wording Experiments: Understanding Why Americans Like Assistance to the Poor More than Welfare." *Public Opinion Quarterly* 77, no. 1, pp. 385–97.

Hundepool, A., J. Domingo-Ferrer, L. Franconi, S. Giessing, E. Schulte Nordholt, K. Spicer, and P.-P. de Wolf. 2012. "Statistical Disclosure Control. http://csus.eblib.com/patron/FullRecord.aspx?p=952417

IBM SPSS. 2014. "SPSS Homepage." http://www-01.ibm.com/software/analytics/spss/ (accessed September 20, 2014).

Internet World Stats. 2014. "Internet World Stats: Usage and Population Statistics." http://www.internetworldstats.com/ (accessed May 19, 2014).

Johnson, T.P., and J.A. Parsons. 1994. "Interviewer Effects on Self-Reported Substance Use among Homeless Persons." *Addictive Behaviors* 19, no. 1, pp. 83–93.

Kane, E.W., and L.J. Macaulay. 1993. "Interviewer Gender and Gender Attitudes." *Public Opinion Quarterly* 57, no. 1, pp. 1–28.

Kennedy, J.M., J. Tarnai, and J.G. Wolf. 2010. "Managing Survey Research Projects." In *Handbook of Survey Research*, eds. P.V. Marsden and J.D. Wright, 575–90. 2nd ed. Bingley, UK: Emerald Group.

Kish, L. 1967. *Survey Sampling*. New York: John Wiley and Sons.

Kreuter, F., S. Presser, and R. Tourangeau. 2008. "Social Desirability Bias in CATI, IVR, and Web Surveys: The Effects of Mode and Question Sensitivity." *Public Opinion Quarterly* 72, no. 5, pp. 847–65.

Krosnick, J.A. 1991. "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys." *Applied Cognitive Psychology* 5, no. 3, pp. 213–36.

Krosnick J.A., and L.R. Fabrigar. 1997. "Designing Rating Scales for Effective Measurement in Surveys." In *Survey Measurement and Process Quality*, eds. P.B.L. Lyberg, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, and D. Trewin, 141–64. New York: John Wiley.

Krosnick, J.A., A.L. Holbrook, M.K. Berent, R.T. Carson, W.M. Hanemann, R.J. Kopp, R.C. Mitchell, S. Presser, P.A. Ruud, V.K. Smith, W.R. Moody, M.C. Green, and M. Conaway. 2002. "The Impact of 'No Opinion' Response Options on Data Quality: Non-Attitude Reduction or an Invitation to Satisfice?" *Public Opinion Quarterly* 66, no. 3, pp. 371–403.

LaPiere, R.T. 1934. "Attitudes vs. Actions." *Social Forces* 13, no. 2, pp. 230–37.

LaPiere, R.T. 2010. "Attitudes vs. Actions." *International Journal of Epidemiology* 39, no. 1, pp. 7–11.

Lazarsfeld, P.F. 1944. "The Controversy over Detailed Interviews-an Offer for Negotiation." *Public Opinion Quarterly* 8, no. 1, pp. 38–60.

Lenburg, J. 2010. "The Facts on File Guide to Research." http://csus.eblib.com/patron/FullRecord.aspx?p=592610

Liebow, E. 1967. *Tally's Corner: A Study of Negro Streetcorner Men.* Boston, MA: Little, Brown and Company.

Likert, R. 1932. "A Technique for the Measurement of Attitudes." *Archives of Psychology* 140, pp. 5–55.

Lind, L.H., M.F. Schober, F.G. Conrad, and H. Reichert. 2013. "Why Do Survey Respondents Disclose More When Computers Ask the Questions?" *Public Opinion Quarterly* 77, no. 4, pp. 888–935.

Marsden, P.V., ed. 2012. *Social Trends in American Life—Findings from the General Social Survey Since 1972.* Princeton, NJ: Princeton University Press.

Maxfield, M.G., and E. Babbie. 1998. *Research Methods for Criminal Justice and Criminology.* 2nd ed. Belmont, CA: Wadsworth Publishing.

McAlevey, L., and C. Sullivan. 2010. "Statistical Literacy and Sample Survey Results." *International Journal of Mathematical Education in Science and Technology* 41, no. 7, pp. 911–20.

McDonald, M.P., and M.P. Thornburg. 2012. "Interview Mode Effects: The Case of Exit Polls and Early Voting." *Public Opinion Quarterly* 76, no. 2, pp. 326–49.

McNabb, D.E. 2014. *Nonsampling Error in Social Surveys.* Los Angeles, CA: Sage.

Miller, D.C., and N.J. Salkind. 2002. *Handbook of Research Design and Social Measurement.* 6th ed. Los Angeles, CA: Sage.

Miller, J.E. 2004. *The Chicago Guide to Writing About Numbers.* Chicago, IL: University of Chicago Press.

Miller, J.E. 2005. *The Chicago Guide to Writing About Multivariate Analysis.* Chicago, IL: University of Chicago Press.

Mohorko, A., de Leeuw, E., and Hox, J. 2011. "Internet Coverage and Coverage Bias in Countries Across Europe and over Time: Background, Methods, Question Wording and Bias Tables." http://www.joophox.net (accessed March 25, 2014)

Mohorko, A., de Leeuw, E., and Hox, J. 2013. "Internet Coverage and Coverage Bias in Europe: Developments Across Countries and over Time." *Journal of Official Statistics* 29, no. 4, pp. 609–22.

Monette, D.R., T.J. Sullivan, and C.R. DeJong. 1998. *Applied Social Research: Tool for the Human Services.* 4th ed. Fort Worth, TX: Harcourt Brace College.

Moore, D.W. 2002. "Measuring New Types of Question-Order Effects: Additive and Subtractive." *Public Opinion Quarterly* 66, no. 1, pp. 80–91.

Neyman, J. 1934. "On Two Different Aspects of the Representative Model: The Method of Stratified Sampling and the Method of Purposive Deletion." *Journal of the Royal Statistical Society* 97, pp. 558–625.

Nilson, A., and R.E. Cossman. 2014. "It's Worse Than We Thought—Landline RDD Unreachable Young Adults in a College Town." *Survey Practice* 7, no. 1, pp. 1–6. http://www.surveypractice.org/index.php/SurveyPractice/article/view/269/html

Oldendick, R.W. 2012. "Survey Reseach Ethics." In *Handbook of Survey Methodology for the Social Sciences*, ed. L. Gideon, 23–35. Heidelberg: Springer. http://www.eblib.com

Petrolia, D.R., and S. Bhattacharjee. 2009. "Revisiting Incentive Effects: Evidence from a Random-Sample Mail Survey on Consumer Preferences for Fuel Ethanol." *Public Opinion Quarterly* 73, no. 3, pp. 537–50.

Pew Research Center. 2014. "Pew Research Increases Share of Interviews Conducted by Cellphone." http://www.pewresearch.org/fact-tank/2014/01/15/pew-research-increases-share-of-interviews-conducted-by-cellphone/ (accessed March 16, 2014).

Peytchev, A., E. Peytcheva, and R.M. Groves. 2010. "Measurement Error, Unit Nonresponse, and Self-reports of Abortion Experiences." *Public Opinion Quarterly* 74, no. 2, pp. 319–27.

Pollner, M. 1998. "The Effects of Interviewer Gender in Mental Health Interviews." *Journal of Nervous and Mental Health Disease* 186, no. 6, pp. 369–73.

Porter, S.R. 2004." Raising Response Rates: What Works?" In *Overcoming Survey Research Problems,* ed. S.R. Porter, pp. 5–21. San Francisco, CA: Jossey-Bass.

Preisendorfer, P., and F. Wolter. 2014. "Who is Telling the Truth? A Validation Study on Determinants of Response Behavior in Surveys." *Public Opinion Quarterly* 78, no. 1, pp. 126–46.

QuestionPro. 2014. "QuestionPro Survey Panels." http://www.questionpro.com/surveypanels/survey-panel.html (accessed September 4, 2014).

QSR International. 2014. "NVivo 10 for Windows." http://www.qsrinternational.com/products_nvivo.aspx (accessed March 16, 2014).

Richardson, S.A., B.S. Dohrenwend, and D. Klein. 1965. *Interviewing: Its Forms and Functions*. New York: Basic Books.

Riley, M.W. 1963. *Sociological Research I: A Case Approach*. New York: Harcourt, Brace & World.

Rosen, J.A., J. Murphy, A. Peytchev, T. Holder, J.A. Dever, D.R. Herget, and D.J. Pratt. 2014. "Prioritizing Low-propensity Sample Members in a Survey: Implications for Nonresponse Bias." *Survey Practice* 7, no. 1, pp. 1–8. http://www.surveypractice.org/index.php/SurveyPractice/article/view/258/html

Roth, S.B, D. Han, and J.M. Montaquila. 2013. "The ABS Frame Quality and Considerations." *Survey Practice* 6, no. 4, pp. 1–6. http://surveypractice.org/index.php/SurveyPractice/article/view/73

Salant, P., and D.A. Dillman. 1994. *How to Conduct Your Own Survey*. New York: John Wiley & Sons.

Schuldt, J.P., S.H. Konrath, and N. Schwarz. 2011. "'Global Warming' or 'Climate Change'? Whether the Planet is Warming Depends on Question Wording." *Public Opinion Quarterly* 75, no. 1, pp. 115–24.

Schuman, H. 2002. "Sense and Nonsense About Surveys." *Contexts* 1, no. 2, pp. 40–47.

Schuman, H. 2008. *Method and Meaning in Polls and Surveys*. Cambridge, MA: Harvard University Press.

Schuman, H., and J.M. Converse. 1971. "The Effects of Black and White Interviewers on Black Responses in 1968." *Public Opinion Quarterly* 35, no. 1, pp. 44–68.

Schuman, H., and S. Presser. 1981. *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. Orlando, FL: Academic Press, Inc.

Scientific Software Development GmbH. 2014. "Atlas TI Homepage." http://www.atlasti.com/index.html (accessed September 20, 2014).

Sincero, S.M. 2012. "Survey Response Scales." https://explorable.com/survey-response-scales (accessed September 22, 2014).

Singer, E., J. Van Hoewyk, and M.P. Maher. 2000. "Experiments with Incentives in Telephone Surveys." *Public Opinion Quarterly* 64, no. 2, pp. 171–88.

Singer, E., J. Van Hoewyk, and R.J. Neugebauer. 2003. "Attitudes and Behavior: The Impact of Privacy and Confidentiality Concerns on Participation in the 2000 Census." *Public Opinion Quarterly* 67, no. 3, pp. 368–84.

Smith, T.W. 1987. "That Which We Call Welfare by Any Other Name Would Smell Sweeter: An Analysis of the Impact of Question Wording on Response Patterns." *Public Opinion Quarterly* 51, no. 1, pp. 75–83.

Snijkers, G., G. Haraldsen, and J. Jones. 2013. *Designing and Conducting Business Surveys*. Hoboken, New Jersey: John Wiley & Sons. http://www.eblib.com

Stat Trek. 2014. "Introduction to Survey Sampling." http://stattrek.com/survey-research/survey-sampling.aspx (accessed February 10, 2014).

Steeh, C. 2008. "Telephone Surveys." In *International Handbook of Survey Methodology*, eds. E.D. de Leeuw, J.J. Hox and D.A. Dillman, pp. 221–38. New York: Psychology Press.

Streb, M.J., B. Burrell, B. Frederick, and M.A. Genovese. 2008. "Social Desirability Effects and Support for a Female American President." *Public Opinion Quarterly* 72, no. 1, pp. 76–89.

Sudman, S. 1983. "Applied Sampling." In *Handbook of Survey Research*, eds. P.H. Rossi, J.D. Wright, and A. Anderson, 145–94. Orlando, FL: Harcourt Brace Jovanovich, Publishers.

Sudman, S, and N. Bradburn N. 1973. "Effects of Time and Memory Factors on Response in Surveys." *Journal of the American Statistical Association* 68, pp. 805–815.

Survey Police. 2014. "Survey Panel Rankings." https://www.surveypolice.com/rankings (accessed September 3, 2014)

Tippett, L.H.C. 1927. *1952 Random Sampling Numbers*. Cambridge: Cambridge University Press.

Tourangeau, R., R.M. Groves, and C.D. Redline. 2010. "Sensitive Topics and Reluctant Respondents: Demonstrating a Link Between Nonresponse Bias and Measurement Error." *Public Opinion Quarterly* 74, no. 3, pp. 413–32.

Tourangeau, R., and T.J. Plewes, eds. 2013. *Nonresponse in Social Science Surveys: A Research Agenda*. Washington, D.C.: National Academies Press.

Tufte, E.R. 1983. *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press.

Tufte, E. R. 2006. *Beautiful Evidence*. Cheshire, CT: Graphics Press.

U.S. Census Bureau. 2009. "Design and Methodology: American Community Survey." https://www.google.com/search?newwindow=1&q=ACS+design+and+methodology&oq=ACS+design+and+methodlogy&gs_l=serp.12..33i21.25925.36033.0.38379.26.24.0.2.2.0.134.1976.19j5.24.0....0...1c.1.37.serp..10.16.1288.-aCo8W_O0x8

U.S. Census Bureau. 2014. "Overview of U.S. Census Surveys." https://www.census.gov/programs-surveys/are-you-in-a-survey/survey-list/household-survey-list.html# (accessed August 30, 2014).

U.S. Department of Health and Human Services. 2014a. *Health Information Privacy*. http://www.hhs.gov/ocr/privacy/ (accessed August 7, 2014).

U.S. Department of Health and Human Services. 2014b. "Code of Federal Regulations," 45 CFR 46. http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.html (accessed September 9, 2014).

Valliant, R., A. Mercer, and W. Hicks. 2014. "Efficient Use of Commercial Lists in U.S. Household Sampling." *Journal of Survey Statistics and Methodology* 2, no. 2, pp. 182–209.

Warnecke, R.B., T.P. Johnson, N. Chávez, S. Sudman, D.P. O'Rourke, L. Lacey, and J. Horm. 1997. "Improving Question Wording in Surveys of Culturally Diverse Populations." *Annals of Epidemiology* 7, no. 5, pp. 334–42.

Warwick, D.P., and C.A. Lininger. 1975. *The Sample Survey: Theory and Practice*. New York: McGraw-Hill.

Webb, E.J., D.T. Campbell, R.D. Schwartz, L. Sechrest, and J.B. Grove. 1981. *Nonreactive Measures in the Social Sciences*. 2 ed. Boston, MA: Houghton Miflin.

Weisberg, H.F. 2005. *The Total Survey Error Approach: A Guide to the New Science of Survey Research*. Chicago, IL: University of Chicago Press.

Whitmarsh, L. 2009. "What's in a Name? Commonalities and Differences in Public Understanding of 'Climate Change' and 'Global Warming.'" *Public Understanding of Science* 18, no. 4, pp. 401–20.

Whyte, W.F. 1943. *Street Corner Society: The Social Structure of an Italian Slum*. Chicago, IL: University of Chicago Press.

Whyte, W.F. 1975. "Conflict and Cooperation in Andean Communities." *American Ethnologist* 2, no. 2, pp. 373–92.

Wikipedia. 2014. Tuskegee Syphilis Experiment. http://www.ask.com/wiki/Tuskegee_syphilis_experiment?o=2801&qsrc=999&ad=doubleDown&an=apn&ap=ask.com#cite_note-timeline-1 (accessed September 30, 2014).

Worthy, M., and D. Mayclin. 2013. "Implementing a Mixed-mode Design for Collecting Administrative Records: Striking a Balance Between Quality and Burden." *Survey Practice* 6, no. 2, pp. 1–6. http://www.surveypractice.org/index.php/SurveyPractice/article/view/26

Ye, C., J. Fulton, and R. Tourangeau. 2011. "More Positive or More Extreme? A Meta-analysis of Mode Differences in Response Choice." *Public Opinion Quarterly* 75, no. 2, pp. 349–65.

Zuell, C., N. Menold, and S. Körber. 2014. "The Influence of the Answer Box Size on Item Nonresponse to Open-ended Questions in a Web Survey." *Social Science Computer Review*. doi: 10.1177/0894439314528091

Index

An Introduction to Survey Research

Ernest L. Cowles • Edward Nelson

An Introduction to Survey Research is for those who want an overview of the survey process. It is intended to describe fundamental survey components to help both students and managers understand and use surveys effectively and avoid the pitfalls stemming from bad survey construction and inappropriate methods. The authors discuss how best to identify the information needed and the best approach to get that information. They also highlight the processes commonly involved in conducting a survey including the value of both obtaining a representative sample and dealing with the types of errors that can distort results. Each chapter focuses on one of the key components of constructing and carrying out a survey, including the elements to consider when developing a survey, the modes of survey delivery, writing good questions, conducting the survey, and presenting the results.

Ernest L. Cowles is professor emeritus of sociology and past director of the Institute for Social Research at California State University, Sacramento (CSUS). He received his PhD in criminology from Florida State University. Prior to his appointment at CSUS, he served as the director of the Institute for Legal and Policy Studies at the University of Illinois for 10 years and interim director of the Institute for Public Affairs, Abraham Lincoln Center for Policy Studies for two years. He has received more than $25 million in research funding and has authored numerous publications and research reports.

Edward Nelson is professor emeritus of sociology at California State University, Fresno. He received his PhD in sociology from the University of California, Los Angeles, specializing in research methods. He was the director of the Social Research Laboratory at CSU, Fresno from 1980 to 2013 and directed more than 150 surveys. He taught research methods, quantitative methods, critical thinking, and computer applications. He has published books on observation in sociological research and using SPSS, a statistical computing package widely used in the social sciences.

Quantitative Approaches to Decision Making Collection

Donald N. Stengel, *Editor*

ISBN: 978-1-60649-818-7

90000

9 781606 498187

BUSINESS EXPERT PRESS