has shown that visual layout affects respondents in many different ways. These effects were the main focus of the 2007 update (Dillman, 2007) and are developed further throughout this third edition of the book.

In addition, we also devote a chapter to longitudinal surveys and a new kind of tailored design, the Internet panel survey. In that chapter we discuss the unique challenges faced by each of these types of surveys as well as some challenges that are common to both. Another new chapter discusses how sponsorship affects survey design, with an emphasis on the regulatory impacts of institutional review boards and the Office of Management and Budget.

CONCLUSION

When interviewers walked city streets or traversed the countryside by car in the 1940s to sample households and interview the people who lived in them, they were pursuing the same objective that would later be pursued by surveyors in the 1970s who sent written questionnaires to postal addresses or who called telephone numbers randomly to contact people, and by surveyors in the 2000s who sent e-mail requests asking people to complete a survey on the Internet. The scientific tenets of surveying—that is, getting a sample of a few hundred or thousand respondents to allow estimates with known precision to be made for the population from which they are selected—have remained the same. However, the methods used for asking questions and obtaining answers and knowledge of the factors that influence the accuracy of those answers and the proportion of people who will provide them have changed dramatically.

This book is about obtaining high-quality responses in the early twenty-first century using the technology available and results of decades of research on factors that influence people to answer survey questions and to respond accurately. Each of the remaining 12 chapters covers some aspect of data collection and how to tailor to different survey populations and situations. We begin in Chapter 2 with the tenets of tailored design, focusing on scientific and theoretical knowledge about what makes people likely to respond to surveys and how to create effective interaction with respondents to encourage cooperation and valid answers to survey questions.

CHAPTER 2

The Tailored Design Method

CONSIDER FOR a moment the following examples of the enormous variety of circumstances surveyors now face, each of which captures a situation we have encountered recently:

- Each month a federal agency conducts a survey of thousands of businesses about their number of employees. The survey is very short, only six questions, but the agency has to collect the data and report it in less than 2 weeks.
- A television rating organization collects people's viewing habits from each of the nation's several hundred television markets to describe those habits accurately for each market for a specific week of the year.
- The elected leaders of a U.S.-based professional organization changed the program for its annual meeting and now wants to know whether attendees liked or did not like the change. One challenge was that the number of attendees from other countries had increased significantly, and the leaders wanted to survey enough of them to know whether their satisfaction level was different from that of U.S. participants.
- A federal agency that is surveying universities faces the challenge of collecting information that is known in individual academic departments but is not generally available from a centralized office.

It should be apparent, even from this small list, that the same procedures will not work for all surveys. But how does one go about deciding which procedures to use and not use, and by what criterion does one choose certain methods for collecting data over others? Also, under what conditions should one choose a single survey mode, and under what conditions is it better to

At first glance, this situation may hardly seem different from that faced for decades by survey researchers. However, the dizzying array of mode possibilities now available, individually and in combination with one another and each with quite different cost implications, adds to the complexity of the situation. In addition, the dramatic changes described in Chapter 1 with regard to the presence or absence of human interaction, trust in the legitimacy of the survey, and changes in respondent control over whether and how they can be contacted, make what once may have been a simple survey design situation much more difficult. In this chapter we introduce tailored design as a means of helping sort out survey procedures that are effective from those that are ineffective given specific survey contexts.

Tailored design involves using multiple motivational features in compatible and mutually supportive ways to encourage high quantity and quality of response to the surveyor's request. It is developed from a social exchange perspective on human behavior, which suggests that respondent behavior is motivated by the return that behavior is expected to bring, and in fact, usually does bring, from others. It assumes that the likelihood of responding to a self-administered questionnaire, and doing so accurately, is greater when the respondent trusts that the expected rewards will outweigh the anticipated costs of responding.

Underlying this general approach are three fundamental considerations. First, tailored design is a scientific approach to conducting sample surveys with a focus on reducing the four sources of survey error—coverage, sampling, nonresponse, and measurement (Groves, 1989)—that may undermine the quality of the information collected. Second, the tailored design method involves developing a set of survey procedures (including the contact letters or e-mails to respondents and the questionnaire) that interact and work together to encourage all people in the sample to respond to the survey. Thus, it entails giving attention to all aspects of contacting and communicating with respondents—few, if any, aspects of this process can be ignored when using a tailored design strategy. Finally, tailoring is about developing survey procedures that build positive social exchange and encourage response by taking into consideration elements such as survey sponsorship, the nature of the survey population and variations within it, and the content of the survey questions, among other things.

REDUCING TOTAL SURVEY ERROR

Conducting surveys that produce accurate information that reflects the views and experiences of a given population requires developing procedures that minimize all four types of survey error—coverage, sampling, nonresponse, and measurement (Groves, 1989). Reducing survey error means selecting the

the entire population (low coverage error) and from which a larg
random sample of the desired population can be drawn (minimizes
error), designing an implementation system that encourages most
the sample to respond (reduces nonresponse error), and approachin
dents in the contacts and the questionnaire itself in a way that en
and enables them to provide thoughtful and honest answers (decrea
surement error).

Coverage error occurs when not all members of the populatio
known, nonzero chance of being included in the sample for the su
when those who are excluded are different from those who are inc
measures of interest. Coverage error can occur because the choic
vey mode may not provide adequate coverage of the population,
case with Internet surveys where a significant number of people
populations do not have access to the Internet. It can also occur whe
from which the sample is drawn does not include everyone in the po
such as when a sample frame based only on households with listed t
numbers is used or when a list is not current and excludes people wh
or joined in the last year (for a more in-depth discussion of covera
see Chapter 3).

The extent to which the precision of the survey estimates is limited
not every person in the population is sampled is described as the
error. The power of random sampling, which is discussed in detail
ter 3, is that estimates with acceptable levels of precision can usually
for the population by randomly surveying only a small portion o
in the population. For example, a researcher can survey about 100 r
of the U.S. general public and achieve estimates with a margin of
±10%, with a 95% confidence level (i.e., 95 out of 100 times the
will be within ±10%). Surveying 2,000 people reduces the margin
to about ±2%. Surveying 100, or even 2,000, people rather than the
mately 303 million who live in the United States represents an enorm
desirable cost savings, but doing so means that one has to be willin
with some error in the estimates. In other words, sampling error res
surveying only some rather than all members of the population and
a part of all sample surveys.

Another source of error, nonresponse error, stems from not ge
eryone who was sampled to respond to the survey request. *Non
error* occurs when the people selected for the survey who do not
are different from those who do respond in a way that is importa
study. For example, a survey of voting intentions for a presiden
tion would be rife with nonresponse error if Democrats were sign
less likely to respond than Republicans. Minimizing nonresponse
topic covered in considerably more depth in Chapter 7) involves

questionnaires from people of different sociodemographic groups or with other characteristics that may be important to the study (i.e., from different types of people).

Lastly, *measurement error* occurs when a respondent's answer is inaccurate or imprecise. Measurement error is often the result of poor question wording or design and other aspects of questionnaire construction (see Chapters 4–6). Because interviewers are not present to help respondents navigate the survey and understand the meaning of the questions in self-administered surveys, questionnaire layout and the design and wording of individual questions are extremely important in ensuring that respondents give accurate and precise answers.

Figure 2.1 presents four surveys, each of which failed to achieve its objectives as originally designed but for very different reasons. These surveys illustrate how each of the four sources of error can undermine survey projects and how designing effective surveys involves paying attention to reducing total survey error.

PERSPECTIVES ON WHY PEOPLE RESPOND TO SURVEYS

THE ECONOMIC EXCHANGE VIEW OF SURVEY RESPONSE

Several years ago an investigator for a very large and well-financed national mail survey about the retirement resources and plans of the survey population called to ask a single question: "I am doing a mail survey that will ask about people's retirement resources and will be about 16 pages long. How much do I have to pay them in order to get them to respond?" His question was framed from the perspective of economic exchange, much like one might ask what price needs to be set in order to get everyone to buy a book, a piece of software, or even a meal in a particular restaurant.

There are several problems with attempting to set a price for what responding to a survey is worth to a respondent and with using money as the primary motivation for seeking a response. People's price points, if they have one, are likely to vary widely, so that the responses of those who judge a given payment as adequate may differ significantly from the responses of those who do not. In addition, very few surveyors have the resources to pay an amount that would guarantee a response from most or all respondents, so in general economic exchange is not a usable model for surveyors. Furthermore, considerable research (discussed in detail in Chapter 7) suggests fairly substantial payments to respondents are not as effective in raising response rates as small token cash incentives given in advance. Also, many survey sponsors in widely different settings—government, universities, and the pri-

Figure 2.1 Why surveys fail.

A designer of health rehabilitation devices wanted to understand potential demand for its products among the U.S. general public and how having health insurance might affect the ability to purchase those devices. They commissioned a survey of more than 10,000 people who were members of a volunteer Internet panel.

Why did this survey fail? About 30% of Americans do not have access to the Internet, and the Internet panel mostly included very young and highly educated people. This was of particular concern for this study becau health needs and insurance coverage have been shown to differ by age and education. Therefore, the sponso could not accurately assess the demand for the products or the influence of having health insurance on purchasing them.

Coverage error results from all members of the population not having a known, nonzero chance of being included in the sample and from those excluded differing from those included.

A PhD student working on her self-financed dissertation spent many months designing a survey of high schoo students to compare differences across schools. However, the student could only afford to mail out 100 survey with no reminders, for each randomly selected school in her sample. Because of these limitations, it was likely that only 25 to 30 students from each school would respond.

Why did this survey fail? The largest difference the student expected to find among the surveyed groups was percentage points, and most differences were likely to be even smaller. The small completed sample from ea school meant that the margin of error for her estimates (±15%) was larger than the differences she expected between schools, meaning that significant differences could not be detected.

Sampling error results from surveying only some, rather than all, members of the survey population.

A vice president at a major university administered a 40-minute web survey to all faculty, staff, and students a the university. The survey used a standard questionnaire that was used at other universities and that contained long series of questions about bias and harassment repeated for many different minority groups found through the country. In addition the survey required an answer for each question.

Why did this survey fail? Early survey takers became upset at the length of the survey and at being forced to answer questions about minority groups that were not present, or that they did not know about, at their university. This dissatisfaction became a topic of conversation in offices and classes, with the result that oth were discouraged from completing the survey. In the end, only people who were very interested in the topic completed the survey, so only their experiences were represented in the final data.

Nonresponse error results when people selected for a survey who do not respond are different in a way th is important to the study from those who do respond.

To encourage careful thinking about each question in a customer satisfaction survey conducted by a public agency, response scales were varied across questions. On one page, a scale began with very satisfied and ende with very dissatisfied, but a later scale on the same page began with not at all satisfied and ended with complet satisfied (i.e., different labels and ordering of categories). Later, respondents were asked to enter a numeric response into a box with 5 meaning very dissatisfied and 1 meaning very satisfied.

Why did this survey fail? People made mistakes when responding to the questionnaire, sometimes marking a satisfied response while intending to register dissatisfaction because the scales were not listed in the same direction across questions and because their expectation that larger numbers are associated with the higher er of the scale was not met. Additionally, the use of different category labels for different services undermined t ability of the sponsor to compare satisfaction across the services.

Measurement error results from inaccurate answers to questions and stems from poor question wording, survey mode effects, or aspects of the respondents' behavior.

their surveys. Finally, there are many other influences on whether people will respond to a survey, and focusing on only one while ignoring others would seem to unnecessarily limit one's ability to obtain responses.

Nonetheless, as Internet surveying has grown and surveyors increasingly find themselves with e-mail but not postal addresses of respondents, the practice of offering cash payments contingent on receiving a completed questionnaire and a postal address has increased substantially. In addition, panel surveys (discussed in Chapter 9) often recruit people based on promises of such postsurvey payment for each survey. When a European conductor of one such panel was asked in late 2006 about the normal payment, he responded without hesitation, "1 Euro for 10 minutes is the norm." An economic exchange with a singular focus on cash payment contingent upon response is not the model we follow in our social exchange framework.

GENERAL PSYCHOLOGICAL MODELS OF SURVEY RESPONSE

Others have attempted to apply general psychological models to understanding survey response. The foundations of these models come from a variety of sources, many of which (e.g., behavior reinforcement, Maslow's hierarchy of human needs, and cognitive dissonance) were described by Cape (2006) with respect to their potential application to survey methodology. The general premise here is that people are motivated by many different considerations, both extrinsic and well as intrinsic, in developing their responses to different situations.

In an early application of a psychological model to survey response, Groves, Cialdini, and Couper (1992) applied Cialdini's (1984) influence theories to the conduct of interview surveys. Cialdini argued that concepts such as scarcity of opportunity, consistency with previous behavior, desire to reciprocate, enjoyment of task, and social proof (i.e., what other people have done or are perceived as doing in the face of similar opportunities) are all social psychological elements upon which one draws in deciding whether to comply with a request to do an activity. Groves et al. argued that these elements influence decisions about survey requests in particular and, based on this, made the case that they should be built into the survey process whenever possible to encourage higher response.

In another approach, Comley (2006) used transactional analysis, a psychoanalytic theory developed by Eric Berne, to develop specific recommendations for how one should and should not interact with respondents throughout the survey process. In particular, Comley pointed out that to maintain their sense of well-being people need units of positive recognition but that surveyors often use an adult-to-child interactional style that turns respondents off rather than an adult-to-adult style that would give them the posi-

answers are invalid. Please review all the questions on this page" is a
to-child-style message that will likely leave some respondents cold
crease nonresponse.

THE LEVERAGE-SALIENCY THEORY OF SURVEY RESPONSE

Building upon the ideas developed through the application of Ci
work to the survey setting, Groves, Singer, and Corning (2000) dev
the leverage-salience theory, in which they proposed that respond
interview surveys are differentially motivated to respond by different
of the survey (i.e., leverage) and by how much emphasis is put on each
by the surveyor (i.e., salience). For some, cash incentives may be imp
but for others, topic and sponsorship or personal community involv
may be important. Each design feature will have varying amounts of le
on the decision to cooperate for different people. In addition, the am
emphasis that the survey sponsor puts on a design feature (i.e., how sal
sponsor makes it) will interact with the importance the respondent pl
it in affecting the final decision to respond or not. If the respondent is s
influenced to participate by the topic of the survey and the surveyor
the topic very salient throughout the implementation process, the resp
will be highly likely to respond. If, however, the topic of the survey is
off for the respondent and the surveyor makes the topic very salie
respondent will be highly unlikely to respond unless some other as
the survey or the request to complete it exerts enough positive lever
the respondent and is made salient enough to overcome this negative
Using this approach, Groves et al. (2000) argued that advance letters fr
sponsoring organization may have more effect on some than others ba
the legitimacy of the sponsoring organization. They also proposed that
people may respond differently than older people to various incentiv
some may be more greatly affected by whether the interview is by tele
or whether they are given the possibility of a data collection mode tha
more privacy.

The leverage-saliency approach suggests that an overemphasis on a
appeal that is attractive to some potential respondents but not othe
produce serious nonresponse error if it is related to an important v
of the study. Additional research on the leverage-saliency theory wa
ducted to determine whether an appeal based upon topic would pr
nonresponse bias (Groves, Presser, & Dipko, 2004). Although people
ested in the topic were much more likely to respond, the impact on
estimates was mixed. The main implication we draw from this ver
ductive line of research is that it is important for appeals to respond
be broadly based in an attempt to encourage all types of survey reci

SURVEY RESPONSE AS A SOCIAL EXCHANGE

Social exchange theory underlies the tailored design method and provides the overarching framework within which we attempt to identify and implement ways of increasing the likelihood of response. As it was first applied to survey methodology in the 1970s (Dillman, 1978), social exchange, like leverage-saliency, focused on appealing to potential respondents in multiple ways. However, in this revision we have sought to update the social exchange perspective on survey nonresponse with knowledge gained from the various frameworks mentioned previously and, in particular, to place an emphasis on utilizing more mutually supportive response-inducing factors in an attempt to appeal to the many different types of respondents that can exist within any survey population.

At the most basic level, social exchange posits that people's voluntary actions are motivated by the return these actions are expected to, and often do, bring from others (Blau, 1964). People engage in a social exchange with others when the perceived rewards outweigh the expected costs. *Rewards* are what one expects to gain from a particular action, and *costs* are what one will have to give up, or spend, to receive the rewards. Entering into social exchanges requires *trust* that the other party will provide a valued reward in the future, although not necessarily to the respondent. In other words, trust is the expectation that the sponsor will carry through on the promises made.

As discussed previously, social exchange should not be equated with economic exchange. Social exchanges are different from economic ones both because there is only a general expectation of a return and because they involve various types of rewards and costs. In contrast, economic transactions rely exclusively on the exchange of money, where exact monetary values are used to measure worth and explicit time periods are often specified for when the rewards should be provided. In social exchanges, the exact nature of the benefits and when they will be provided are often not specified in advance; instead, they are left open to the discretion of the person providing the reward. For example, after hosting a dinner party for a group of friends, one may expect others to reciprocate but cannot specify when or what type of party they should have or even guarantee that others will have a party at all.

An example of how social exchange differs from economic exchange can be illustrated by the research on providing survey incentives. Small token incentives provided with the request to complete the survey significantly improve response rates, whereas a promise to pay people after completing the survey has minimal or no effects on whether people respond (Church, 1993; James & Bolstein, 1990, 1992; Johnson & McLaughlin, 1990). Providing people an incentive in advance encourages participation because people feel they should reciprocate for the reward they receive by completing the survey.

APPLYING SOCIAL EXCHANGE TO SURVEY D

Social exchange is a subtle but powerful method for motivating pe
spond to surveys and is particularly useful for surveyors because th
that they can typically offer are relatively small. The theory of socia
suggests three key questions about how the design of the questio
the implementation process can motivate people to respond to the

1. How can the perceived rewards for responding be increased
2. How can the perceived costs of responding be reduced?
3. How can trust be established so that people believe the rew
   outweigh the costs of responding?

There are many different ways to establish trust and increase ben
decreasing costs; our suggestions were developed based on tradi
ments of social exchange as developed originally by Blau (1964)
(1961), and Thibaut and Kelley (1959) and on more recent ideas ab
ence in a social context from others whose works were reviewed p

WAYS OF INCREASING THE BENEFITS OF PARTICIPATION

According to social exchange, people are motivated to act by th
they expect to receive. As people receive more and more requests to
surveys, the rewards of participation in any one survey may decrea
who have already participated in one or two surveys may feel they
their part as survey participants (Groves & Magilavy, 1981). Thus,
must work hard to distinguish their survey request from other req
emphasize the benefits of responding to the survey.

Provide Information about the Survey

Sending information to potential respondents about the survey an
results of the survey will be used to benefit them and others can e
survey participation (Groves et al., 1992). In particular, sending i
as prenotice letters and informative brochures that explain why t
is being conducted and that highlight the importance of particip
influence people's decision to participate.

Ask for Help or Advice

Many people feel a sense of reward from knowing they have he
ers (Blau, 1964; Homans, 1961). Appealing to people's helping tend
norms of social responsibility can encourage them to respond to t
(Groves et al., 1992). Survey requests that ask people for their hel
vice demonstrate how respondents' assistance is needed and subor

survey sponsor to potential respondents. For example, Mowen and Cialdini (1980) found that including the phrase "it would really help us out" at the end of their survey request increased survey participation by 19 percentage points.

Show Positive Regard

Similar to asking for people's help, showing positive regard can also encourage people to participate in the survey. Thibaut and Kelley (1959) noted that many people feel rewarded from being regarded positively by another person. Personally addressing contacts, providing a toll-free number to call with questions, and providing various ways for people to respond can show positive regard and respect for potential respondents, which may help motivate them to participate in the survey.

Say Thank You

Verbal appreciation can be an important reward in social exchanges (Blau, 1964). Phrases such as "We appreciate your help" or "Many thanks in advance" can be added in contact e-mails or letters to increase the likelihood of people responding. In addition, a postcard follow-up designed as a thank you for the prompt return of "the important questionnaire we sent to you recently" has been found in some surveys to increase survey participation almost equal to the initial mailing with the questionnaire (Dillman, Christenson, Carpenter, & Brooks, 1974).

Support Group Values

Most people identify with certain groups, such as being an American citizen or a dues-paying member of the Nature Conservancy. By tailoring communications to the survey population, sponsorship, and topic, one can appeal to values shared widely by those in the group. Supporting people's values can convey a sense of reward in them (Blau, 1964). Showing support for shared values also underlies efforts to appeal to potential respondents on the basis of a study's social usefulness (Dillman, 1978; Slocum, Empey, & Swanson, 1956).

Give Tangible Rewards

As discussed earlier, providing potential respondents with token financial incentives (ranging from $1 to $10) with the survey request has been shown to significantly increase the number of people who respond. Other types of incentives, such as charity donations and ballpoint pens, have also been shown to have modest effects on response when provided with the request to complete the survey. Providing token incentives in advance evokes a sense of reciprocal obligation such that people feel the need to respond to the reward

Make the Questionnaire Interesting

Designing questionnaires with questions that a wide variety of p
interesting will encourage higher response rates. Liking to do
can be a powerful determinant of human behavior (Cialdini, 19
the topic of the survey is highly salient to potential respondent
more likely to complete the questions than when the topic is of lo
(Heberlein & Baumgartner, 1978). Even when not all potential re
may be interested in the topic, questionnaires can be made more i
by improving their visual layout and design, ordering questions so
engaging ones are placed early in the questionnaire, and crafting
that are easy to understand and answer.

Provide Social Validation

Knowing that other people similar to themselves have completed
can strongly influence people to participate in a survey (Groves et
Because people frequently use the actions of others as standards c
ison for their own actions and feel rewarded when they see ther
similar to most others in a group, they are more likely to comply wit
when they believe others like them would as well. For example, in
contacts, telling people that many others have already responded e
them to act in a similar way and respond to the survey.

Inform People that Opportunities to Respond Are Limited

People perceive rewards as more valuable when opportunities becc
(Groves et al., 1992). Telling people, in a friendly and nonpatroni
that there are relatively few opportunities to respond and that the
have an opportunity to respond unless they do so quickly can
them to participate (Petrie, Moore, & Dillman, 1998). Likewise, exp
potential respondents how only a small number of people are selec
survey can have a similar motivating effect.

WAYS OF DECREASING THE COSTS OF PARTICIPATION

Oftentimes reducing the costs of participating in a survey is close
to increasing the rewards. However, there are also particular strat
surveyors can use to decrease the perceived costs of responding to
independent of rewards.

Make It Convenient to Respond

Perhaps one of the most effective ways of decreasing costs is making
as possible for people to respond. This may involve offering a desi

and including a link that, when clicked, will open their browser and take them directly to the survey to make it more convenient for them. Likewise, including a prepaid return envelope with postal questionnaires has been shown to increase response rates because it is easier for respondents to return the completed questionnaire when they do not have to find, address, and stamp an envelope (Armstrong & Luske, 1987).

Avoid Subordinating Language

People prefer not to feel that they are dependent upon others, and Blau (1964) argued they will expend great effort to avoid feeling subordinated. Consider these contrasting statements that might be included in a letter or e-mail to potential respondents:

- "For us to help solve the school problems in your community, it is necessary for you to complete this questionnaire."
- "Will you please be a part of helping to solve the school problems in your community? Your responses can assist this community in fully understanding the issues facing schools here."

The first statement subordinates the respondent to the surveyor using what Comely (2006) might consider an adult-to-child style, whereas the second statement makes the respondent feel that the surveyor is dependent on him or her. Asking a person for help or assistance subordinates the sponsor to the potential respondent rather than vice versa and is one way to decrease the costs and increase the rewards of participation.

Make the Questionnaire Short and Easy to Complete

Questionnaires that appear short and easy to fill out reduce the perceived costs of responding. For most people, one of the biggest costs of responding to survey requests is the time it takes to complete the survey. Research has shown that longer questionnaires achieve slightly lower response rates (Heberlein & Baumgartner, 1978) but does not confirm that using more pages for the same number of questions decreases response rates (Leslie, 1997). Other research has shown that respondent-friendly questionnaires, with carefully organized questions in easy-to-answer formats, can improve response rates (Dillman, Sinclair, & Clark, 1993). Thus, designing questionnaires to enhance usability and minimize respondent burden, as well as keeping questionnaires short, can decrease the costs of responding to the survey.

Minimize Requests to Obtain Personal or Sensitive Information

Many survey questions ask for information that some people do not want to reveal to others. For example, surveyors often ask people about their income and other financial information, their health and medical history, and their

past sexual behavior or drug use. Surveyors should only incl
for sensitive information when the responses are essential to
objectives. When these types of information are an importan
survey, efforts can be made to provide explanations for why respo
questions are important and how the information will be kept
or even anonymous. In addition, the choice of question wordi
"soften" the requests for personal information.

Emphasize Similarity to Other Requests or Tasks to Which a Person H
Already Responded

People have a strong desire to appear consistent in their attitu
and actions. Therefore, people who have committed themselves
are more likely to comply with requests to do something consiste
position (Groves et al., 1992). This inclination to behave consist
that arguments can sometimes be offered that point out how
to a particular survey is consistent with something one has al
For example, a survey of members of a particular organization r
the following statement: "We really appreciate your support t
recent payment of dues, and we want to be responsive to your e
Completing this web survey will give us guidance on how best t
and your fellow members."

The need to be consistent may also explain why, in panel su
people respond to the initial request, it is much easier to get them
to subsequent requests (Otto, Call, & Spenner, 1976). Consistenc
explain why the "foot-in-the-door" technique, where people are
to perform a large task if they are first asked to perform a sr
is effective (Mowen & Cialdini, 1980). A survey of national pa
successfully used this technique by first asking people to respo
short questions upon entering the park and then asking them to
questionnaire at the end of their visit (Dillman, Dolsen, & Machli

WAYS OF ESTABLISHING TRUST

Because social exchange involves future obligations, one must tr
other party will follow through and provide the return or reward a
Under conditions of social exchange, there is no way to ensure tha
survey sponsor has promised as a benefit of the study will actu
For example, if a surveyor says "This survey will help our com
better job of providing service to its customers," or "This survey
state legislatures make decisions about how to allocate funding
education," there is no way to guarantee that the results of the s
actually deliver the return as expected. Thus, potential respondent

the rewards as promised. Trust is critical to believing that in the long run the benefits of completing the survey will outweigh the costs of doing so.

Obtain Sponsorship by Legitimate Authority

People are more likely to comply with a request if it comes from an authoritative source that has been legitimized by larger society to make such requests and expect compliance (Cialdini, 1984; Groves et al., 1992). Therefore, it is not surprising that government-sponsored surveys achieve higher response rates than those sponsored by marketing research firms (Heberlein & Baumgartner, 1978). In addition, the unique authority that some government agencies have to inform people that their response to a survey is mandatory helps improve response for government surveys of both businesses (Tulp, How, Kusch, & Cole, 1991) and individuals (Dillman, Singer, Clark, & Treat, 1996). A more complete discussion of sponsorship can be found in Chapter 11.

Provide a Token of Appreciation in Advance

A few dollars included with a survey request increases rewards, but it also creates value in the social exchange process by establishing trust. By providing the incentive with the request, before the survey is completed, the researcher demonstrates trust in potential respondents—who can pocket the money without completing the survey—and encourages their trust by demonstrating that the surveyor will provide the promised rewards. In addition, emphasizing that the incentive is a "small token of appreciation" is consistent with conveying trust and respect for the respondent.

Make the Task Appear Important

Many surveys try to appeal to people on the basis that something important will ultimately happen as a result of the survey. Making each contact appear important can help establish trust in the survey sponsor and that the results will have the impact the surveyor says they will. Printing personalized cover letters on letterhead stationery, including a carefully chosen color picture on the front of the questionnaires, and providing brochures or other materials about the survey project can make the survey appear credible and help establish trust in the survey sponsor. In contrast, form letters produced on copy machines and questionnaires that are poorly designed or contain questions that are difficult to understand suggest that a survey, and the sponsor, is relatively unimportant.

Ensure Confidentiality and Security of Information

Of considerable concern for some survey respondents is how the information they provide will be used and who will have access to it, particularly if they are disclosing information that is personal or sensitive. The rise of the

transmitted via the Web, especially with respect to whether su
guarantee that people's responses will remain confidential and
way of establishing trust is by explaining the efforts that will
ensure the confidentiality and security of people's survey respo

Box 2.1: Trust at the Industry Level

The issue of trust is one that extends far beyond any one survey
respondent, group of respondents, or even survey project. Establi
maintaining trust is a fundamental concern for the survey industry a
and it is an area in which the industry is constantly being challeng
following example demonstrates.

The year 2004 saw one of the closest state gubernatorial electio
history, with the disputed winner edging out her opponent by a mere
about 2.9 million votes cast. The final result was a contentious one, rea
after three ballot recounts and a state supreme court ruling. During th
wrangling over the election, some absentee voter ballots were qu
resulting in these voters signing affidavits to ensure their ballots were
The election outcome was still in contention in January and Febru
when backers of the losing candidate sent a "Home Ownership Surve
than 400 residents who had signed the affidavits in one particularly
county in the state. Along with the three-question survey, the sponso
$10 incentive check. The survey, it turned out, was a ploy. When res
returned their signed surveys or endorsed and cashed their incenti
these signatures were compared to the signatures on their ballot
If they did not match, the sponsors alleged that the respondent's
fraudulent (Postman, 2005). When the story of this survey ploy was
largest newspaper in the state, several respondents explained that
different signatures for different purposes. Others explained that the
signed and cashed the check for them. Nearly all were upset at the
nature of the survey, as undoubtedly were many un-surveyed reside
state who learned of the ploy.

This example, although somewhat sensational, is only one of man
some scams utilizing surveys are so common that they have their ov
within the industry.

- *Frugging: Fundraising under the guise of survey research.* In 2
  the Republican and Democratic parties sent voters "surveys" a
  their views on central party issues. Embedded within the surve
  ever, were appeals for sizable financial contributions to the partie

(c

tactics have been used by a whole slew of groups ranging from local media outlets to wildlife and parks organizations.

- *Sugging: Selling under the guise of survey research.* A prominent DVD marketing company sent out a survey and as thanks for filling out the survey told respondents they could pick 5 DVDs of their choice for just $0.49 each. The fine print, however, revealed that those who returned the survey would be automatically entered into a club and obligated to buy more DVDs.
- *Phishing: Tricking people into providing personal information, usually over the telephone or Internet.* In 2007, unsuspecting taxpayers were sent e-mails that appeared to come from the Internal Revenue Service. The e-mail explained that they had been randomly selected to participate in a customer satisfaction survey and promised to pay $80 for their participation. Embedded in the survey were requests for personal information that would presumably be used to contact respondents and transfer the $80 to them. Instead, the information was to be used to access the respondents' bank accounts, run up their credit card bills, and take out loans in their names. The respondents were to become victims of identity theft.
- *Push polling: "A form of negative campaigning disguised as a political poll"* (AAPOR, 2007). A recent one-question poll of 9,000 people within a school district asked whether respondents supported a construction project "that will result in higher taxes, while not improving education." The goal of the poll, admittedly, was not to collect public opinion. Rather, it was to back certain candidates in a local school board election by influencing public opinion about the construction project.

Each of these self-serving scams abuses the public's trust and undermines the credibility of legitimate survey research. As such, each impacts the ability of all survey sponsors, from the U.S. Census Bureau to the unfunded graduate student, to conduct good survey research. Therefore, as individuals, we all have a vested interest in working to eliminate these scams. Such scams are addressed by the following organizations: the American Association for Public Opinion Research (AAPOR), the Council for Marketing and Opinion Research (CMOR), and the Council of American Survey Research Organizations (CASRO).

THE IMPORTANCE OF USING A VARIETY OF APPEALING SURVEY FEATURES

One of the features of a social exchange that is sometimes overlooked is how communication between people should change over time as one person tries

request over and over in everyday conversations is unconventio
quite irritating, and is unlikely to be effective. The following set
can be illustrative here:

Week 1: Ted, I need the book that I loaned to you.
Week 2: Ted, I need the book that I loaned to you.
Week 3: Ted, I need the book that I loaned to you.
Week 4: Ted, I need the book that I loaned to you.

In this case, the request for Ted to return the book is being made i
way over and over. However, if Ted does not comply with the
first time, the stimulus in this type of request is used up and, t
unlikely to be effective the second and third times it is used. Anoth
requests are often made is illustrated in the following set of appe

Week 1: Ted, this is to let you know that I will require that new H
book that I loaned to you back by Friday.
Week 2: Ted, I just wanted to remind you that you have not br
the Harry Potter book that I informed you needed to be re
Friday.
Week 3: Ted, did you forget to return the Harry Potter book that i
It was not that I really needed it then, but some people do not
quickly as they should.
Week 4: I suppose you still have not remembered the Harry Pot

Here the request is varied throughout the contacts, but each include
demanding, and somewhat patronizing wording that is reminis
adult-to-child conversational model described by Comley (200
most insulting is the final communication, which says in a defe
that it is okay not to respond. In contrast, the following set of requ
a more conventional, positive way of encouraging Ted to return t

Week 1: Hi Ted, have you had a chance to read that book? I ho
enjoying it.
Week 2: How are you doing with the new Harry Potter book? C
asking me if I had a copy, and I said I would check and see i
finished.
Week 3: Hey Ted, did you watch the playoff game last weeke
way, when you are finished with the Harry Potter book, I ha
one that I think you would like.
Week 4: I need your advice on a project report. Carolyn thought
be helpful on it. Oh, she also asked me about the Potter book.

Here the stimulus is varied over time, and Ted is respected in each contact. This final strategy is consistent with both the leverage-salience approach to making requests and the tailored design approach.

Although survey requests do not lend themselves to the informal conversation and introduction of side topics to the extent shown in this final set of examples, it is equally important to vary the stimulus in order to appeal to different types of respondents and to write requests in an adult-to-adult conversational style. Comments from respondents to a recent mail survey illustrate well the importance of attending to multiple design and implementation features in order to attract different types of respondents. In the fall of 2006, the survey "Family Farming and Ranching in Washington: A Woman's Perspective" was sent to a random sample of cattle and wheat operations across Washington State. Several measures were taken in the implementation of this survey to make its appeal as broad as possible. First, both the survey title, which was displayed prominently on the cover page of the 12-page booklet-style questionnaire, and the message in the accompanying cover letter stressed the importance of hearing the opinions of farm and ranch women. A shortened description of the survey ("Women's views on farming and ranching survey") included in the return address carried this same message over to the envelope the survey was sent in. Second, an appealing picture of a mailbox mounted decoratively on a very old piece of farm equipment and set next to a green wheat field was displayed on the front cover. Third, a $2 token incentive was included with the initial questionnaire mailing. Fourth, because Washington State University (WSU) is the well-respected land grant university in the state and offers many services to the state's agricultural sector, and because the survey was being conducted from there, the WSU logo was included on the outgoing envelopes and the letters were printed on university letterhead. Finally, the back cover of the 12-page booklet-style questionnaire was devoted almost entirely to space for respondents to write any additional comments or opinions they may have had that could not be expressed in the more structured survey questions (Smyth, 2007; for more information, see "Family Farming and Ranching in Washington Research Results" at www.crs.wsu.edu/1-07-farmranchwa.pdf).

Several respondents took advantage of this final feature to make comments that can shed light on what features of the survey and implementation materials convinced them to respond. Some comments included "Go cougs!" (a phrase commonly used by supporters of WSU and its Cougar athletic teams, which was left by several respondents), "Thank you for giving me the opportunity to vent!" "Thank you, thank you for asking farm women for their opinions on their farm lives," "Thank you for the $2 bill. Here I can only afford my 2¢," and "Your $2.00 could have been sent in pennies. It would have taken me less time to count the @%#& pennies than to fill out your

woman's voice heard." These comments suggest that for some
it was the connection to WSU that influenced their decision to
others the focus on women's opinions—communicated in the s
the cover letter, and on the outgoing envelope—was importan
the $2 bill caught their attention and made a positive impressio
were offended by the $2 bill but felt the topic of the survey w
enough that they should respond.

This last case, in particular, demonstrates very strongly the i
drawing on multiple features of questionnaire design and imple
persuade as many sampled respondents as possible to complet
the survey. Yet when looking at letters used to encourage peop
to surveys, one often sees virtually, if not completely, identica
for multiple follow-ups, and in many instances the letters are s
distant that all aspects of humanness on the surveyors' end
get omitted. Another common tendency is for surveyors to u
in both the initial contact and the follow-up contacts. Howev
with the argument we are making here, research has shown
survey situations the repetition of an incentive in a questionnaire
mailing does not increase response over what can be obtained wit
alone (Tortora, Dillman, & Bolstein, 1992). Thus, as both the leve
approach and the tailored design perspective would suggest, it
to change the look, feel, and content of later contacts rather tha
same requests over and over.

DEVELOPING A TAILORED SURVEY DESIG
THE SITUATION

In considering how to apply social exchange ideas for tailored
important to realize the breadth of the opportunities for chang
to increase rewards and trust and minimize costs. Tailored desig
exchange are not limited simply to the cover letters that one
the recipient. In fact, beautifully composed and executed cove
be nullified by poorly designed questionnaires, poor timing, a
other poorly designed survey features. Here are some examples
we have received that illustrate a singular focus:

- I need responses to a web survey within 24 hours. Is it okay
  reminders the same day we send the first e-mail?
- I am only going to send one mailing. I need to know ho
  response rate will improve if I use a colorful commemorati
- We have a lot of nice pictures. Will using them improve our re
- What color paper or background colors should I use to
  response?

These requests and dozens more like them have included only a few details (and sometimes none at all) about other aspects of the survey plan, but the answers we might provide will often differ depending on other aspects of the survey design. Generally, we would like to know who is being surveyed and by what mode or modes, how respondents will be contacted, how many contacts are planned, the anticipated timing of contacts, what (if any) incentive will be sent, what the topic of the survey is, and how many questions will be asked. For the specific inquiries above, we might also ask why the survey has to be done in one day, why only one mailing is being used, what else besides nice pictures will be used, and so on.

Once we understand more about the overall survey design, we can better answer people's questions about particular aspects of their survey and make practical suggestions to fit their situation. When asked about whether they should contact nonrespondents by telephone, we might suggest instead including a token incentive with the first contact. Or after hearing about their overall design, we might conclude that a telephone follow-up would work or that they should also offer a web option to encourage people to respond who may not otherwise. Similarly, we might recommend varying the background color in a web survey because the cost is minimal, but designing a mail survey with the questionnaires printed on white instead of colored paper to save money and so that the front page can include a color picture that resonates with members of the population.

When developing a tailored survey design for self-administered questionnaires, surveyors need to consider a number of survey features and their effects on potential respondents and to keep in mind that when one specific social exchange feature cannot be used in a particular survey situation, many others can likely be used instead. In Figure 2.2 we present some examples of essential features that should be considered when devising a survey plan and different ways each feature can be tailored to the particular survey situation. As the figure suggests, an effective survey design is the result of many decisions that need to fit together and support one another in a way that encourages most people to respond. As a result, although the merits of each individual survey feature should be considered carefully given one's particular survey situation, it is also important to consider how each feature fits into the larger survey system. After all, whether an action evokes a sense of cost, reward, or trust is related to how it interacts with other features of the system, not just how it appears in isolation (e.g., a well-designed questionnaire may never be viewed if it is mailed in a poorly designed envelope that makes it look like junk mail).

Applying the tailored design approach to each feature in Figure 2.2 requires consideration of rewards, costs, and trust as well as recognition that some survey features may invoke multiple social exchange elements within one's

Figure 2.2 Features of the survey design that can be tailored to the situ

Survey mode	• Choice of mode or any combination of modes
Sample	• Type of sample (random, stratified, etc.) • Number of units sampled
Contacts	• Number of contacts • Timing of initial contact and between contacts • Mode of each contact • Whether each contact will be personalized • Sponsorship information • Visual design of each contact • Text or words in each contact
Incentives	• Type of incentive • Amount or cost of incentive • Whether to provide before or after the survey is completed post)
Additional materials	• Whether to provide them at all • Type of materials (brochures, pamphlets, research reports, • Visual design of the materials • Text of the materials
Questionnaire	• Topics included • Length (duration, number of pages/screens, number of que • First page or screen • Visual design and layout of pages/screens • Organization and order of questions • Navigation through the questionnaire
Individual questions	• Topic (sensitive, of interest to the respondent, etc.) • Type (open vs. closed) • Organization of information • Text or wording • Visual design

is sent in advance it also promotes trust). Because what constitutes a reward or cost often depends on the survey context, it is important to take into account and plan around differences in survey content, sponsorship, and populations through the use of different social-exchange-inducing features.

For most surveys, for example, it is typical to receive only one or maybe two contacts in a 10-day period; any more begin to become irritating. However, people who have been asked to keep a diary of the specific television programs they watch in a given week might not find it unusual or too burdensome to be contacted five times in a 10-day period. First they would be mailed an introductory letter, followed quickly by a questionnaire, and then a postcard to tell them their "diary week is about to begin." Then midweek they might get a call to see if they are having any difficulties, and finally another postcard would be sent indicating that their diary week is over and thanking them. This sequence, which is similar to the one used for many national television viewing and other diary surveys, makes perfect sense in this context in which respondents are being asked to record their behavior each day for a specific week. In other survey situations, this same contact strategy may have the opposite effect and turn people off because it is so intense.

Another example is that sending a token financial incentive to individuals can greatly improve response; however, sending incentives to business respondents may not have the same effect because they may think that there is an ethical problem or that they may have to complete extensive paperwork to accept the money. In this case, what seems like a reward from the sponsor's perspective may actually produce a cost for some recipients.

How a survey is designed and administered should also depend on the length and topic of the questionnaire, as both influence respondents' perceptions of rewards and costs. Longer questionnaires or those about particularly dry or sensitive topics increase costs to the respondents. Getting people to respond to these types of surveys often requires extra effort, perhaps an added incentive, extra explanation as to why their participation is important, additional follow-up contacts, or administration through a survey mode that provides more anonymity (i.e., self-administered instead of interviewer administered). Adding more interesting questions, changing the order and format of questions to make them more interesting, or easing into sensitive topics are other tactics that might be used with these types of questionnaires. As a slightly different example, a topic requiring a respondent to read a detailed statement before answering or to evaluate a lengthy list of response options may be better suited for self-administered surveys where the respondent can review the information several times, if needed, before responding.

Sponsorship can also influence people's motivations to respond. Government sponsorship of surveys is likely to improve response, whereas market

can appeal to legitimate authority, avoid the use of financial incentiv
often require that people respond. In contrast, market research organi
generally have little "legitimized" authority, so they may need to s
centives for people to respond. Sponsorship can also interact with q
topic to influence participation decisions. For example, it is understa
that market researchers might ask about one's affinity for different
and where one likes to shop, but it might seem more of an intrusion
ask for detailed financial information. However, people might be wi
provide detailed financial information to the U.S. Census Bureau.

Tailored design also involves tailoring procedures to the particul
ulation being surveyed. Knowledge about the target population o
similar populations can help surveyors develop effective design str
that motivate response. For example, some populations can only b
questionnaires if they are handed to them in person: visitors to muse
national parks, or voters exiting the polls. In some situations, mem
the population may only have access to one mode of communication
greatly limits how they can be contacted. In other situations, it may
possible to survey all of those sampled from the population using a
mode, so multiple modes may need to be used. Compared to surveys
general public, surveys of more specialized populations, such as st
employees, or military personnel, usually achieve higher response ra

In addition to these larger population considerations, it is impor
consider differences among potential respondents within the populati
how different aspects of the survey design may have different amo
leverage on respondents' decisions to participate (Groves et al., 2000).
deciding whether to respond to a survey, each person evaluates the
request based on his or her own values and beliefs, positions in s
current emotional state, and other psychological dispositions. Beca
these differences, each feature of the survey design may be more imp
to some people than to others and can have positive or negative infl
on people's decisions to respond. A summary of the tailored design r
can be found in Figure 2.3.

The examples that follow show how some surveyors have succe
tailored their surveys to address the various types of issues that hav
introduced throughout this chapter and how multiple concerns hav
brought together to achieve desired results.

- An organization was interested in using the Internet to the extent
  ble (i.e., to reduce costs) to survey a large number of respondent
  the general public on a variety of topics. To get a probability s
  they used a random-digit-dial telephone contact followed by
  questionnaire to create a panel of many thousands who were wil
  complete surveys every 2 to 3 weeks either by mail or Web dep

Figure 2.3 Overview of the tailored design method.

A. Tailored design is the development of survey procedures that work together to form the survey request and motivate various types of people to respond to the survey by establishing trust and increasing the perceived benefits of completing the survey while decreasing the expected costs of participation.

B. Successful tailored design attends to the multiple sources of survey error—coverage, sampling, measurement, and nonresponse—with a focus on minimizing overall survey error.

C. Tailored design involves customizing survey procedures for each particular survey situation based on knowledge about the topic and sponsor of the survey, the types of respondents who will be asked to complete the survey, and the proposed budget and time frame for reporting the results.

D. Multiple aspects of the implementation process and the questionnaire can be combined in different ways to encourage respondents to participate by creating trust in the sponsor and influencing the perceived expectations of the benefits and costs of responding to the survey.

To establish trust	To increase benefits of participation	To decrease costs of participation
▪ Obtain sponsorship by legitimate authority	▪ Provide information about the survey	▪ Make it convenient to respond
▪ Provide a token of appreciation in advance	▪ Ask for help or advice	▪ Avoid subordinating language
▪ Make the task appear important	▪ Show positive regard	▪ Make the questionnaire short and easy to complete
▪ Ensure confidentiality and security of information	▪ Say thank you	▪ Minimize requests to obtain personal or sensitive information
	▪ Support group values	▪ Emphasize similarity to other requests or tasks to which a person has responded
	▪ Give tangible rewards	
	▪ Make the questionnaire interesting	
	▪ Provide social validation	
	▪ Inform people that opportunities to respond are limited	

on their Internet access, frequency of computer and Internet use, and stated preferences (i.e., they reduced coverage error by including both people with and without Internet access). The surveys by mail and Web (two visual modes) contained identically worded questions and were constructed similarly to minimize potential measurement differences. To maintain interest in the panel and to encourage respondents to complete surveys (i.e., to minimize nonresponse and panel attrition), they routinely thanked panel members for their responses, occasionally sent token gifts, and sometimes gave summaries of results from earlier surveys in which the respondents had participated. For more information

- A graduate student wanted to survey community leaders th two states for her dissertation, but like many other PhD stuc faced budget constraints because her project was self-financed sampled communities across the two states and then used th to identify key community officials who she would then ask to the 12-page mail questionnaire containing 50 questions (some parts). To distribute her labor and financial costs over severa she divided the sample into three subsamples, which made i for her to do all of the mailing preparations herself, includi ple follow-ups (i.e., reducing nonresponse). Follow-ups wer mail in most cases; however, to get several sample members t she had to send the questionnaire as an e-mail attachment. Se same questionnaire, but with a different mode of delivery, help nonresponse while minimizing measurement error. She comp survey with a 72% response rate (360 of 500; Crowe, 2008).
- A 25- to 30-question annual Internet survey about the under student experience at WSU was conducted six times from 200 with response rates consistently reaching the 50% to 60% level number of shortcomings in the sample frame from which stud randomly selected. In particular, many students did not up contact information after being admitted to the university. M number of students did not use their university-provided e-mai and only 60% to 70% provided the university with a differe address through which to contact them. As a result of both problems, implementation procedures were developed whereb $2 token of appreciation was sent (i.e., reducing nonresponse) mail (i.e., reducing coverage error that would have resulted fr e-mail contacts) along with letters requesting that students Web and complete the survey (i.e., using only one data collect eliminated the possibility of mode effects in the data). For stud known e-mail addresses, an e-mail was also sent after the letter t a more convenient hyperlink for accessing the survey (i.e., nonresponse). Finally, up to two follow-ups by postal and/ were sent at 2- to 3-week intervals to encourage students to resp Smyth, Dillman, Christian, & Stern, 2005, for a more complete de of the methods used).

None of these surveys were perfect; however, each represents a c effort to contemplate and minimize errors due to coverage, sampl response, and measurement. Each of them involved multiple con were carefully constructed to encourage the sampled individuals to Some involved the Web only, some mail only, and some used both m

the particular study, respondent preferences, what contact information was available, and the resources budgeted for the project.

CONCLUSION

Tailoring to the population, survey sponsorship, mode considerations, and other features of one's survey situation is not something one does just in the cover letter or e-mail. Nor is it something one does in just the questionnaire or in the decision to use or not use an incentive. It starts when one is making a decision to do a survey and should occur throughout the entire process of designing and implementing the survey. It is not something one does by attending to only a few parts of the survey process; it is something one thinks about when designing each and every part.

Recently we described the use of the social exchange framework and tailored design to a class of students. After summarizing the major trends of the past decade toward greater impersonalization in the survey process, less time spent focusing on each respondent as an individual, greater automation of the survey process, and the use of mass appeals that now characterize many web surveys, one student seemed perplexed. She then asked if applying social exchange was consistent with the times or if attending to so many details was akin to getting water to run uphill. Her question was reminiscent of the time when the original total design method book was written, and when the mass society technologies of the time made focusing on individual respondents difficult. Our response: Now that we have improved technological ability, perhaps part of the solution to the growing problem of decreasing response rates and their effect on nonresponse error is to change the trajectory of how we as survey researchers attempt to relate to respondents and their needs.

CHAPTER 3

Coverage and Sampling

A GRADUATE student came into the office one day and announced thusiasm, "I now know how to do the perfect survey on discrimin harassment against minorities at this university." This particular st very comfortable with large data sets, had recently learned to pro surveys effectively, and understood that access could be obtained sity e-mail address files. The rest of the conversation unfolded in the way:

QUESTION: How will you make it perfect?
ANSWER: It's simple. Instead of surveying a sample of some kir just get all of the faculty, staff, and student e-mail addresse them all to fill out a web survey. It's only 25,000 people, an that big and even bigger are now done all the time because it to send mass e-mail and collect data on the Web.
QUESTION: Can you get everyone's e-mail address, and will al use their e-mail?
ANSWER: Well, the university gives everyone an e-mail addres must be able to use them.
QUESTION: What about some of the staff who work at the phys and others who choose not to use their university e-mail addr also, what about people who have kept their e-mail address aft the university?
ANSWER: I don't think that would be a big enough problem to wo and a good survey can't do everything.
QUESTION: Do you think you would get a good response rate?

developed, with formulas for measuring sampling error and determining needed sample sizes that are applicable to most types of surveys and survey populations. In addition to providing an introductory discussion of sampling theory, we have attempted to offer some strategies for putting it into practice. Yet for a survey to truly be successful, one has to give attention to other sources of error as well: All four sources of survey error have to be reduced to acceptable levels. In the next four chapters we turn our attention away from coverage and sampling error and focus instead on strategies for minimizing measurement error through effective survey design and nonresponse error, first through survey design and then through implementation procedures.

CHAPTER 4

The Basics of Crafting Good Questions

ON THE surface, writing a question may seem simple. For example: "
many hours per day do you typically study?" Yet the decision to ask s
question immediately raises a whole host of issues. Will the question pr
different answers in different modes? Should one provide answer cate
or an open space where respondents can write their answer howeve
desire? If one chooses to provide answer categories, what should the
how should they be ordered, and how should they be arranged on the

Underneath these questions are more fundamental methodological
tions. From what does a respondent draw meaning when reading and
preting a question? How can all parts of the question work together? Ho
the words and the visual layout influence how respondents comprehen
ultimately respond to the question? Does it make a difference what the
cific categories are? Are answers influenced by providing other categ
Do some question structures produce higher nonresponse than others?
challenges are illustrated by a recent experiment we conducted (S
Dillman, & Christian, 2007a).

We asked three randomly selected subsets of a sample of student
number of hours per day they studied in three different ways:

1. A low range consisting of five categories from 0.5 hours or less
   to 2.5 hours plus a sixth category for those who studied more tha
   hours.
2. A high range consisting of an option for those who studied 2.5 h
   or less plus five categories for those studying from 2.5 to more

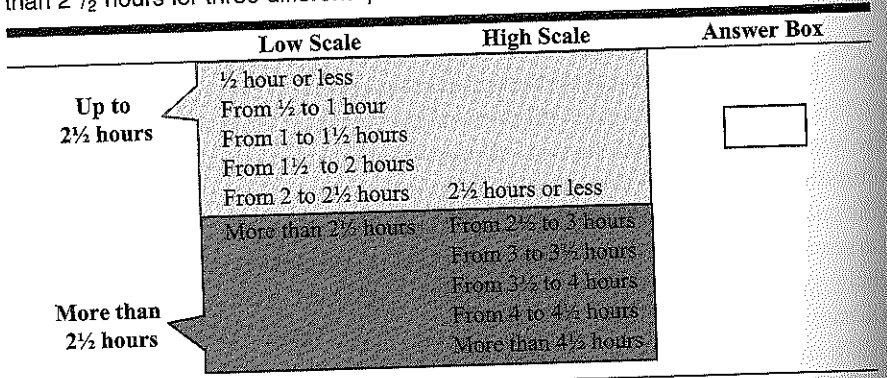
3. An open-ended answer box with no categories at all so that the respondent could choose what to report.

Technically, all three ways of asking the question could accommodate students studying anywhere from 0 to 24 hours, but as Figure 4.1 demonstrates, the scales emphasized different, slightly overlapping portions within that range whereas the answer box did not emphasize any of the range.

As a result of these differences in construction, the percentage of students reporting in a web survey that they study more than 2.5 hours per day ranged from 30% when the low scale was used to 71% when the high scale was used. This is a difference of 41 percentage points! The answer box version produced a more moderate estimate: 58% of students reported studying more than 2.5 hours. When the same experiment was conducted over the telephone, the results were very similar in that students who received the high scale were much more likely to report studying more than 2.5 hours per day than were those who received the low scale.

A previous mail experiment reported in the second edition of this book found a similar result, with 23% of respondents to the low-scale version reporting that they study 2.5 hours or more per day and 69% of respondents to the high-scale version reporting studying this much (Dillman, 2000a,

Figure 4.1 Percentage of students reporting studying $2^1/_2$ hours or less and more than $2^1/_2$ hours for three different question formats.



Percentage Reporting Studying Up to and More than 2½ Hours by Format			
Up to 2½ hours	70%	29%	42%
More than 2½ hours	30%	71%	58%

Source: "Context Effects in Web Surveys: New Issues and Evidence" (pp. 427–443), by J. D. Smyth, D. A. Dillman, and L. M. Christian, in *The Oxford Handbook of Internet Psychology,* A. Joinson, K. McKenna, T. Postmes, and U. Reips (Eds.), 2007a, New York: Oxford University Press.

pp. 32–34). Clearly, the way these questions are constructed influence ability to accurately measure the number of hours students study. But What is producing such differences?

If respondents use only the words in the question stem and the cate numbers to interpret the meaning of the question, the answers across three versions of the question should be the same. Because the answers widely, it is apparent that respondents are drawing extra information the response categories and using that information to help formulate answers. In other words, respondents are using information well be the numbers that are provided to define the parameters of each resp option.

For many students, knowing exactly how many hours per day they s is not something they can recall in the same way that they can recall wh they live in a student dormitory or own a car. As a result, they probably ha estimate how many hours per day they study, and doing so probably req them to average across days during the week and on weekends and prob even across times of year (i.e., a normal day vs. a day during the week of exams). When respondents have to do this type of mental work to form their answers, they often look to the question and its accompanying resp options for clues.

When asked how many hours per day they study, respondents migh sume that the range emphasized by the scale represents how many h most students study. As a result, someone who gets the low scale might clude that *most* students study between 0.5 and 2.5 hours, whereas som who gets the high scale might conclude that most students study betv 2.5 and 4.5 hours. Another assumption that respondents may make is the middle option(s) represents the amount that the *average* student stu thus, those who receive the low scale might assume that the average dent studies between 1 and 2 hours, but those who receive the high s would assume that the average student studies from 3 to 4 hours. Ra than actually counting the hours they study, respondents can instead de whether they study more, the same, or less than most typical students. In type of estimating, different assumptions made based on the scale range midpoint are bound to influence answers.

This example illustrates the challenge of crafting good survey quest that every potential respondent will be willing to answer, will be abl respond to accurately, and will interpret in the way the surveyor inte Stated quite simply, one must think about many things at once to wri good question, and failure to do so can have significant effects on how question performs. Factors to consider include what type of question to w (e.g., open- vs. closed-ended, single vs. multiple answer, etc.), how to word question stem, what response options to offer and how to word them, h

to visually present the questions, what type of answer spaces to provide, and whether and where to provide additional sources of information (i.e., instructions).

In this chapter, we address how to answer these questions and others that may arise when developing survey questions. First we describe a holistic approach to crafting survey questions that considers what question format is best and how multiple aspects of the question wording and layout need to work together to reliably provide accurate data about the concept of interest. Then we discuss general guidelines for question wording and visual presentation that apply to nearly all survey questions. In the chapter that follows, we turn to more specific guidelines for particular types of questions. However, before we start with guidelines, we discuss four issues that need to be considered for each question.

ISSUES TO CONSIDER WHEN CRAFTING SURVEY QUESTIONS

1. WHAT SURVEY MODE(S) WILL BE USED TO ASK THE QUESTIONS?

How one writes a survey question should depend strongly on how that question is going to be delivered to respondents. The key point to keep in mind here is that different survey modes rely on different communication channels. In telephone interviews, respondents give and receive information through spoken words and the hearing system, whereas on the Web and in mail questionnaires, information is transmitted through the visual system. As a result, words take on extra importance in telephone surveys, and memory becomes a significant factor to be considered. In mail and Internet surveys, visual design elements become important. In this chapter we focus mostly on writing questions for mail and Internet surveys, with only brief comments on telephone surveys. We discuss writing questions for mixed-mode surveys in Chapter 8.

2. IS THIS QUESTION BEING REPEATED FROM ANOTHER SURVEY, AND/OR WILL ANSWERS BE COMPARED TO PREVIOUSLY COLLECTED DATA?

The answer to this question will influence how much, if any, the question can be changed. If a particular question has been used in another survey and the main objective is to replicate the previous survey or make the new results comparable in some other way, usually no changes or only minimal changes can be made. Examples are government surveys that have asked the same question repeatedly, sometimes for decades, to produce time-series data. For self-administered surveys, this means trying to replicate not only the question wording but also the other aspects of the visual design and

layout of the questions. Thus, it is important to ask whether question be repeated from other surveys or previous waves of data collection so, whether they can be changed.

3. WILL RESPONDENTS BE WILLING AND MOTIVATED TO ANSWER ACCURATELY?

Ensuring that respondents are motivated to respond to each questic major concern in self-administered surveys because there is no interv present to encourage respondents to carefully select and report cor answers. Without proper motivation, respondents may ignore instruc read questions carelessly, or provide incomplete answers. Worse yet, the skip questions altogether or fail to complete and return the questionna

In some instances motivational problems stem from poor question d such as when questions are difficult to read and understand, instructio hard to find, or the response task is too vague. In other instances, the qu topic itself may be the source of motivational problems. This is often th with questions pertaining to personal financial information. For exa people are more likely to report their income when provided with categories from which to choose rather than asked to provide an exact however, sometimes a survey, such as the U.S. Decennial Census, requi exact number, and anything else is unacceptable.

Respondents are also often reluctant to answer questions about c behaviors that they may find embarrassing or threatening, such as sex criminal activity. When asking for sensitive information about people or current behavior, changing the wording of the question can enco reluctant respondents to answer. For example, instead of asking "Hav ever shoplifted anything from a store?" one might ask "Have you ever anything from a store without paying for it?" Another strategy is to in the question with others, such as "How often do you go shopping? "What types of stores do you shop at?" so that it may appear in conte seem less objectionable. Although steps can be taken to improve the of sensitive questions, it may still be difficult to collect accurate inform from all respondents.

4. WHAT TYPE OF INFORMATION IS THE QUESTION ASKING FOR?

It is easier to get accurate answers for some types of survey question others. For example, almost everyone knows how old they are, as p are frequently asked to give that information to others. Because peo ready have an answer in their head when asked about their age, ass they are willing they can easily provide it and can do so in a num different ways, as shown in Figure 4.2. Because the information is r

Figure 4.2 What type of information is the question asking for?

A question that people can easily answer regardless of how it is asked

What year were you born?

☐ Year born

How old are you?

☐ Age

What is your date of birth?

☐ Month ☐ Day ☐ Year

A question for which people are often more likely to be influenced by context

In your opinion, how effective do you think citizens groups are in helping to solve environmental problems?

- ☐ Very effective
- ☐ Somewhat effective
- ☐ A little effective
- ☐ Not at all effective

In your opinion, how effective or ineffective do you think citizens groups are in helping to solve environmental problems? (Please mark an "X" on the line.)

Very effective ←——→ Very ineffective

In your opinion, how effective are citizens groups in helping to solve environmental problems?

- ☐ A great deal
- ☐ A fair amount
- ☐ Not very much
- ☐ Almost none at all

available to the respondent, most surveyors can accurately collect such data as age and other *factual or demographic information* regardless of how they ask for it; however, as we discuss later in this chapter, people can still be encouraged, through question design, to report the answer in a particular format.

Surveys also frequently ask for information that people may have thought little about and will need more time to answer, such as questions about *attitudes and opinions*. The second example in Figure 4.2 provides an example in which respondents are asked how effective they think citizens groups are in helping to solve environmental problems. Most respondents will not

considerable work to formulate one. Some may consider generally wh they think people can effect change, whereas others will think about exar of environmental citizens groups or various types of environmental prob that people have tried to help solve. For these types of questions, mo than for factual or demographic questions, respondents can be substan influenced by the context of the question as they work their way throug question–answer process of comprehending the question, recalling rel information, forming a judgment, and reporting their answer (Touran 1992). Different elements of the question that can influence the answ process include what type of response is being asked for (e.g., choos category, mark an X on the scale), the wording of the question and resp options (e.g., "a great deal" or "somewhat effective"), and visual layou

Other questions that are prone to such context effects are those a about *behaviors and events*. Surveyors often ask about many aspects of pe behavior, such as what they have done, how often (number of times or re frequency), and when. Frequently survey designers want responde provide far more detail about past behaviors than can be recalled ar a result, they write questions respondents find difficult, if not impos to answer. Doing this causes respondents to draw even more on fea of the questions' context rather than their real experiences in formul their answers, as is the case in Figure 4.1. To avoid this tendency, surv should consider three recall problems. First, memory tends to fade over Second, individual episodes or occurrences of regular and mundane e are generally not precisely remembered (Rockwood, Sangster, & Dil 1997). And third, people usually do not categorize information by p month or year. Given these limitations, respondents are unlikely to b to accurately report how many days they drove more than 5 miles d the past 6 months. But they can probably very accurately report how days they drove their car during the past week or drove more tha miles at a time in the past 3 months. Asking questions about behavior people can easily recall because they are recent or more memorabl help improve the accuracy of the information people report. In addit choosing an appropriate reference period for the type of behavior, definitions and examples can also help improve recall (Schaeffer & Pr 2003). However, definitions must be easy for respondents to understan examples should be selected carefully so as not to influence respon answers in unintended ways.

A HOLISTIC APPROACH

Throughout this and the next chapter, we discuss a holistic approa crafting effective survey questions for mail and Internet surveys. By h we mean that multiple aspects of the wording and design of the questio

to work together to convey meaning. This approach considers what type of question structure best measures the concept of interest, how questions are composed of multiple parts that work together, and how both the words and the visual presentation of questions are important.

CHOOSE THE APPROPRIATE QUESTION FORMAT

There are two broad types of question formats: open-ended and closed-ended questions. *Open-ended* question formats provide a blank space or box where respondents type or write in their response using their own words (or numbers), whereas *closed-ended* question formats or *scalar* questions provide respondents with a list of answer choices from which they must choose to answer the question.

The strength of the open-ended question format is that it allows respondents to freely answer the question as they want without limiting their response. Thus, this format is preferable when the surveyor does not want to influence respondent answers by providing a set of answer choices; when the goal is to collect rich, detailed information from respondents; and when the surveyor is questioning about topics for which little information is known ahead of time. Additionally, an open-ended format in which respondents provide a numerical response can sometimes be easier for respondents and yield more precise information because respondents report an exact number rather than choose from categories with vague labels or ranges of values.

However, there are also several limitations to open-ended question formats. In self-administered surveys, more respondents skip open-ended question formats than closed-ended formats because the former require more work to answer. Issues of item nonresponse bias arise because some types of respondents may be more likely to skip these questions than other types. If respondents do answer the question, they may provide only a short response. In addition, responses to open-ended questions must be entered and coded before they can be analyzed; however, web surveys make this less time consuming because the responses are already in electronic form. In addition to the time to code responses to open-ended questions, there is often a lot more variation in respondents' answers, so it may be more difficult to analyze and interpret the data, and a variable may not be able to be created based on the responses. In contrast, responses to closed-ended questions can be analyzed immediately (or with minor transformations to the data), and data results can be produced quickly.

Closed-ended question formats should be used when surveyors want respondents to provide an answer after considering or evaluating a set of answer choices. Because researchers provide answer categories in closed-

on how respondents interpret the questions. Closed-ended question fo
can utilize nominal or ordinal scales. In *nominal* scalar questions, respo
are asked to compare a set of categories with no natural order unde
the categories. Because the categories lack an inherent ordered relatio
the difficulty of processing nominal scales increases as the number o
gories that need to be compared at one time increases. An adaptation
nominal closed-ended format allows respondents to select multiple an
(e.g., check-all-that-apply and ranking questions). Examples of nomina
ables include grocery stores one has visited, web sites one frequently
and brands of personal care products purchased. Surveyors may orde
variables alphabetically or group them by type, but any such order
grouping is subjective and is usually to help make answering the qu
easier for respondents. One difficulty discussed below is that such
ing and grouping can sometimes have unintentional effects on answe
actually make responding more difficult.

In contrast, *ordinal* scale questions provide an ordered set of answ
egories (but the intervals between categories is unknown), and respo
must decide where they fit along the continuum. Because there is an in
order to the categories, respondents are particularly influenced by ho
categories are distributed and by the overall layout of the response
A common ordinal scale asks about levels of satisfaction (e.g., comp
satisfied, very satisfied, somewhat satisfied, not at all satisfied), wher
category represents a higher degree or level of satisfaction. Someone
"completely satisfied" is more satisfied than someone who is "very satis
but it is not necessarily known how much more satisfied. Another co
type of ordered scale asks about frequency of behaviors or events (e
of the time, most of the time, some of the time, none of the time),
each category represents a greater or lesser frequency. One concern w
dered closed-ended questions is that researchers often use vague quan
to describe the answer categories. For example, some people may co
walking every day to be "all of the time," whereas other people may co
walking three times a week to be "all of the time."

One of the fundamental writing tools that exists for creating s
questions is to shift questions from one format to another. Having a wo
knowledge of different question formats can help surveyors craft eff
survey questions because often questions that do not work in one qu
format can be converted to another format to more effectively measu
concept. For example, one of us was once asked to help a university
mittee that was preparing a questionnaire to evaluate a dean's perform
All of the questions proposed by the committee were nominal closed-e
questions with unordered response categories, similar to the qu
structure commonly used in student examinations. The first question, s

Figure 4.3 Choosing the appropriate question format.

Nominal closed-ended

Which of these five statements best describes the dean?

- ☐ Innovative but lacking leadership qualities
- ☐ About the same on innovation and leadership qualities
- ☐ Stronger on leadership than innovation
- ☐ A born leader
- ☐ A real innovator

Ordinal closed-ended for each concept

To what extent has the dean demonstrated strong leadership qualities?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ None of the time

To what extent has the dean demonstrated an ability to innovate?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ None of the time

Nominal closed-ended—revised to achieve direct comparison of concepts

Which one of the following do you feel best describes the dean?

- ☐ A strong leader
- ☐ A strong innovator
- ☐ Both a strong leader and innovator
- ☐ Neither a strong leader nor innovator

Open-ended for each concept

How would you describe the dean's leadership abilities?

How would you describe the dean's ability to innovate?

would have made it difficult to interpret the results. The proposed solution to the university committee was to break the question apart and to ask two ordered closed-ended questions that focused on how often or to what

and innovate are directly compared. Separating leadership and innov into separate questions and asking the direct comparison question all the committee to test its stated objectives of finding out how the fa evaluated the dean separately on leadership and innovation and on v attribute she performed better.

Another option for revising the question would have been to replace it two open-ended questions that asked separately about the dean's abil lead and innovate and perhaps still follow up with the direct compa of leadership and innovation in the nominal closed-ended question. W strategy should be used depends on the ultimate purpose for askin question. The open-ended questions would have produced more descri data on how faculty evaluated the dean's abilities independently; how the ordered closed-ended questions would have allowed the committ measure the dean's abilities using a common scale so that results coul easily summarized and compared.

A partially closed question format is a hybrid of the open- and clo ended formats that includes a set of response categories and an "other sponse, thus allowing respondents who do not fit into the provided resp categories to specify a different category that they do fit. This format is c used when it would be too burdensome to ask respondents about the e set of items. For example, in the question in Figure 4.4, if a respond favorite sport were basketball, she would select that sport and move to next question; however, if her favorite sport were rowing, which is not on the options provided, she would select "other" and write or type "row in the adjacent answer space. The value of this question format is th reduces the number of items respondents have to consider at once and collects data for the key items of interest. However, respondents are n likely to select the options provided than to write or type their own o responses. Hence, categories should be included for all of the key item interest, and care should be used in drawing conclusions about voluntee categories versus those that are explicitly provided (i.e., it is not accurat

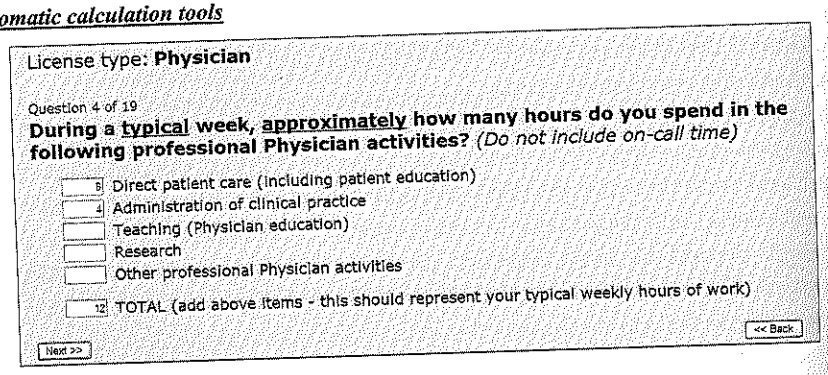
Figure 4.4 Example of a partially closed question format.

Which of the following is your favorite college women's sport?

- ☐ Basketball
- ☐ Gymnastics
- ☐ Soccer
- ☐ Softball
- ☐ Swimming
- ☐ Tennis
- ☐ Volleyball
- ☐ Other: Please specify

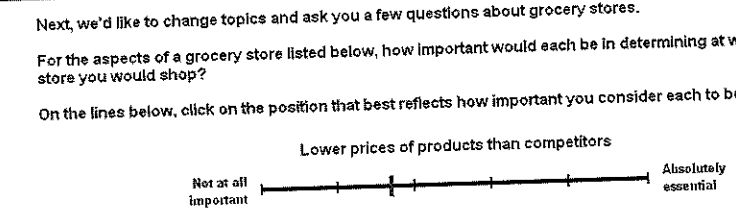
Figure 4.5 Example of new web response mechanisms.

Automatic calculation tools



Drop-down menus

Visual analog scales



Source: Visual analog scale image courtesy of "A Comparison of Visual Analog and Graphic Ratings Scales," by R. K. Thomas and M. P. Couper, March 2007, paper presented at the General Online Research Conference, Leipzig, Germany.

draw conclusions that compare listed to unlisted options, such as "respondents were 10 times as likely to say basketball is their favorite sport than rowing").

In web surveying, new response mechanisms such as automatic calculation tools, drop-down menus, and visual analog scales have been developed (see Figure 4.5). Each of these formats uses one of the question formats we discussed above, but the way that respondents report their answers differs in each case. *Automatic calculation tools* are usually composed of a series of numerical response open-ended questions for which the computer helps calculate a running total to make responding easier and reduce the number of

A *drop-down menu* provides respondents with a list of options whe select one or sometimes multiple responses, similar to a nominal ended question or a multiple-answer question. The main difference respondents cannot view the options until they click on the menu, an they often have to scroll to find the answer they want to select. Heerwe Loosveldt (2002a) and Couper, Tourangeau, Conrad, and Crawford found that response times were not significantly different between a button format and drop-down box. However, Couper et al. (2004) also that respondents were more likely to select the visible response options provided with a drop-down box with half of the response options sh than when presented with a drop-down box where they had to scroll the response options. Thus, it is important to not preset drop-down with only some response options visible. Instead, one should label the with what type of information is being requested and only make s items visible once the box is clicked, such as in the example in Figure

Visual analog scales usually use an ordinal closed-ended question f in which the respondent can interactively slide a marker to the po on the scale that best describes his or her answer. The parallel in a survey would be placing a mark such as an X on a continuum. In two studies of visual analog scales on the Web, Couper, Tourangeau, Conra Singer (2006) found no differences in the response distributions for analog scales than for scales presented horizontally with radio button Thomas and Couper (2007) found similar validity and self-reported acc ratings between visual analog scales and scales presented as a list of o vertically with radio buttons. Both studies found that the visual analog took longer to complete than the scales with radio buttons.

Although these new response mechanisms provide alternative wa respondents to report their answer or help them provide an answer, of them are not new question formats, so we discuss them with the sp question structure to which they apply.

THE ANATOMY OF A SURVEY QUESTION

Survey questions are made up of multiple parts that must work togeth concert to produce high-quality data about the topic of interest. If one of the question fails or provides a conflicting message with another p can undermine the accuracy of responses. Crafting good survey ques requires understanding how each component of the question conveys m ing independently to respondents as well as how all of the parts work tog to convey meaning.

The most important part of any survey question is the *question* or the words that form the actual query itself. In Figure 4.6, the que

Figure 4.6 Examples of the components of open- and closed-ended question formats.

Open-ended question

Question stem with additional verbal and numeric instructions

How many years have you lived in Washington?
Please report only whole numbers. (For example, if you have lived in Washington 20 months, please round to 2 years.)

Answer space with additional verbal and symbolic instruction

☐ # of years

Closed-ended question

Question stem with additional verbal instruction

How many years have you lived in Washington?
Please choose the category that best describes the total number of years you have lived in Washington during your lifetime.

Answer choices

○ Less than 2 years
○ 2 to 5 years
○ More than 5 years, but less than 10 years
○ 10 or more years

stem provides the most explicit and direct information about what the question is asking (e.g., how long have you lived in Washington) and how respondents should provide their answer (e.g., in years). It may also include *additional instructions*, definitions, or examples that will help respondents comprehend the meaning of the question or of key concepts. Additional instructions might include verbal instructions (e.g., "select only one" or "please round to the nearest whole year"), or they may consist of numbers, graphics, or symbols that further inform respondents how to answer (e.g., $, #, YYYY). Each question has *answer spaces or choices* that provide additional information to respondents about what responses are possible and how to record their answer. Answer choices limit the available possibilities from which respondents can choose (e.g., less than 2 years, 2–5 years, etc.) and often provide cues about the type and number of answers to provide (e.g., size of answer space, reminder of units requested). Throughout this chapter and the next, we use the specific terms italicized above for each of the particular parts, and we use the term *question* to mean the entire anatomy of the survey question, including all of the parts.

Words used in the question stem, additional instructions, and response options are the primary sources of meaning that respondents draw upon when comprehending the meaning of survey questions. In addition, numbers, symbols, and graphics can also influence how respondents answer questions. Finally, the visual layout and presentation of questions can also have significant

question stem, any instructions that accompany the stem or answer and the visual display of the answer spaces (e.g., size, location, etc.). In ended questions, both the wording and the visual display of the choices, in addition to the question stem and any instructions provi influence how people respond. Therefore, crafting survey questions i both choosing words to form the questions and deciding how to present the questions, including each of the component parts, to respo

In the remainder of this chapter, we present guidelines for crafting questions that apply to nearly all types of survey questions. First we guidelines for choosing words and writing good survey question we discuss further how the visual layout of survey questions can in responses to questions in mail and web surveys. Lastly, we present guidelines for the visual presentation of survey questions that apply t all types of survey questions.

GUIDELINES FOR CHOOSING WORDS AND FORMING QUESTIONS

Guideline 4.1: Make sure the question applies to the respondent

Imagine you are responding to a survey that asks "What type of connection do you have in your home?" but you do not have an connection. How do you respond? What if you ate dinner at a res last night but are asked "If you made dinner at home last night, abo many minutes did it take to prepare the meal?" (see Figure 4.7). Son people who write questions for mail surveys try to reduce the nu

Figure 4.7 Make every question require an answer.

A question that does not require an answer from every respondent

If you made dinner at home last night, about how many minutes did it take to prepare the meal?

☐ minutes

A revision that uses a filter question

Did you make dinner at home last night?

☐ Yes
☐ No

If yes, how many minutes did it take to prepare the meal?

☐ minutes

questions (i.e., save space) and avoid skip instructions by asking questions such as these.

These two questions have a common problem (i.e., they do not require answers of every respondent), but the source of the problem is slightly different for each. The first question contains an embedded assumption of having an Internet connection that may not be true for all respondents. The second avoids making assumptions by including an "if" statement but does not apply to those who do not fit the "if" criteria. Even if a "does not apply" box were provided, the use of the word *if* implies that no response is needed from those who ate out the previous night.

This type of question can be particularly damaging in web surveys when respondents are required to enter an answer for every question before they are allowed to advance to the next question. In this situation, the respondent has to choose between two bad options: knowingly entering false information or quitting the survey altogether. However, questions that do not apply to every respondent are still problematic in questionnaires that do not require answers. Aside from potentially confusing respondents, the methodological problem with these questions is that it is impossible to distinguish between those who did not respond because they were unmotivated (i.e., nonrespondents) and more motivated people who, nonetheless, did not respond because the question did not apply to them. Moreover, in order to be able to estimate the distribution of a characteristic in the sample population, the surveyor must give respondents the opportunity to answer every question they are asked. Therefore, a good rule to apply is that in order for an inquiry to constitute a survey question, it must require an answer from each person of whom it is asked. The two questions above should only be asked of respondents who answer "yes" to filter questions such as "Do you have an Internet connection in your home?" or "Did you make dinner at home last night?" (see Figure 4.7). Techniques to ensure that respondents follow the required skip instructions properly are discussed in Chapter 6.

Guideline 4.2: Make sure the question is technically accurate

Asking a question that is not technically accurate can confuse respondents and make answering difficult. For example, many avid horse people might be confused by this seemingly simple question: "How many feet tall is your horse?" This is because horses are often measured in a different unit—hands (one hand equals 4 inches). A more appropriate way to ask this question would be "How many hands tall is your horse?" Ensuring that questions are technically accurate becomes more challenging when one is asking questions about topics that apply to very specialized populations such as equestrians or business executives. Failing to do so, however, can compromise the quality

the surveyor, possibly resulting in reduced motivation or even break-o the part of respondents.

Guideline 4.3: Ask one question at a time

On first take, the advice to ask one question at a time seems like a no-br yet it is striking how often what appears to be one question actually cor two components about which respondents may feel differently. Con for example, the question in Figure 4.8. This question actually contains questions: "Do you subscribe to any periodicals, magazines, newsle etc. . . . ?" and "Do you regularly read any periodicals, magazines, new ters, etc. . . . ?" As this question is written, it poses problems for respond who subscribe to but do not regularly read these items, or for those regularly read them but do not subscribe. Such respondents will not k for which component to provide their response. The question also p a problem for the survey or anyone using the resulting data, as they not know which component the respondents were referring to when marked "yes" or "no."

One possible solution to this problem, as demonstrated in the first revi is to untangle the original question by asking the two questions separa If even more precision is desired, each of the two questions can be writte a forced-choice style so that respondents are asked to indicate whether subscribe to and whether they read each type of literature, as shown ir second revision in Figure 4.8.

Guideline 4.4: Use simple and familiar words

One way to establish legitimacy and credibility with respondents is to pre them with a formalized and professional questionnaire. Generally thi good practice. One way it can backfire, however, is if efforts to forma the questionnaire lead to the use of complex words or phrases and techn terminology that not all respondents will understand. Many complex wo and phrases can be easily replaced by more generally understood terms seen in Figure 4.9. When drafting questions it may be advisable to consu grammar book or writing manual, as they commonly provide more extens lists of replacement terms for complex words and wordy phrases. A good r of thumb is that when a word exceeds six or seven letters, a shorter and m easily understood word can probably be substituted. However, it should automatically be assumed that all shorter words are acceptable. For exam it would not be advisable to substitute "deter" for "discourage."

Another common tendency, especially in government surveys, is to in vertently use abbreviations or specialized phrases that are commonplace the survey sponsor but require some translation for respondents. An exa ple is shown in Figure 4.10. Although the survey sponsors may know w

Figure 4.8 Ask one question at a time.

A double-barreled question

Do you personally subscribe to, or regularly read, any periodicals, magazines, newsletters, etc. that are specifically related to your occupation?

☐ Yes
☐ No

A revision to ask each question separately

Do you personally subscribe to any periodicals, magazines, newsletters, etc. that are specifically related to your occupation?

☐ Yes
☐ No

Do you regularly read any periodicals, magazines, newsletters, etc. that are specifically related to your occupation?

☐ Yes
☐ No

Another possible revision to collect more specific information

Please indicate whether or not you personally subscribe to each of the following sources of information specifically related to your occupation.

	I do subscribe	I do not subscribe
Periodicals	☐	☐
Magazines	☐	☐
Newsletters	☐	☐

Please indicate whether or not you regularly read each of the following sources of information specifically related to your occupation.

	I regularly read	I do not regularly read
Periodicals	☐	☐
Magazines	☐	☐
Newsletters	☐	☐

what IRS stands for, further confusing them and making it even harder to figure out what form SS-4 is. A clearer statement of the question is provided in the revision, in which the form is referred to by its full name and "IRS" is replaced with "Internal Revenue Service."

In most instances it is desirable to replace complex and specialized words, but there are instances when this is not necessary. Virtually all occupational groups share a particular vocabulary that is not understood by outsiders but that facilitates efficient communication within the group. Replacing this specialized vocabulary with simpler words would only confuse matters for these

Figure 4.9 Words and phrases that can be simplified.

Replacing complex with simple words

Exhausted	Tired
Candid	Honest
Top priority	Most important
Leisure	Free time
Employment	Work
Courageous	Brave
Rectify	Correct

Replacing complex with simple phrases

Occupants of this household	People who live here
Your responses	Your answers
Post-school extracurricular activities	What you do after school
Subnational region	Area of the country

to talk about "annexation" instead of "an addition." Similarly, in a physicians it seems reasonable to talk about "pharmaceutical comp stead of "companies that sell medicines." To do otherwise may eve a lack of knowledge and understanding of the topic of the survey.

However, the fact remains that people who write question far more likely to overestimate than underestimate the knowl vocabulary of respondents. Thus, when in doubt, it is prudent t simpler of the available alternatives. Ultimately, though, the best w termine if the vocabulary is appropriate is to pretest questions with of the population of interest to identify potential difficulties, a topic to in Chapter 6.

Figure 4.10 Use of simple and familiar versus complex or specia words.

Unnecessary use of specialized words and abbreviations

Have you filed Form SS-4 with the IRS?

☐ Yes
☐ No

A simplified revision

Have you filed an application for an employer identification number (Form SS- with the Internal Revenue Service?

☐ Yes
☐ No

Guideline 4.5: Use specific and concrete words to specify the concepts clearly

Suppose you are interested in different elements of family cohesiveness and you pose the question in Figure 4.11 to a sample of mothers with school-age children. One strength of this question is that it clearly specifies a reasonable time referent and the units in which one should report an answer. The problem, however, is that it contains several vague concepts. What does "eat" mean? Should we count snacks? What if we all gathered for a smoothie a couple hours before dinner; does that count, even though it is a thick drink? And what does "together as a family" mean? Does the lunch we got from the drive-thru and ate in the car on the way to grandma's house count? What about the pizza we ate in the living room while watching a movie on Saturday night? Do we count the end-of-season potluck for the baseball team at which we were all present but interspersed with other families?

One respondent might take a very liberal interpretation of this question and include all gatherings where her whole family was present and food or drinks were consumed, whereas another may take a very conservative view and count only full meals for which the family was gathered at home around the kitchen table. The problem is that the two interpretations would most likely result in quite divergent answers that are due to different interpretations of the question, not differences in family cohesiveness. This question could be improved by specifying that you are interested in meals consumed at home, as in the first revision. It could be made even more specific by specifying that you are interested in sit-down meals shared as a family, as in the second revision.

Figure 4.11 Use specific and concrete words to specify the concepts clearly.

Question with vague concepts

How many times did you eat together as a family last week?

☐ # of times

A revised question with more specific and concrete concepts

How many meals did you eat together as a family at home last week?

☐ # of meals

A more specific revision

How many meals did you sit down to eat at home as a family last week?

☐ # of meals

This example illustrates a common problem for many writers of s questions. Once the units are specified, many concepts such as age, h and weight are very straightforward. Others, however, are not as str forward as they seem. It seems like everyone would know what it mea eat together as a family, but as the example shows, once one begins to in the complexity of family life in this day and age, the concept of e together as a family is opened up to much interpretation. Thus, it is in tant to make sure the concepts in survey questions are clearly defined communicated in order to minimize the amount of interpreting and def that respondents need to do.

Guideline 4.6: Use as few words as possible to pose the question

Part of keeping questions simple is keeping them short and to the point longer the question, the more information the respondent has to take in process, and the higher the likelihood for misunderstanding or misrea When presented with the question in Figure 4.12, for example, one respon in a cognitive interview answered, "I don't have any idea how many pe live in the United States." As a result of this and other interviews, the intentioned second sentence that explained the reason for the directions a who to include and who to exclude was removed (Dillman & Allen, 199

The goal of keeping questions short sometimes contradicts the previo stated goals of using familiar and simple words and using specific concrete words to specify concepts clearly. Substituting several sim words for a more complex word or carefully specifying concepts lengthen questions. In these instances, we recommend subordinating

Figure 4.12 Use as few words as possible to pose the question.

Long question with potentially confusing information

How many people were living or staying at this residence on Saturday, March 3rd, 2000? To make sure each person in the United States is counted only once, it is very important to:

Include everyone who lives here whether related to you or not, and anyone staying temporarily who has no permanent place to live;
But not include anyone away at college, away in the Armed Forces, in a nursing home, hospice, mental hospital, correctional facility, or other institution.

A shorter revision with potentially confusing information removed

How many people were livingor staying at this residence on Saturday, March 3rd, 2000? Please:

Include everyone who lives here whether related to you or not, and anyone staying temporarily who has no permanent place to live;
But not include anyone away at college, away in the Armed Forces, in a nursing home, hospice, mental hospital, correctional facility, or other institution.

Figure 4.13 Eliminating wordy and redundant expressions.

Due to the fact that	Because
At this point in time	Now
A small number of	A few
A considerable number of	Many
Small in size	Small
Has the ability	Can
Ascertain the location of	Locate
Concerning the matter of	About
If conditions are such that	If
In the majority of instances	Usually
Make a decision	Decide
Take into consideration	Consider

goal of keeping the question short to the goals of using simple and familiar words and using specific and concrete words. Once one is sure that any words chosen are understood by virtually all respondents and the concepts are clearly specified, one can attempt to keep the question short.

There are several ways to do this. One way is to replace wordy and redundant expressions such as those shown in Figure 4.13 with simpler wording. More comprehensive lists of commonly used wordy expressions and their replacements can be found in most grammar and writing manuals.

Another technique is to avoid including answer categories in the question stem, as in the following example: "Are you very likely, somewhat likely, somewhat unlikely, or very unlikely to visit Glacier National Park again?" If the respondent will be able to see the response options after reading the question stem, it is unnecessarily redundant to include these options in the stem as well. Such redundancy across many questions is a particularly strong indicator to respondents that it is okay to skip words, and it may result in the rest of the sentence also being unevenly read. Therefore, this question should be shortened to "How likely or unlikely are you to visit Glacier National Park again?" However, it is important that this technique only be used in single-mode self-administered survey designs in which respondents will be reading the question for themselves. In surveys where an interviewer administers the survey or when self-administered surveys are combined with interviews in mixed-mode studies, the answer categories may need to be included with the question stem for better respondent comprehension and consistency across modes (see also Chapter 8).

Guideline 4.7: Use complete sentences with simple sentence structures

It is tempting to save space by using incomplete sentences for paper surveys.

Figure 4.14 Use complete sentences.

Common use of incomplete sentences

Number of years lived in Idaho

☐ Years

Your city or town

☐ City or Town

Your County

☐ County

A revision using complete sentences

How many years have you lived in Idaho?

☐ Years

In what city or town do you live?

☐ Name of City or Town

In what Idaho county do you live?

☐ Name of Idaho County

However, the series of questions in Figure 4.14 once caused many respondents to provide erroneous answers to the second and third questions. Nearly 20% of the respondents listed the number of years they had lived in the city or town and the county. In addition, several other respondents listed "United States" for *county*, a word that is only one letter different from *country*. Writing each question as a complete sentence would have helped solve both problems. In addition, in the revision "county" is changed to "Idaho county" in order to minimize the possibility of listing the United States as the respondent's county of residence.

Using complete sentences is even more important in web surveys with page-by-page construction. When there is only one question per screen, respondents can easily lose track of the context in which an inquiry is made. Here, incomplete sentences become isolated in ways that can make the meaning even less clear.

When asking for very specific information it is tempting to add extra clauses onto the sentence to help specify the focus of the question; however, these may confuse respondents or result in their misunderstanding the

question. The problem with doing this is that reading sentences with multiple clauses or a complex sentence structure requires more skill than reading sentences with simple structures. Just as with wording, it is advisable to avoid complex sentence structures and replace them instead with simple structures.

Guideline 4.8: Make sure "yes" means yes and "no" means no

It seems obvious that questions should not include double negatives, or, in other words, require a respondent to say "yes" to mean "no" as in the following question: "Should the city manager not be directly responsible to the mayor?" Yet such questions are commonly asked in surveys. One of the reasons they are so prevalent is because voters are often asked in elections to vote for measures where a yes vote would result in something not being done, as illustrated by the tax approval question in Figure 4.15. Surveyors are often reluctant to pose the question differently than it will be expressed on the ballot. However, because people tend to read questions quickly, it is likely that some people will miss the word "not." In addition, the mental connection of favoring a "not" is difficult for most people.

Two different solutions for this problem might be considered. The first revision simply asks whether people favor or oppose requiring 60% approval by voters in order to raise state taxes. To help clarify, the answer categories specify what favor and oppose mean for the purposes of this question. This

Figure 4.15 Make sure "yes" means yes and "no" means no.

A question containing a double negative

Do you favor or oppose not allowing the state to raise taxes without approval of 60% of the voters?

☐ Favor
☐ Oppose

A revision with no double negative

Do you favor or oppose requiring 60% approval by voters in order to raise state taxes?

☐ Favor
☐ Oppose

A revision that preserves important wording

In the September election, you will be asked to vote on this referendum: "No state tax can be raised without approval by 60% of those voting in a statewide election." If the election were held today, would you vote for or against approval?

☐ For
☐ Against

wording would seem appropriate during discussion of an issue before it h reached the ballot measure stage. A second revision, indicating that a vo will be taken, specifies the measure exactly as it will appear on the ball and asks whether respondents are for or against approval of the measu The switch of categories from favor/oppose to for/against is also an attem to bring the language of the question more in line with that of the voti situation.

Guideline 4.9: Be sure the question specifies the response task

Another way of stating this guideline is that the response task should be cle to the respondent after having only read the question stem. There are tw pieces to this guideline. First, *the question stem has to clearly state the respon task.* This means that the question stem needs to ask for exactly the kind information and level of detail (e.g., units) the surveyor wishes to collect. Se ond, *the response format and/or options provided must match the task as it is stat in the question stem.* In other words, do not change the rules in the middle the game. A yes/no question should only have "yes" and "no" as substanti response options (nonsubstantive options such as "don't know" or "not ap plicable" are still appropriate). Similarly, if the question asks how many da something occurred, the response options should be numbers appropriate t the time referent (e.g., 0–7 if the referent is a week) or a number box labele "number of days." Asking for the number of days in the question stem an then providing options such as "Always," "Most of the time," "Sometimes, "Rarely," and "Never" represents a mismatch between the question stem an the response options that forces the respondent to undertake the extra, an sometimes difficult, step of determining which category fits best (i.e., is 5 day "most of the time" or "sometimes"?).

HOW DOES THE VISUAL PRESENTATION OF SURVEY QUESTIONS MAKE A DIFFERENCE?

Prior to the 1990s, most of the survey methodology research focused on how the wording of questions influences how respondents answer them; the visua design and layout of questions was mostly viewed as an art form. Since then researchers have drawn on research from the science of visual perceptio (Hoffman, 2004; Palmer, 1999; Ware, 2004) to formulate theoretically based rationales for how the design of questions and questionnaires can be im proved to help respondents process survey questions and navigate through self-administered questionnaires (Christian, Dillman, & Smyth, 2007a; Jenkins & Dillman, 1997; Redline & Dillman, 2002; Redline, Dillman, Dajani, & Scaggs, 2003). In addition, a growing body of experimental research has

various aspects of the visual design and layout of survey questions influence how people respond to paper and web surveys (Christian & Dillman, 2004; Dillman & Christian, 2005; Smyth, Dillman, & Christian, 2007b; Smyth, Dillman, Christian, & McBride, in press; Smyth, Dillman, Christian, & Stern, 2006b; Tourangeau, Couper, & Conrad, 2004, 2007).

Visual design and layout influences respondents as they organize information presented in the survey questionnaire and as they focus their attention on responding to individual survey questions. Visual design features can help guide respondents to self-administered surveys, much like an interviewer would do in a face-to-face or telephone survey, to ensure that each respondent receives the same question stimulus delivered in the same way. Thus, this research is requiring survey methodologists to apply new visual design concepts to the practice of crafting survey questions and constructing questionnaires.

In this chapter, we briefly introduce the visual design concepts, summarized in Figure 4.16, necessary for understanding how survey respondents process and respond to survey questions. We follow this discussion with guidelines that can be applied to the visual presentation of all types of survey questions to help ensure that respondents attend to each component and process the parts of the question in the intended order. In Chapter 5, we present guidelines for the wording and visual presentation of specific types of open- and closed-ended questions. We continue our focus on visual design in Chapter 6, where we describe how respondents visually process and navigate entire questionnaire pages. There we present guidelines for questionnaire construction to help respondents organize the information presented on each questionnaire page and to encourage them to process questions in the intended order.

Survey questionnaires include four types of visual design *elements* that communicate meaning to respondents: words, numbers, symbols, and graphics. *Words* are the most powerful source of meaning that respondents draw upon when answering survey questions; however, *numbers* used in the question stem, in instructions, and in or as labels for answer choices also communicate additional meaning to respondents. *Symbols* can also be used to add special meaning, often without occupying very much physical space. For example, an arrow may communicate to respondents where to focus their attention next. Finally, *graphics* (e.g., text boxes, squares, html boxes, circles, and radio buttons) are another type of visual design element that can be used in designing survey questions. In addition, graphics can include more complex images and logos that layer various elements. The use of these types of graphics has increased, particularly in web surveys, because of the ease and affordability of including them, but also in mail surveys because the cost of printing them has decreased.

Figure 4.16 Visual design concepts that guide question design.

Visual design elements that communicate information to respondents

Words are the fundamental source of meaning that help respondents understand what is being asked of them.

Numbers are used to convey meaning and sequence or order to respondents.

Symbols are figures that add special meaning based on what they represent to respondents.

Graphics are shapes and visual images that can be simple or complex and convey meaning to respondents.

Visual design properties that modify how elements are presented visually and the meaning respondents assign to them

Size: Changes in the size of elements influence how elements are perceived and whether they stand out visually.

Font: Changes in the shape and form of elements influence the legibility of words and how elements are perceived.

Brightness/Contrast/Color: Changes in shading and color influence how elements are perceived and whether they stand out visually from the background.

Location: How near or far elements are from one another (the spacing and alignment) influences whether they are perceived as related or unrelated.

Gestalt grouping principles that guide how respondents perceive relationships among information

Pragnanz: Elements that are organized into the simplest, most regular, symmetrical objects will be easier to perceive and remember.

Proximity: Placing visual elements closely together will cause them to be perceived as a group.

Similarity: Elements sharing the same visual properties (color, shape, size, orientation, etc.) will be grouped together.

Elemental connectedness: Elements connected by other elements will be grouped together.

Common region: Elements within a single closed region will be grouped together.

Continuity: Visual elements that can be seen as continuing smoothly will be perceived that way.

Closure: Elements that together create a "closed" figure will be perceived that way.

Common fate: Elements that move or imply movement in the same direction will be grouped together.

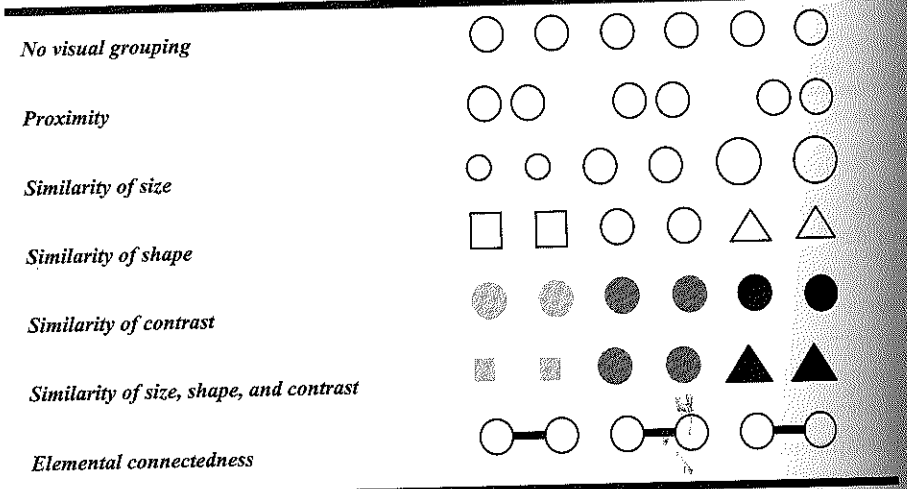
Each of the four visual design elements (words, numbers, symbols, a graphics) can be presented in different ways: They can be large or small, lig or **dark**, close to one another or far apart, static or in motion, grayscale in color. Thus, in addition to using the four types of visual design elemen surveyors can also manipulate the *properties* of each of these types of elemer to increase or decrease the attention and change the meaning responder assign to them. Visual design properties include size, font, brightness contrast, color, location or proximity, shape, orientation, and motion. F example, attention can be drawn to particular words, numbers, or symbols changing their size, contrast, or color in relation to the surrounding text (e.

bolding or *italicizing* an important word or phrase in the question stem). A dollar symbol ($) could flicker (motion) to remind respondents to report their income in dollars. Symbols can also be located in proximity to the answer spaces where respondents will need to use them at the time of response. In addition, differently shaped graphics for response scales, for example a ladder versus a pyramid shape, can influence how respondents interpret and respond to the scales (Schwarz, Grayson, & Knäuper, 1998).

When respondents are presented with a questionnaire, they first take in the entire scene and organize the information presented to them. During this process, respondents begin to distinguish the various visual elements on the page, and the properties of these elements influence whether they are noticed (Ware, 2004). After respondents organize the information on the page, they focus on answering the questions. Here, the visual processing is attentive and conscious, the visual field narrows to about 2 degrees or 8 to 10 characters in width (the *foveal* view), and attention is focused on only a few elements (for definitions of *attentive processing* and the *foveal view*, see Figure 6.4 in Chapter 6). During focused attention, the properties of the visual design elements or how they are displayed can strongly influence the meaning respondents assign to them.

The Gestalt psychology principles of pattern perception can help surveyors understand how respondents perceive groups among visual elements with shared properties and then assign meaning by viewing the grouped elements as conceptually related (see Figure 4.17). According to the principle of *proximity*, locating elements that should be grouped closer to each other

Figure 4.17 Examples of Gestalt grouping principles.

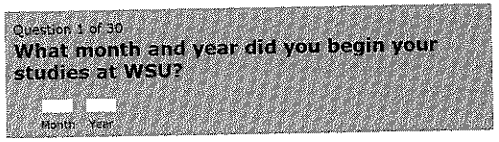
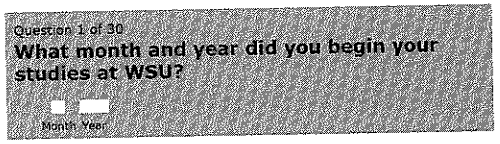
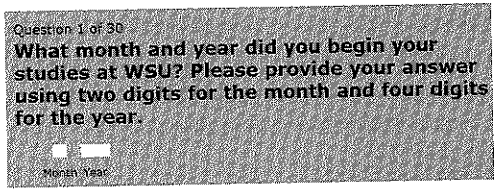
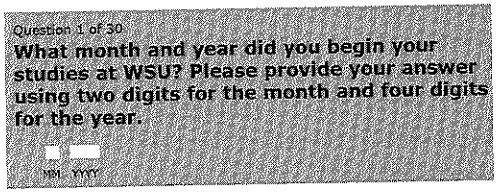


than to other elements encourages respondents to perceive them as rel
Surveyors can also use *similarity* of contrast, color, size, or shape to encou
respondents to perceive elements as a group. Respondents also per
elements that are enclosed in a *common region*, such as a box or an area w
common background color, as a group. Likewise, connecting visual elem
by using another element, such as a line, encourages respondents to per
the *connected* elements as a group. Finally, elements that continue smoo
will be perceived as a group; *continuity* is probably used in surveys m
in presenting complex graphics where multiple elements are layered
continuous manner so that they are perceived as a group. These Ge
principles can help surveyors in deciding how to present visual de
elements and which properties to apply to them. In addition, using tw
more of these principles to layer properties on the same element(s) can
a stronger stimulus to make them appear as a group than varying only
property of the elements.

The following example illustrates how visual design elements and p
erties used in accordance with the Gestalt grouping principles can imp
responses to a simple survey question. In a recently published article b
on a series of experiments, we found that visual design changes to the q
tion "What month and year did you begin your studies at WSU?" incre
the percentage of respondents reporting their answer in the desired fo
(i.e., two digits for the month and four digits for the year) from 55% to
an increase of 41 percentage points (Christian et al., 2007b). In a follow
study, Christian (2007) reported the sequential impact of a series of visual
verbal manipulations to the same question. The results of her study ca
seen in Figure 4.18. In the initial version at the top of Figure 4.18, respond
were provided with two equal size boxes for the month and year, and
44% provided their answer using the desired format. However, when
size of the month box was reduced by half (consistent with the expecta
that the month be reported in half the number of digits as the year), n
respondents reported the year using four digits, raising the percentage u
the desired format to 57%. Thus, the size of the answer boxes communic
additional meaning to respondents, beyond the graphics alone, about
many digits they should use in providing their answer.

In the next manipulation, adding the verbal instruction to "Please
vide your answer using two digits for the month and four digits for
year," resulted in a 21 percentage point increase, bringing the percen
using the desired format to 78% and demonstrating that respondents w
processing the words in the instruction and applying them when pro
ing their response. Providing the verbal instruction directly after the q
tion stem helped to ensure that respondents would see and process it
prior to providing their answer. Finally, replacing the word labels "Mo

Figure 4.18 Example of the influence of visual design on how respondents report date information.

Experimental Treatment	Used desired format (2-digit month, 4-digit year)
	44%
	57%
	78%
	94%

Source: *How Mixed-Mode Surveys Are Transforming Social Research: The Influence of Survey Mode on Measurement in Web and Telephone Surveys*, by L. M. Christian, 2007, Pullman, WA: Washington State University. Unpublished doctoral dissertation.

and "Year" with a symbolic instruction MM YYYY beneath their respective boxes increased the percentage using the desired format to 94%. The symbolic instruction was designed so that the letter *M* was used to represent month and *Y* to represent year, with the number of letters indicating the number of digits to use when reporting the month and year. In this version, respondents gained additional meaning from the symbolic instruction about the number of digits to use in their response. Locating the instruction near the answer spaces and within their foveal view also helped to ensure that respondents would notice and apply the instruction when reporting a

In the previous experiment reported by Christian et al. (2007b), pr
ing the symbolic instruction with the answer spaces had the highest im
on use of the desired format (increased use by 35 to 42 percentage p
in the two experiments), but there was no additional instruction prese
with the question stem. That the instruction was so important in Christ
(2007) experiment (see Figure 4.18) without the MM YYYY symbol to
cate the desired number of digits suggests that verbal instructions are
ticularly effective in the absence of adequate visual information. Resul
telephone experiments confirm this conclusion: The instruction to use th
sired format on the telephone, where there is no visual information avail
raised the percentage reporting in the desired format from less than 1
59% (Christian, 2007). Nevertheless, in self-administered surveys, it is
clear that the use of visual information is key to obtaining desired respo
and can contribute to this effort above and beyond question and instruc
wording.

GUIDELINES FOR THE VISUAL PRESENTATION OF SURVEY QUESTIONS

Because the visual presentation of survey questions influences how pe
answer them, choosing words and forming clear questions is not enou
surveyors also need to think about how to put all of the components of
question together. We now turn to describing general guidelines for the vi
presentation of survey questions that apply to designing nearly all type
questions. We continue to use the visual design concepts we just discus
in presenting these guidelines.

The poorly designed question in the top panel of Figure 4.19 would
doubtedly turn many respondents off. Some might not even be able to un
stand that this is a survey question or what it is asking, making it difficult
most respondents to provide an answer. For the next five guidelines, we
this question as an example. The overall problem with this question is th
is unorganized and cluttered, so there is no clear message sent by the vis
design. But such a broad observation does not necessarily give us enou
information to start revising the question. When we look closer, howe
we can identify a number of more specific problems that we can begi
address:

- It is difficult to tell where the question stem ends and the respo
  options begin.
- The response options run together.
- It is not immediately clear how one should mark an answer.
- Certain options stand out more than others, making them more likely

Figure 4.19 Implementing general visual design principles to construct individual questions.

Poor design

Which one of the following best **describes** the reason for your most *recent visit* to the Southgate Mall? ♡ **Shopping** for fun/entertainment ▷ Shopping for a needed item [ ]MALL WALKING/Exercise ☐ Other _____ ○Dining at the mall ☆Hanging out with friends Meeting new people ♧ Conducting Business

Revision with improved design

Which one of the following best describes the reason for your most recent visit to the Southgate Mall?

- ☐ Shopping for fun/entertainment
- ☐ Shopping for a needed item
- ☐ Mall walking/exercise
- ☐ Dining at the mall
- ☐ Hanging out with friends
- ☐ Meeting new people
- ☐ Conducting business
- ☐ Other _____

Poor design

How much do you favor or oppose implementing a merit-based pay system for elementary school teachers? [1] Very much in favor [2] Somewhat in favor [3] Neutral [4] Somewhat oppose [5] Very much oppose

Revision with improved design

How much do you favor or oppose implementing a merit-based pay system for elementary school teachers?

- ☐ Very much in favor
- ☐ Somewhat in favor
- ☐ Neutral
- ☐ Somewhat oppose
- ☐ Very much oppose

- The purpose of the bolding, underlining, and reverse print are unclear.
- One would have to process the "other" option before processing all of the options provided by the surveyor.

The revision is clearly an improvement on the original design. To construct

Guideline 4.10: Use Darker and/or Larger Print for the Question and Lighter and/or Smaller Print for Answer Choices and Answer Spaces

The first thing we needed to do was create subgrouping within the q tion. Good subgrouping helps the respondent quickly recognize and pro the parts of the question. We used the design property of *contrast* to cr separation between the question stem and the response options. To cr differences in contrast, we bolded the question stem but not the respo options. This is the standard use of bolding that we have adopted for n of the examples throughout this book. If we wanted to reserve bolding another purpose in our questionnaire, however, another property we co manipulate is text size (see the bottom example of Figure 4.19). Increas the size of the text in the question stem but not the response options h differentiate these two parts of the question.

Guideline 4.11: Use Spacing to Help Create Subgrouping within a Question

The Gestalt psychology principle of *proximity* states that items located cl to one another will be perceived as belonging to a group, and items loca farther apart will be perceived as not belonging together. We applied principle to help reinforce the subgrouping within the question. We sta by moving the first response option onto its own line of text and add some extra space between it and the question stem. We then moved e response option onto its own line and arranged them vertically underne the question stem so that they would no longer blend together. To help cre the impression that the response options were all part of one group, we pla them in close vertical proximity to one another and spaced them equally. also indented them a few spaces to the right underneath the question stem reinforce the subgrouping we were creating. Grouping and subgrouping multiple questions is discussed in more depth in Chapter 6.

Guideline 4.12: Visually Standardize All Answer Spaces or Response Options

Another problem with the poor design that we needed to address was t some response options stood out visually more than others, making th more likely to be seen and selected. "Shopping" stood out because it w bolded, "dining" because it was in reverse print, and "hanging out w friends" because it was underlined. Our solution to this problem was standardize the design properties of all of the response options. The fi thing we did was make sure they were all the same readable size with t same character spacing. We then changed them all to the same font.

We chose Times New Roman because of its readability and professionalis (i.e., compared to the Disney Script used for "conducting business"). Final we removed variations due to color, contrast (bolding), underlining, a

process and helps ensure that they will be processed equally. Incidentally, the similarity across response options also helps them appear as a subgroup within the larger question group (i.e., the Gestalt psychology principle of *similarity* says that items that appear regular and similar will be perceived as belonging together). In addition to making these changes, we reordered the response options so the "other" option was located at the end of the list so that respondents would process all of the response options before getting to it.

Guideline 4.13: Use Visual Design Properties to Emphasize Elements that Are Important to the Respondent and to De-Emphasize Those that Are Not

In the poor design, the words "describes" and "recent visit" in the question stem are emphasized with bolding and italics, respectively. However, these words seem no more important to the respondents' understanding of the response task or the question than any others in the stem. As a result, the bolding and italics were removed in the revised design. Instead of emphasizing these words, we opted to use underlining to emphasize the word "one" in order to draw the respondents' attention to the fact that they should select only one of the response options. The choice of underlining for this purpose works quite well for paper surveys but should be carefully considered for web surveys because underlining already has a predefined meaning on the Internet, especially when combined with the color blue. Underlining on the Web often denotes a clickable link, although many web designers use underlining inconsistently, and sometimes links are not underlined.

Nevertheless, we use underlining in the same way in the merit-based pay example in the bottom of the figure to draw respondents' attention to the fact that this question is about elementary school teachers only. In this example we also face another common problem: the need to include extra information for survey processing reasons. In this case, the extra information is numbers located inside the check boxes to assist with data entry. Because they are unimportant to the respondent, these numbers should be deemphasized if they cannot be eliminated altogether. We do this by manipulating the properties of size, contrast, and location. The numbers are made smaller and lighter to make them less obvious but still visible to the astute data enterer. They are then relocated from the center to outside the check box, where they are less likely to be noticed by respondents but can still easily be used for data entry.

Guideline 4.14: Use Design Properties with Consistency and Regularity

This general guideline may be the most important. Even if the meaning of a design element or property is not immediately intuitive, the respondent has a better chance of learning its meaning and applying it throughout the questionnaire if it is used consistently, both within and across questions. However, if design elements and properties are used inconsistently, like the

their meaning at each use. Doing so may require more patience and me energy than some respondents are willing to expend.

A good rule of thumb is to use each design element or property for one purpose. For example, no matter what question or what part of the q tion it is used in, underlining is only used to draw attention to impor words, white square boxes are only used as answer spaces, bolding is used to distinguish between the question stem and the answer options pending on the particular needs of one's survey. One can choose to use th design properties to convey different meanings than in this example, but important thing is that they be used with consistency and regularity.

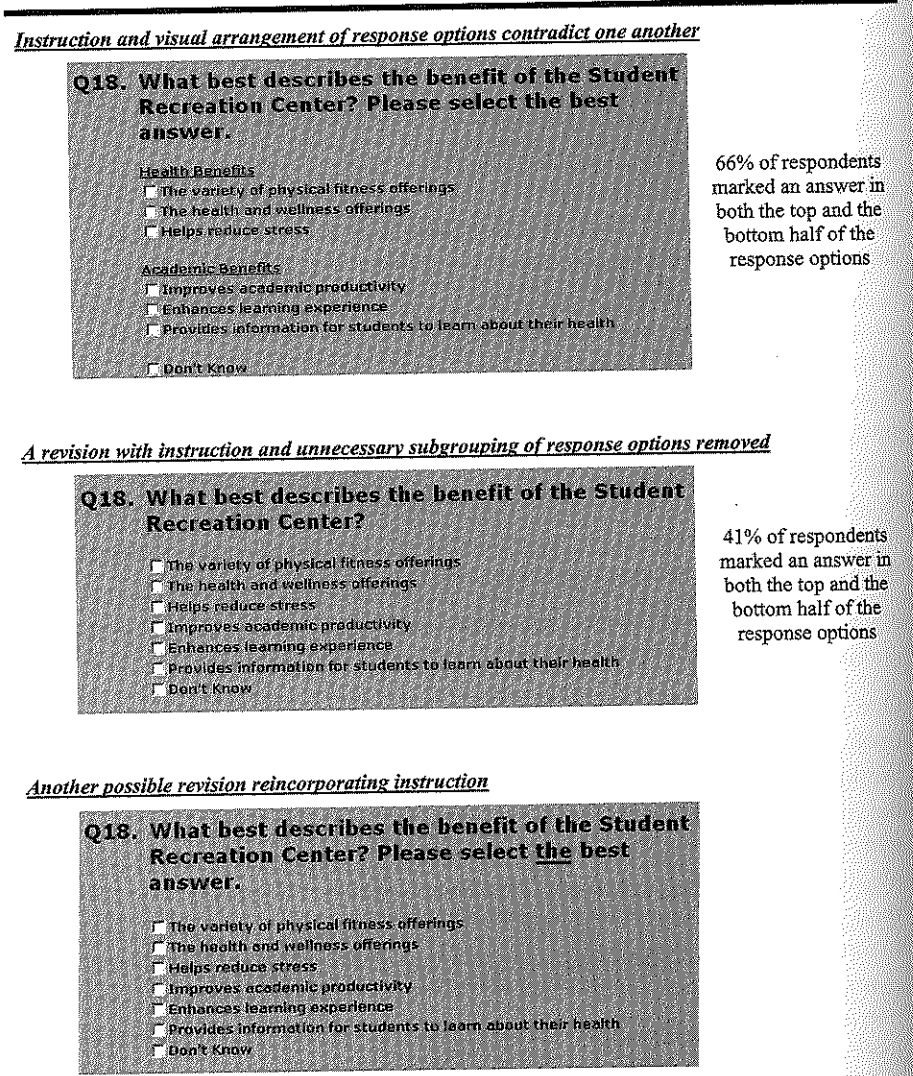
Taken together, these visual design guidelines (4.10–4.14) helped us ganize all of the information in the questions in Figure 4.19 to make th questions more easily perceived and processed and, most important, to m it easier for respondents to provide a response. In addition to easing response task, these changes also added an air of professionalism to the qu tions, thereby increasing the likelihood that they would be taken seriou by potential respondents (i.e., perhaps increasing rewards and trust). In remainder of this section we present several more general guidelines deal with instructions, the response task, and organization.

Guideline 4.15: Make Sure the Words and Visual Elements that Make Up the Question Send Consistent Messages

One of the biggest lessons we have learned since the last edition of t book was published is just how influential visual design elements can We know of multiple cases where surveyors made design decisions w the intention of easing the response task but with the result that the des features they introduced biased the results. In one national survey, for exa ple, the designers sorted related response options into two groups based their content. Pretesting revealed that the subgrouping caused respondei to make mistakes in answering the question. Many respondents attempted mark multiple answers, but because the survey was only designed to acc one answer, they inadvertently erased their first response when they enter their second.

Figure 4.20 shows examples of this type of design from a set of web surv experiments we undertook to examine the effects of subgrouping respor options in this way (Smyth, Dillman, Christian, & Stern, 2006b). When t response options were subgrouped with no instruction to select the best a swer (not shown in figure), 70% of respondents marked answers within bo groups compared to 41% when the response options were not subgroupe Adding the instruction "Please select the best answer" to the subgroup version reduced the number of respondents marking answers within bo groups to 66%. The instruction also reduced the mean number of optio

Figure 4.20 Make sure verbal and visual design elements send a consistent message.



Source: "Effects of Using Visual Design Principles to Group Response Options in Web Surveys," by J. D. Smyth, D. A. Dillman, L. M. Christian, and M. J. Stern, 2006b, *International Journal of Internet Science, 1*(1), pp. 6–16.

one option from each subgroup instead of multiple options from each subgroup. Overall, these findings suggest that the subgrouping communicated to respondents that they should select answers from both subgroups. Within

visual information provided by the subgrouping influenced how th
tion was interpreted.

This experiment provides an excellent example of how verbal a
elements can contradict one another, leading to errors in respor
haps more important, though, it demonstrates the importance of
back and looking at question construction holistically to ensure tha
words and the visual design of the question are sending a consistent
about the meaning of the question and the response task. In many
guidelines we present in this chapter and the next provide the tools
just that.

Guideline 4.16: Integrate Special Instructions into the Question Where T Be Used Rather than Including Them as Free-Standing Entities

Frequently it is necessary to provide a special instruction to clarify a
This leads to the undesirable practice of placing instructions outsi
question and emphasizing them with boxes or perhaps a different c
problem with this practice is that once people have gotten into the r
completing a questionnaire, the marking of an answer leads to the in
search for the next question. As a result, free-standing instruction
be skipped entirely. The example in Figure 4.21 shows that such ins
are most likely to be properly applied if they are expressed as pa
query itself rather than placed as a separate entity (Christian &
2004).

In the first layout, an instruction to skip this question and mo
the next if it does not apply is located below both the question s
the response options. In this design, 40% of respondents marked
only 5% left the question blank and moved on. In the second la
instruction is moved up to a more integrated location between t
tion stem and the response options. Placing the instruction here re
19% of respondents marking "no" and a full 26% leaving the questi
and moving on. When the instruction was located where it was n
help respondents decide whether and how they should answer t
tion, more people were able to successfully apply it. In contrast, v
instruction was located as a free-standing entity outside the quest
and answer categories, many respondents had probably already m
answer before they even noticed it. The fact that 11% of responden
version (compared to 3% when the instruction was integrated) left
question (Question 9) blank suggests that when some respondents
instruction located below the response options, they applied it to th
question altogether. The third layout, which has shown some promis
nitive interviews, is another possible way the instruction could be in
with the question stem.

As this example demonstrates, it is not enough to simply move

Figure 4.21 Integrate special instructions into the question stem.

Instruction placed outside of the navigational path

8. Have one-on-one meetings with professors contributed significantly to your WSU education?

☐ Yes
☐ No

40% marked No
5% left Question 8 blank
11% left Question 9 blank

If you haven't had many one-on-one meetings, just skip to Question 9.

A revision with the instruction placed within the navigational path

8. Have one-on-one meetings with professors contributed significantly to your WSU education?

If you haven't had many one-on-one meetings, just skip to Question 9.

19% marked No
26% left Question 8 blank
3% left Question 9 blank

☐ Yes
☐ No

Another possible revision with the instruction integrated with the question stem and visually distinguished using italics

8. Have one-on-one meetings with professors contributed significantly to your WSU education?
If you haven't had many one-on-one meetings, just skip to Question 9.

☐ Yes
☐ No

Source: "The Influence of Graphical and Symbolic Language Manipulations on Responses to Self-Administered Questions," by L. M. Christian and D. A. Dillman, 2004, *Public Opinion Quarterly, 68*(1), pp. 58–81.

appropriate question subgrouping. Rather, even within a single question, in order to be effective, instructions need to be strategically located where they will be used (the location of instructions is discussed in more depth with respect to establishment surveys in Chapter 12).

Guideline 4.17: Separate Optional or Occasionally Needed Instructions from the Question Stem by Font or Symbol Variation

When respondents begin to fill out a questionnaire, they are learning how the questionnaire works, including what must be read and what can be skipped. Requiring them to read through a great deal of material that does not apply or that can be skipped without negative consequences encourages the habit of skipping words and phrases. For these reasons a distinction should be made between words that are essential for every person to read and those that may

reading a particular instruction may be optional. Perhaps it is bec
instruction "put an X in the appropriate box" is the same instructi
for a previous question, and many respondents will remember that
also be that only a few respondents need the information, such a
case of the instruction used in Figure 4.21 ("If you haven't had ma
on-one meetings, just skip to Question 9"). To avoid presenting info
that respondents already know, or that applies to relatively few
distinguish this information from the query by the use of either it
shown for the second revision in Figure 4.21) or a symbol variati
putting it in parentheses).

Guideline 4.18: Organize Each Question in a Way that Minimizes the Nee to Reread Portions in Order to Comprehend the Response Task

The goal underlying this guideline is efficiency for the respondent
ure 4.22, a recreated excerpt from the 1993 U.S. Census of Agricult
inefficient for respondents to read in great depth about what land
sider in their answer before even knowing what the question is ask
inevitable result is that it will be necessary to reread the informati
discovering what the question is asking. The drawback to such ineffi
that respondents may become frustrated and unwilling to retrace tho
and therefore may give a wrong answer or no answer at all. In this
problem is confounded by a visual layout that makes it somewhat
what navigational path is to be followed (i.e., what information is to
in what order).

A more effective organization of the information is shown in t
sion in the bottom panel of Figure 4.22. The revision allows resp

Figure 4.22 Poor information organization with unclear navigational path.

Poor information organization and lack of navigational path

CENSUS USE ONLY | 035 | 036 | 037 | 038 | 039 | 040 | 041

SECTION 1 — ACREAGE IN 1992 – Report land owned, rented, or used by you, your spouse, or by the partnership, corporation, organization for which you are reporting. Include ALL LAND, REGARDLESS OF LOCATION OR USE – cropland, past
S1 *rangeland, woodland, idle land, house lots, etc.*

If the acres you operated in 1992 changed during the year, refer to the INFORMATION SHEET, section 1.

None ☐ Num 043

1. All land owned.....

Better information organization and creation of clear navigational path

1. How many acres of land did you own in 1990? You should report all land (crop land, pastur rangeland, woodland, idle land, house lots, etc.) regardless of location, owned by you, your s or by the partnership, corporation, or organization for which you are reporting. ***(If the acres operated in 1990 changed during the year, refer to the information sheet, Section I.)***

_____ Number of acres owned

to know at the beginning that they are being asked to report the number of acres they own; they are then given instructions on what to include and exclude. The important implication of this principle is that no amount of visual redesign can compensate for poorly worded questions or unorganized information, which, once read, leave the respondent unclear about precisely what to do.

Guideline 4.19: Choose Line Spacing, Font, and Text Size to Ensure the Legibility of the Text

Even a very well-worded question can be difficult for respondents to process if it is not designed in a legible way. Enhancing legibility means choosing an appropriate font, font size, and line length. With respect to fonts, one should avoid script fonts because they can be very difficult to read (e.g., *Brush Script MT*, *Walt Disney Script*, *Edwardian Script ITC*, *Freestyle Script*). Instead, serif or sans serif fonts should be used. Examples of serif fonts are Times New Roman, Garamond, Century, and Georgia. Each of these fonts has added detail, or *serifs*, at the end of the strokes that make up the structure of the letters. In contrast, sans serif fonts do not have the added details (e.g., Arial, Verdana, Tahoma, and Latha). Although both serif and sans serif fonts work well on paper, sans serif fonts are commonly preferred for web readability. Generally, one should also choose proportionally spaced fonts (e.g., Arial or Times New Roman) rather than monospace fonts (e.g., Courier New).

To some degree, the font size one chooses will depend on the survey population. A good rule of thumb is to use 10- to 12-point fonts for most populations but larger fonts for older populations. An additional consideration for web surveyors is that font preferences are often set on the user's computer, giving the designer little control. Additionally, font sizes appear different on screen than on paper, and the same font size may appear larger or smaller depending on factors such as the user's screen resolution. Thus, web designers are advised to seek additional resources on how to ensure legibility of text (more on this in Chapter 6).

With respect to line length, readers may have difficulty tracking along the lines, reading evenly, and finding their place at the beginning of the next line on the return sweep when text lines are too long. In comparison,

excessively short
lines of text require
almost constant eye
motion and frequent
return sweeps that
can become overly
burdensome.

CONCLUSION

Although crafting survey questions may seem simple, we have demons in this chapter how it requires attending to many details at once t ensure that respondents process all of the component parts and compr the question as intended so that they can report an accurate answer. W also discussed a holistic approach that highlights the many aspects sur need to think about when crafting survey questions. This approach co whether an open- or closed-ended question is best for each concept of i in the survey and requires thinking about how all of the componen question (the question stem, any additional instructions, and answer or response options) work together to form the entire question stimul

Drawing on the considerable amount of research on the importance visual design and presentation of survey questions, we have demons how crafting effective survey questions for mail and web surveys in not only choosing words to form clear questions but also deciding h components of the question are presented visually to respondents. Witl holistic framework, we have offered general guidelines to help survey they choose the words to form questions and visually design and p the components of the question. This second step is an important one often neglected in other guides for crafting questions.

The guidelines we have presented in this chapter apply to nearly all s questions. However, this is only the first of two chapters devoted to c survey questions. In the next chapter we shift our focus to crafting s types of open-ended and closed-ended questions. Because each of different question types has a different goal and a slightly differen figuration of question components, each is subject to its own stre and challenges. Although the general guidelines presented in the c chapter still apply to these question types, we take them a step furt focus on both the wording and visual presentation of the question ster additional instructions, the answer spaces, and response options th unique to these question types.

LIST OF GUIDELINES

Guidelines for Choosing Words and Forming Questions

(continu

Guideline 4.3: Ask one question at a time
Guideline 4.4: Use simple and familiar words
Guideline 4.5: Use specific and concrete words to specify the concepts clearly
Guideline 4.6: Use as few words as possible to pose the question
Guideline 4.7: Use complete sentences with simple sentence structures
Guideline 4.8: Make sure "yes" means yes and "no" means no
Guideline 4.9: Be sure the question specifies the response task

Guidelines for the Visual Presentation of Survey Questions

Guideline 4.10: Use darker and/or larger print for the question and lighter and/or smaller print for answer choices and answer spaces
Guideline 4.11: Use spacing to help create subgrouping within a question
Guideline 4.12: Visually standardize all answer spaces or response options
Guideline 4.13: Use visual design properties to emphasize elements that are important to the respondent and to deemphasize those that are not
Guideline 4.14: Use design properties with consistency and regularity
Guideline 4.15: Make sure the words and visual elements that make up the question send consistent messages
Guideline 4.16: Integrate special instructions into the question where they will be used rather than including them as free-standing entities
Guideline 4.17: Separate optional or occasionally needed instructions from the question stem by font or symbol variation
Guideline 4.18: Organize each question in a way that minimizes the need to reread portions in order to comprehend the response task
Guideline 4.19: Choose line spacing, font, and text size to ensure the legibility of the text

CHAPTER 5

Constructing Open- and Closed-Ended Questions

TO SUCCESSFULLY undertake almost any task, one has to attend to man at once. For example, going for any type of bicycle ride requires atte balance, pedaling, and steering. Failure to properly attend to any or (especially balancing!) can have significant and far-reaching consec However, to have a *great* bicycle ride, one has to look far beyond the mon details and focus on how the mechanics of the specific bicycle wo the type of ride desired. If the bicyclist wants to ride on flat paved ro would need smooth thin tires, a light frame, and high gears to have ride. The same bike, however, would not hold up to mountainous terr this type of riding, she would need wide and knobby tires, a sturdie and lower gears to get the best ride.

There are many parallels between planning for a great bicycle ri constructing survey questions. A set of common guidelines, such as th cussed in Chapter 4, is useful in that it provides a general framework which to craft all types of survey questions, but, as with the bicycle ex constructing a set of great survey questions also requires carefully de each component of every question in light of its unique measuremen Just as a bicycle rider needs to put the right components together in th of different types of bikes for different types of rides, a survey designe to know how to put together the components of different types of questions to effectively measure different types of constructs. In this c we pay detailed attention to how different types of questions—open- closed-ended nominal, and closed-ended ordinal—need to be const to perform optimally. We continue to apply the techniques discus