
CHAPTER 2

Writing Questions

The goal of writing a survey question for self-administration is to develop a query that every potential respondent will interpret in the same way, be able to respond to accurately, and be willing to answer. However, in practice, producing good questions is often difficult.

A methodologist friend once described the question-writing task as similar to driving in freeway traffic while drinking a cup of hot coffee and answering an emergency call on his cell phone. Many things are competing for attention, and failure to heed any of them can spell disaster. It is the need to consider many competing things at once that makes it difficult to write questions for self-administered surveys. But unlike freeway driving while tending to coffee and telephone calls, which is not advisable, many things need to be considered when writing questions. Moreover, there is no alternative to understanding and dealing with each of the competing concerns that is addressed in this chapter.

Consider, for example, responses to the question, “How many hours per day do you typically study?” asked in a mail survey of university students (Rockwood, Sangster, and Dillman, 1997). As shown in Figure 2.1, different response categories, a high set and a low set, were used. Five of the six categories in the low set cover the same range of hours (less than 2½ per day) as that covered by only one of the six categories in the high set (Figure 2.1). Use of the low set resulted in 23% of the respondents reporting that they studied more than 2½ hours per day, compared with 69% when the high set was used. Clearly, the answers to this question were influenced by more than the words used to describe each response category.

Two other samples of students were interviewed by telephone for this experiment, with half being presented with each set of categories. Among those presented with the lower set of categories, 42% (compared to 23% by mail) reported studying more than 2½ hours per day. Thus, the choice of surveying by mail versus telephone also influenced people’s answers. These results could

Figure 2.1 Low and high sets of categories used to ask students how many hours they 1) studied and 2) watched television each day, and the results (Rockwood, Sangster, and Dillman, 1997).

<i>Version A (low) Categories</i>		<i>Version B (high) Categories</i>	
Less than .5 hour per day			
.5–1 hour			
1–1.5 hours			
1–2 hours			
2–2.5 hours per day		Less than 2.5 hours	
More than 2.5 hours		2.5–3 hours	
		3–3.5 hours	
		3.5–4 hours	
		4–4.5 hours	
		More than 4.5 hours	

Reported Hours/Day	Version A (low)		Version B (high)	
	Mail	Telephone	Mail	Telephone
<u>Studying:</u>				
2.5 hours or more	23%	42%	69%	70%
Less than 2.5 hours	77%	58%	31%	30%
<u>Watching Television:</u>				
2.5 hours or more	17%	17%	32%	31%
Less than 2.5 hours	83%	83%	68%	69%

have placed the surveyor in the unenviable position of reporting to the study sponsor that as few as 23% or as many as 70% of the students at this university were studying more than 2½ hours per day, depending upon which version of the questionnaire and survey method she wanted to believe. Unfortunately, the measurement challenge did not end here.

Another, nearly identical question included in these experimental surveys asked about how many hours each day students watched television. For this question the alternative sets of categories had similar effects on responses; 17% reported watching television 2½ hours or more per day in response to the low set, compared with about 32% when the high set was used. However, in contrast to the question about studying, whether the survey was done by mail or telephone made absolutely no difference in how many people selected each answer category.

Why did these results for the alternative sets of categories and the two questions differ so dramatically? Which answers should be trusted? In addition, how can you ever be sure that whatever questions are asked in surveys obtain answers that provide the best possible estimates of the distribution of the characteristic of interest in the population? These issues and more are the topics of this chapter. In the course of this discussion, I will return to the likely reasons that these questions produced such divergent answers. The purpose of this chapter is to provide principles, as well as specific tools, to help you write questions for self-administered surveys in ways that will produce answers you can trust.

CRITERIA FOR ASSESSING EACH SURVEY QUESTION

Survey questions fail in their purpose for many reasons, ranging from use of the wrong words or an inappropriate structure to simply not being answerable. In addition, a question that will work fine in one survey may not be satisfactory for another. Attempting to help hundreds of people, ranging from first-time surveyors to experienced methodologists, with their surveys leads me to conclude that too often sponsors have not figured out what they want to know from respondents, except in a general sort of way: “I want to know why people have such negative attitudes about. . . .” Other sponsors know fairly precisely the questions they want to ask, and still it is very difficult to progress from a list of draft questions to a set of *good* survey questions. Some questions are easy to write and ask, while others require draft after draft and much testing. Once a working draft of proposed survey questions and response choices (if any) has been prepared, ask each of the following eight inquiries about each of the proposed survey questions. Answers to these questions will help diagnose problems and guide you towards the structural and wording decisions that are appropriate for your study.

QUESTION 1: DOES THE QUESTION REQUIRE AN ANSWER?

A survey question is more than a general inquiry. It is the surveyor’s tool for gaining responses from subjects in a survey sample that will make it possible to determine the distribution of a characteristic (an attitude, belief, behavior, or attribute of each respondent) in the survey population. In order for an inquiry to constitute a survey question, it must require an answer from each person to whom the question is asked. Neither of the following questions meet that criterion.

If you fixed dinner at home last night, did you eat meat as part of that meal?

- ☐ Yes
- ☐ No

When you go out to eat, which type of food do you most prefer?

- ☐ American
- ☐ Italian
- ☐ Chinese
- ☐ Other

A respondent who ate out the previous night could not answer the first question. A respondent who never eats out could not answer the second one. Use of the introductory words “if” or “when,” invites some respondents not to provide an answer. Inasmuch as not all respondents are required to provide an answer to the above question, we cannot distinguish nonresponse from those to whom the question does not apply. Even if a “does not apply” box were provided, the wording of the first question implies that no response is needed from those who ate out the previous night. In addition, for us to be able to estimate the distribution of a characteristic in the sample and (through it) the population, respondents must be given the opportunity to answer every question they are asked.

QUESTION 2: TO WHAT EXTENT DO SURVEY RECIPIENTS ALREADY HAVE AN ACCURATE, READY-MADE ANSWER FOR THE QUESTION THEY ARE BEING ASKED TO REPORT?

Some survey questions are easier to get accurate responses for than are others. For example, virtually everyone knows their age. People are frequently asked how old they are (birthdays are special events in the lives of most people) and age is a number that people are expected to know. People are also able to report age accurately in response to many ways of asking the question, as shown in Figure 2.2. Assuming a willingness to report age, it does not make a lot of difference how the question gets asked. From the standpoint of expected precision of the answer, the question with three broad categories will, of course, obtain a less precise answer. However, from the perspective of whether the age question will be answered accurately, the choice of question structure makes little difference.

However, consider the other question shown in Figure 2.2, concerning whether tall people are more likely to be elected president. Most people do not have a ready-made answer that they can immediately report. Some might wonder what is meant by “more likely” in the stem of the question. Others might wonder what the difference is between “somewhat” and “strongly” in the answer choices. To answer, they not only have to give definition to the question and answer choices, but they may also try to recall the height of current and past presidents and the relative heights of their opponents. Some might spend a long time and think about a lot of presidents, while others think

Figure 2.2 Examples of questions to which respondents already have a ready-made answer (which produce accurate answers) versus one that asks about an informed opinion that produces inconsistent answers.

Respondent's age; the exact wording makes relatively little difference to response accuracy:

33. How old are you?

_____ Age

33. What is your current age?

_____ Years

33. When were you born?

_____ Year born

33. What is your date of birth?

___ Day ___ Month ___ Year

33. Which of the following age category describes you:

- ☐ 35 or younger
- ☐ 36–65
- ☐ 66 or older

Respondent's opinion on issue for which wording changes produce substantial inconsistencies in response choice:

“Tall people are more likely to be elected President of the United States.”

Do you strongly agree, somewhat agree, somewhat disagree, or strongly disagree with this statement?

(or) Do you very strongly agree, strongly agree, agree, disagree, strongly disagree, or very strongly disagree?

(or) On a scale of 1 to 7, where 1 means entirely agree and 7 means entirely disagree, use a number to indicate how strongly you agree or disagree.

only about the current president's height in relation to the opponent in the last election until a satisfactory answer is reached. Still others may not think about any specific presidents or candidates at all.

In contrast to the near automatic answer for a question about age, when people respond to broad opinion questions they must comprehend the question and answer choices and contemplate what each of them means. They must recall information about the topic, make a judgment about the information they retrieve from memory, and then report a response, a process that has

been described in detail by Tourangeau (1992). Each of these stages is influenced by different contextual considerations and processes. People do not have ready-made responses to opinion and belief questions to nearly the same degree that they possess answers to questions about their educational level, the kind of car they drive, or whether they own their home. The vaguer the question, the vaguer the categories (e.g., using only numbers, as in a 1–7 scale), and the more remote these items are from people's experiences, the more likely a question is to produce inconsistent responses if we ask the same person to answer this question at different times.

Sometimes it is necessary to spend a lot of time drafting, writing, and testing alternative wordings of questions, and even then we can obtain only an approximate answer. In other instances, writing questions becomes a very easy task, requiring little time and effort.

Thus, a critical step towards writing good questions is to understand the extent to which respondents have an accurate, ready-made answer, and whether creating an answer demands considerable thought that is subject to myriad influences, including the context in which the question is asked. If people do not have a ready-made answer to a question, getting an accurate answer becomes more difficult.

QUESTION 3: CAN PEOPLE ACCURATELY RECALL AND REPORT PAST BEHAVIORS?

Whereas abstract beliefs and attitudes are hard to measure, it would seem on the surface that people should be able to report past behaviors. However, that is not always the case.

Asking how many hours a person watched television on the first Sunday of the previous month is an example of a behavior that is hard to remember. Respondents are unlikely to be able to recall something so precise from that long ago. To solve this problem, surveyors often ask how many hours per day a person "usually" watches television, an example of which was provided at the beginning of this chapter. To answer, the respondent must recall what she "usually" does and estimate. Recalling the frequency of these routine or mundane behaviors is subject to considerable potential error, as evidenced by the effect of the category choices offered as part of the survey question (Rockwood et al., 1997). Asking people to reconstruct how much time they have spent studying or watching television in the last four hours, on the other hand, is less subject to unintentional error.

Frequently, people who write surveys want respondents to provide far more detail about past behaviors than can be recalled. Thus, determining whether people are able to recall information needed for answering each proposed survey question is important. Keeping recall simple and related to recent events helps to produce high quality survey data.

QUESTION 4: IS THE RESPONDENT WILLING TO REVEAL THE REQUESTED INFORMATION?

The fact that respondents know an answer does not mean they are willing to provide it. Many respondents are reluctant to reveal certain information about themselves, for example, their income. Others may be unwilling to answer questions about previous drug use, or having shoplifted when they were teenagers.

Considerable evidence suggests that people are more likely to give honest answers to self-administered than to interview questionnaires (e.g., de Leeuw, 1992; Fowler, Roman, and Di, 1998; Aquilino, 1994). For example, when asked the question, "How often have you driven a car after drinking alcoholic beverages?" only 52% responded "never" to the self-administered questionnaire versus 63% for the comparable telephone survey (Dillman and Tarnai, 1991).

Although self-administered questionnaires are often selected because of respondents' greater honesty with their answers, there is little doubt that social desirability is somewhat of a problem for this method as well. A recent article suggests that young people responding to computers may provide more truthful answers than with self-administered questionnaires (Turner, Ku, and Rogers, 1998).

QUESTION 5: WILL THE RESPONDENT FEEL MOTIVATED TO ANSWER EACH QUESTION?

Motivation to respond distinguishes self-administered from interview surveys dramatically, and in a way that counts. Designers of interview surveys often write questions independent of any motivational considerations, leaving it to the interviewer to encourage, cajole, or otherwise persuade respondents to carefully select and report complete answers. Unless consideration is given to how people react to questions in self-administered questionnaires, instructions may be ignored and incomplete answers given. Also, the question may be skipped or, even worse, the questionnaire not returned.

For example, questionnaires are sometimes constructed in a way that requires respondents to consult a separate instruction booklet to understand unclear questions. Questions are sometimes presented in matrices with row and column headings that must be connected in order to understand the question. Further, it may be impossible for the respondent to understand the question without consulting a separate instruction booklet to determine what is meant by each row and column heading. Thus, three widely separated items of information have to be connected by the respondent in order to figure out an appropriate answer. Another kind of question with adverse motivational qualities is one that asks people to rank a large number of items

from top to bottom, such as a list of 20 “priorities for economic development.”

Motivation can be encouraged in many ways, ranging from incentives and follow-up reminders to respondent-friendly questionnaire design. In some instances, the questions themselves are the source of a motivational problem and no matter how much one does with other aspects of survey design, wording remains a major impediment to accomplishing the survey objectives. Sometimes, modifying questions (e.g., stating them more clearly) or changing a ranking question to a rating question (e.g., asking the importance of each of the 20 economic priorities on a scale) will improve the likelihood of getting an answer.

QUESTION 6: IS THE RESPONDENT’S UNDERSTANDING OF RESPONSE CATEGORIES LIKELY TO BE INFLUENCED BY MORE THAN WORDS?

Responses to the questions about studying and watching television presented at the beginning of this chapter were clearly influenced by more than words alone. Some respondents appeared to see a set of categories that were viewed as running from low to high. Choice of the highest category may have seemed appropriate to respondents who saw themselves as in the “top” group of TV watchers or studiers. Respondents faced with a situation in which they don’t have an obvious answer may respond to these questions partly in terms of where they see themselves in relationship to other students, for example, “I study more than most, so I should choose one of the top categories.” It is important to recognize that category ranges and visual layout (as discussed in the next chapter) also provide important clues used to select the appropriate answer. Words are only part of the question stimulus.

Attitudinal and belief questions typically rely on vague quantifiers, such as strongly favor to strongly oppose, high priority to low priority, agree to disagree, or even vaguer ones that rely on numbers such as –3 to +3, 1–7, or 1–10. Such numerical scales require respondents to give a certain amount of definition to any category they choose to use. The vaguer the question and answer categories, the greater the potential for measurement error.

QUESTION 7: IS SURVEY INFORMATION BEING COLLECTED BY MORE THAN ONE MODE?

Increasingly, more than one survey mode is used to collect information for a single survey. This means that data collected by each mode needs to be comparable with that collected by another. An illustration is the fact that 42% of the telephone respondents versus only 23% of the mail respondents chose more than 2½ hours in response to the hours of study question posed at the beginning of this chapter (Figure 2.1).

It appears from the available research that several different kinds of mode differences may occur between self-administered and interview surveys (Dillman, Sangster, Tarnai, and Rockwood, 1996), and that the introduction of electronic survey technologies such as Interactive Voice Response, as discussed in Chapter 11, introduces even more challenges (Srinivasan and Hanway, 1999). However, several remedial actions, ranging from restructuring questions to changing question wording, can be done to minimize mode effects. For this reason, Chapter 6 is devoted entirely to the special challenges of mixed-mode surveys.

QUESTION 8: IS CHANGING A QUESTION ACCEPTABLE TO THE SURVEY SPONSOR?

Questions with recognized defects cannot always be changed. Sometimes a particular question has been used in another survey, and the main objective is to replicate the previous survey or make the new data comparable in some other way. Examples are government surveys that have asked the same question repeatedly in order to produce time-series data, sometimes for decades. In other instances the surveyor may be willing to accept a lower response rate or higher item nonresponse in an effort to get the more precise answers only some respondents are able to provide. This is a tendency of some economic surveys, when offering broad income categories is considered an unacceptable alternative to requesting exact income. Often sponsors want information about which respondents have virtually no knowledge and for which formulating a meaningful answer is difficult. Political considerations may also dictate the selection of question and answer categories. Thus, it is important to ask sponsors whether questions that appear troublesome are subject to change and if so, how much.

Writing questions is a difficult challenge precisely because so many factors simultaneously influence whether a proposed question obtains accurate answers. The eight questions listed here constitute the mental checklist I attempt to apply when, inevitably, surveyors asking for help thrust a draft in front of me and ask, "What do you think?" The search for tools and principles for writing questions is a means of overcoming each of these potential problems. This is the topic to which I now turn.

WHICH QUESTION STRUCTURE IS MOST APPROPRIATE?

Fundamentally, there are only three different ways a survey question can be structured. One way is to pose a query as an open-ended question, an item for which no answer choices are provided. The others provide answer choices which can, in turn, be structured as closed-ended in either of two ways—as ordered or unordered response categories. Shifting from one structure to

another is the most fundamental tool available for responding to the kinds of problems suggested above.

OPEN-ENDED QUESTIONS

An inability to get adequate answers to open-ended questions is often identified as a chief disadvantage of self-administered surveys. After all, there is no interviewer to respond to an unclear answer with a probe, such as, "Could you provide a little more detail so that I'm sure I understand what you mean?" However, the issue of whether questions that don't offer answer categories are acceptable for mail surveys is somewhat more complicated. Consider, for example, this open-ended question:

A. Why did you purchase a new automobile?

_____ Reason for purchase

Answers to Question A are likely to be inadequate. This question will not elicit self-administered answers that are as complete as those offered in interview surveys when supported by interviewer probes. People don't necessarily have a ready-made answer to such a question in the way they do for such things as age, and the answers are prone to considerable unintentional error. Although the above question format could be improved somewhat by providing more space, perhaps 2–3 lines to visually suggest that a longer answer is desired, answers are not likely to be as complete as one would like. The fundamental problem with questions of this nature is that the answer depends upon the extent to which respondents are willing to think hard about the question and write a complete answer. In such cases a surveyor would be well advised to list possible reasons for buying a car and turn the question from open-ended to closed-ended.

When we have proposed that kind of solution, two counter arguments are often extended. One is that surveyors do not know what the possible reasons might be, so they cannot list them. Second, the question is viewed as "exploratory," that is, an attempt to get some idea of what reasons might exist, which in a future survey could be built into closed-ended formats. Thus, despite the shortcomings of this type of open-ended question, it is sometimes appropriate to use in self-administered surveys.

Question B is quite different from Question A.

B. What kind (make and model) of automobile do you currently drive?

_____ Make _____ Model of car

Most people can provide the make (e.g., Ford) and, to a lesser extent, the model (e.g., Taurus) of the car they drive. To list all possible makes and models would require an enormous amount of space and require that respon-

dents search the list carefully to find their car. This type of open-ended question works quite well in self-administered surveys. I find it helpful when asking such questions to provide blank spaces followed by the units for reporting (here, make and model) rather than simply printing an unlabeled space.

At the beginning of this chapter, I noted the biasing effects that result from offering response categories for mundane characteristics such as hours of study or watching television. Asking these questions in an open-ended form with a blank space followed by units, for example, “_____Hours/Day,” is a better way to obtain accurate answers.

Frequently, when designing self-administered surveys we have been asked to elicit the respondent’s occupation, as in Question C below.

C. What is your current occupation?

_____ Occupation

This question produces unacceptably vague answers. For example, a “building engineer” might be someone with a college degree in engineering or one who performs custodial services. When faced with the need to obtain occupational information in self-administered surveys, one solution is to break the question into multiple parts, as shown here:

What is your current occupation?

_____ Current occupation

What kind of work do you do?

_____ Kind of work

What is the name of your employer?

_____ Name of employer

Other specific queries, such as the amount of education one normally needs to perform this job, might be added to assure sufficient information for deriving an occupational code. In essence, this format builds in the kinds of probes that an interviewer might ask to get adequate information for understanding a respondent’s occupation.

In sum, open-ended questions are frequently very useful in self-administered surveys, but their usefulness depends upon the nature of the questions as well as the way in which they are structured. Sometimes that means changing to an open-ended format, while in other instances it means avoiding it.

CLOSED-ENDED QUESTIONS

Consider these two questions:

- A. To what extent do you favor or oppose expanding the number of franchises operated by this company?
- ☐ Strongly favor
 - ☐ Somewhat favor
 - ☐ Neither favor or oppose
 - ☐ Somewhat oppose
 - ☐ Strongly oppose
- B. Which one of the following do you think should be our company's highest priority?
- ☐ Opening more franchises
 - ☐ Improving remuneration for current employees
 - ☐ Expanding our product line
 - ☐ Improving skills of existing employees through training

The first of these two questions provides respondents with a response scale and the answering task is to determine where one best fits on that scale. Thus, it consists of carefully ordered categories. The second question presents categories in no particular order, and respondents are asked to pick the one that best describes their opinion.

The mental effort needed for answering each of these questions is quite different. Whereas the first question requires envisioning a scale and figuring out where on that scale one fits, the other requires comparing each discrete category with the others, a task that usually is more difficult. Typically, questions with ordered categories consist of two basic concepts, with one presented in the stem of the question (expansion of franchises) and another in the answer choices (favor versus oppose). Questions with unordered categories may consist of many different concepts which must be evaluated in relation to each of the others. In addition to the question stem concept (top priority) the answer categories reveal four distinctly different concepts (alternative business goals) for consideration. Typically, closed-ended questions with unordered categories are more difficult for respondents to answer because of the amount of information that must be processed and about which decisions must be made. However, each type can be enormously useful in a self-administered questionnaire. They also provide quite different writing challenges.

Closed-Ended Questions with Ordered Response Categories

This type of question is most useful when one has a well-defined concept for which an evaluative response is wanted, unencumbered by thoughts of alter-

native or competing ideas. The list of scalar concepts that might be used to evaluate a concept or idea seems almost endless. Here are a few possibilities:

- Strongly agree to strongly disagree
- Very favorable to very unfavorable
- Excellent to poor
- Extremely satisfied to extremely dissatisfied
- High priority goal to low priority goal
- A complete success to a complete failure
- A scale of 1–7 (or 1–10 or 1–100) where 1 means lowest possible quality and 7 (or 10 or 100) means highest possible quality
- A scale of –3 to +3 where –3 means completely lacks this characteristic and +3 means completely exhibits this characteristic

These answer choices are frequently referred to as “vague quantifiers,” a moniker that tells a lot about their measurement characteristics. Generally, these types of scales request answers the respondent may not have ready-made, and which are therefore subject to considerable measurement error. However, when a surveyor wants to obtain separate respondent evaluations of many different concepts (e.g., 20 possible areas for spending agency funds) and compare preferences across areas, there may be no alternative to this approach.

Using closed-ended, ordered categories of this nature involves making many decisions. One such decision is how many choices should be offered. A second decision is whether to label each choice and, if so, what to call it. The nature of these decisions and factors associated with making a decision can be illustrated by the experiences of two researchers who became dissatisfied with a community satisfaction measure then in common use (Andrews and Withey, 1976). The traditional four-choice scale shown below was quite common, but because most people tended to be “very satisfied,” the clustering of answers into a single cell limited the analyses they could perform. The researchers extended the scale from four to seven answer categories, and relabeled them as shown in alternative B.

A. *Traditional Measure of Community Satisfaction*

How satisfied are you with your community?

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

B. *Proposed Measure of Community Satisfaction*

How satisfied are you with the community where you live?

- ☐ Delighted
- ☐ Mostly satisfied

- ☐ Slightly satisfied
- ☐ Neither satisfied nor dissatisfied
- ☐ Slightly dissatisfied
- ☐ Mostly dissatisfied
- ☐ Terrible

The researchers were successful in dividing people's answers across more categories and accomplished their analytic objective, but they also produced a scale that has not been used much in the ensuing years. I suspect the reason for the lack of use is that the scale labels strike most potential users as a little strange. At first glance the scale appears to convey two distinct concepts, degree of satisfaction plus the delighted-terrible extremes, an inconsistency that one should try to avoid. Moreover, the combination seemed to be unclear to respondents. The intended use of responses may also make a difference with regard to the suitability of these items. If one plans to present results to community officials for discussions on community goals, there is some risk that the terms delighted and terrible might not be taken seriously in that discussion.

In order to appreciate another potential difficulty with this scale, I suggest that you pick up the telephone and ask a friend to allow you to interview him or her. Read the above questions out loud and then assess how easy it is for the scale to be read and comprehended, and whether you think the question is likely to be taken seriously. The movement towards telephone interviewing in the last few decades has encouraged people to use somewhat fewer categories and simplify their presentation to respondents for easier comprehension. Similarly, as discussed in Chapter 6, the need for unimode construction has led me to keep my scales simpler as opposed to more complex.

Simplicity has also been achieved by stripping scales of labels completely and asking people to place themselves on one to seven or other length scales. Most people can use such scales by telephone as well as in self-administered questionnaires. The disadvantage of doing so stems from removing a sense of meaning which a label gives to each point on a scale. One surveyor who frequently conducted policy surveys said he was comfortable going to a city council meeting to report, for example, that 30% of residents were somewhat satisfied with police protection and another 20% were completely satisfied, but did not feel at ease reporting that people were on average 3.1 on a four-point scale. The more abstract the concept and scale presented to respondents, the greater the potential for any given answer to have a less concrete meaning than when a word label is attached.

Closed-Ended Questions with Unordered Response Categories

Some questions that use unordered or nonscalar response categories, such as, "Do you rent or own your home?" are quite simple. However, choosing from among several or many categories is quite complex, as suggested by these two examples:

- A. Show which of these six groups you feel should have the most and least influence in deciding whether the proposed community bypass should be built by putting a “1” in the box for most influence, “2” for second most influence, and so on until you have ranked all six choices:
- ☐ Local chamber of commerce
 - ☐ State highway department
 - ☐ The city council
 - ☐ The mayor
 - ☐ The voters
 - ☐ Local businesses
- B. If the highway bypass is to be built on one of these routes, which would you most prefer?
- ☐ A north route that starts west of the city at Exit 21 (Johnson Road) off Highway 30, crosses Division at North 59th Street, and reconnects to Highway 30 three miles east of the city at River Road.
 - ☐ A modified north route that starts further west of the city at Exit 19, crosses Division at 70th Street, and reconnects to Highway 30 three miles east of the city at River Road.
 - ☐ A south route that begins west of the city at Exit 19, crosses Division at South 24th Street, and reconnects to Highway 30 east of the city at River Road.

Responding to either of these questions requires considerable effort. The first question requires comparing six groups, then five, then four, and so forth to complete the ranking. The second question requires absorbing considerable detail, identifying differences between the choices, and then selecting the most preferred route. In both cases, providing answers requires considerably more effort to comprehend and decide how to answer than is usually the case for closed-ended questions with ordered categories. Yet these are precisely the types of question structures that can sometimes provide the most useful information to survey sponsors.

One of the challenges of writing questions of this nature is to keep the demand on respondents from getting out of hand. Each answer choice adds another concept which must be compared with other choices. Trying to provide detailed information on alternatives, as in the second example above, increases the likelihood of some small detail on a proposal being overlooked (e.g., the exact location of South 24th Street). On paper, a question that lists 15 items to be ranked from top to bottom may not look much more difficult than one which has only six items to be ranked, but the respondent demand is obviously far greater. If all 15 options need to be presented, then the question might be simplified by asking for a ranking of only the top three.

Partially Closed-Ended Questions with Unordered Response Categories

Sometimes a surveyor gets caught between feeling a need to list many categories, despite the difficulty, versus not allowing respondents to give their preferred answer.

Consider, for example, this question about college sports.

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

- ☐ Football
- ☐ Basketball
- ☐ Track and field
- ☐ Baseball
- ☐ Other (Please specify: _____)

Limiting choices to the first four alternatives means that some people will not be able to give the answer they want to provide. Trying to list all of the alternatives means a potential quagmire of making sure to list a huge array of minor sports, such as lacrosse, rugby, bowling, or rowing. Writing a partially open-ended question is a possible solution, but not one that lends itself well to the construction of variables for analysis. It is likely that fewer people will mention other choices if they are not listed. However, the open-ended responses might be analyzed separately to see which other sport is mentioned most often. The difficulty comes when one tries to make statements such as, "Respondents to this survey are 10 times as likely to say football is their favorite sport as they are to say bowling." Such statements cannot be defended because the stimulus of a bowling category was not offered to respondents. Another possible revision may be to change the wording of the question in a way that limits its scope only to the offered categories:

Of the four college men's sports listed below, which one do you most like?

- ☐ Football
- ☐ Basketball
- ☐ Track and field
- ☐ Baseball

In this case, the absence of the "other" category does not force people to give an answer which they think will be misinterpreted with respect to their favorite college sport.

WHY IS A WORKING KNOWLEDGE OF DIFFERENT QUESTION STRUCTURES IMPORTANT?

Survey sponsors often get locked into a type of question structure, making it much more difficult for them to accomplish their survey objectives. In some instances, they ask only open-ended questions because of concern about forcing

people to choose from among alternatives. In other cases, I have been given a confusing mixture of good ideas in formats that intermingled ordered and unordered categories. For example, I was once asked to help a university committee that was preparing a questionnaire to evaluate a dean's performance. All of the questions they proposed were structured as closed-ended questions with unordered response categories, similar to the one shown in Figure 2.3.

The first question sought information about both leadership and innovation skills but seemed destined to produce results that were difficult to inter-

Figure 2.3 Restructuring a question to provide interpretable answers.

An uninterpretable confirmation of closed-ended ordered and unordered categories:

6. Which of these five statements best describes this 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

Revision—Ordered categories for each concept:

6. To what extent has the dean demonstrated strong leadership qualities?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ Seldom
- ☐ Never

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

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ Seldom
- ☐ Never

Revision—Unordered categories that achieve head-to-head comparison of concepts:

6. 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
-

pret. The proposed solution was to break the question apart using both the ordered and unordered question structures, as shown in the first revision in Figure 2.3. Doing so allowed the committee to accomplish its stated objectives of finding out how faculty viewed the dean separately with regard to leadership and innovation skills, and also on which attribute she scored best.

An examination of questions posed about election candidates also illustrates how a knowledge of different kinds of question structures can help get information of greatest use to survey sponsors (Figure 2.4). When one thinks

Figure 2.4 Voting preference questions posed in four structural formats, appropriate to need.

Completely open-ended:

3. Who would you most like to see elected President of the United States in the next election?

_____ Name of person I'd most like to see elected
President of the United States

Closed-ended with scalar categories:

3. For each of the following candidates, please tell how qualified you feel he or she is for becoming president:

Emma Cain:
☐ Very well qualified
☐ Fairly well qualified
☐ Somewhat qualified
☐ Somewhat unqualified
☐ Not well qualified
☐ Not at all qualified
Etc.

Closed-ended with unordered categories:

3. If the election for President were being held today, for which one of these candidates are you most likely to vote?

☐ Martha Holmes
☐ David Badger
☐ Harold Gurwell
☐ John Williby

Partially closed-ended question:

3. If the election for President of the United States were being held today, for whom are you most likely to vote?

☐ Laurel F. Lugar, Democrat
☐ Charles Young, Republican
☐ Other (Write in name of other choice) _____

of election surveys, it is usually expected that head-to-head questions, as represented by the third example listed here, will be used. However, when candidates have not yet declared, and the question of much interest is who has public visibility and exists in people's minds as a potential candidate, then a completely open-ended question may be the best type to ask. If one is thinking about "dimensions" of the candidate, perhaps as a campaign manager devising beneficial advertisements, knowing how one's candidate is perceived and how that differs from perceptions of other candidates would be enormously useful. The scalar question for asking how well qualified each candidate is to be president will provide a sense of "distance" between candidates rather than a comparison of who would be best and worst. The head-to-head comparison, or closed-ended unordered structure, would be of obvious use during an election campaign, providing feedback that is typically used to design campaign strategies. The partially open-ended question is of use when write-in candidates are involved in the race. Thus, each type of question has much usefulness for determining voting preferences.

Changing question structures in this way enables us to formulate appropriate questions for achieving survey objectives. Each type of question structure performs a role that no other type can achieve as well, making an understanding of question structure a fundamental tool for drafting acceptable survey questions.

PRINCIPLES FOR WRITING SURVEY QUESTIONS: THE COMBINING OF WORDS AND STRUCTURE

Words are the building blocks for all question structures, but deciding which words to use and in what order is far from simple. The wrong choice of words can create any number of problems, from excessive vagueness to too much precision, from being misunderstood to not being understood at all, and from being too objectionable to being uninteresting and irrelevant. No one would deny that it is important to find the right choice of words. However, making the right choice in any given situation and knowing when you have achieved it are issues on which agreement is far less likely.

Perhaps no one has summarized the dilemma of writing questions as elegantly as Stanley Payne (1951) nearly 50 years ago in *The Art of Asking Questions*. In one chapter he presents 41 versions of a single question before finding one that he considers acceptable for a survey. Even this question is cautiously labeled "passable," and the reader is admonished that pretesting might shoot this 41st version full of holes. The concluding chapter of the book summarizes Payne's rules for wording questions, with the subtitle of the chapter, "A Concise Check List of 100 Considerations," describing the surveyor's dilemma.

The rules, admonitions, and principles for how to word questions, enumerated in various books and articles, present a mind-boggling array of generally good but often conflicting and confusing directions about how to do it.

There is no shortage of simple admonitions on what to do and what not to do. For example:

- Use simple words
- Do not be vague
- Keep it short
- Be specific
- Do not talk down to respondents
- Avoid bias
- Avoid objectionable questions
- Do not be too specific
- Avoid hypothetical questions

The problem is that these “how to do it” rules often get in one another’s way. “Use simple words” is usually good advice, but frequently interferes with advice to “keep it short.” It is interesting, for example, to note that Payne’s 41st version of the question mentioned previously expanded from only eight words to 28 as difficult words were changed to simple ones. Using simple words also increases the risk of talking down to the respondents rather than communicating with them at the same level. The well-founded advice not to be vague often produces questions that are too specific. The advice to keep questions from being too direct and therefore objectional sometimes results in not heeding the advice to avoid hypothetical questions. In addition, although biased wording is certainly to be avoided, it is precisely such questions that may be required for building many kinds of attitude scales.

The reason that seemingly good advice, taken literally, may turn out to be bad advice is that questions are not written in the abstract. Writing questions for a particular questionnaire means constructing them for a particular population, a particular purpose, and placement next to another particular question. Words that are too difficult for use with some populations may be perfectly acceptable for others. A question that is fairly vague may satisfy the exploratory objectives of one survey, but not the analytic ones of another. A question that makes little sense by itself may be quite clear when asked after the ones that precede it in the questionnaire. A list of admonitions, no matter how well intended, cannot be considered absolute. With these cautions in mind, the remainder of this chapter is devoted to principles I have found useful in making the wording and structural changes necessary for turning an initial draft of respondent queries into acceptable survey questions.

PRINCIPLE 2.1: CHOOSE SIMPLE OVER SPECIALIZED WORDS.

We begin by trying to find synonyms that are likely to be understood by more people, and substituting the first word for the second.

tiredexhausted
 honestcandid
 most importanttop priority
 free timeleisure
 workemployment
 bravecourageous
 correctrectify

When a word exceeds six or seven letters, chances are that a shorter and more easily understood word can be substituted, although it should not automatically be assumed that all shorter words are acceptable. For example, it would not be advisable to substitute “deter” for “discourage.”

Next, we focus on words that can be simplified only by using combinations of shorter words:

people who live hereoccupants of this household
 your answersyour responses to this questionnaire
 what you do after schoolpost-school extracurricular activities
 job concernswork-related employment issues
 area of the countysubnational region

In addition, we search for specialized words or abbreviations that are commonplace for the survey sponsor, but require some translation for survey respondents. Such terms are especially prevalent in government surveys. For example:

Oil inventory questionnaireForm 822
 Bureau of Labor StatisticsBLS

Sometimes it is unnecessary to find substitutes for what appear to be difficult words. Virtually all occupational groups share a particular vocabulary that is not understood by outsiders. The use of such a vocabulary facilitates efficient communication, and the use of simpler words would only confuse matters. In a survey of city planners it seems quite reasonable to talk about “annexation” instead of “an addition.” Similarly, in a survey of physicians it seems reasonable to talk about “pharmaceutical companies” instead of “companies that sell medicines.” To do otherwise may even suggest a lack of knowledge and understanding of the topic of the survey.

However, the fact remains that people who write questionnaires are far more likely to overestimate than underestimate the knowledge and vocabulary of respondents. Thus, when in doubt, it is prudent to use the simpler of the available alternatives. A thesaurus is an indispensable tool for finding synonyms that might be used. In addition, there is no substitute for asking

someone with less education than the survey designer to go through a questionnaire and identify words that are confusing. Pretesting with actual respondents is also very important, because it helps the surveyor identify the commonly shared vocabulary of the study population.

PRINCIPLE 2.2: CHOOSE AS FEW WORDS AS POSSIBLE TO POSE THE QUESTION.

At first glance this goal may seem to contradict the frequent necessity of using several words as a substitute for a single, more complex word. That is the reason I place this goal second, rather than first. Initially, we must be sure that any words we choose are understood by virtually all respondents. Having done that, we attempt to keep questions short. Consider the following example tested for a U.S. Census questionnaire:

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 cognitive test of this question resulted in one respondent appearing somewhat embarrassed. He then said, "I don't have any idea how many people live in the United States." As a result of this and other interviews, the well-intentioned second sentence that explained the importance of the inclusion/exclusion definition was removed (Dillman and Allen, 1995).

The problem with long questions stems from the fact that when people read questions they attempt to be efficient, thus giving uneven attention to each word. Important words get missed and unimportant ones sometimes receive undue emphasis, especially when the questionnaire language is a second language for some of the respondents.

Figure 2.5 shows a question in which the answer categories are given in the stem of the question and listed separately below. Such redundancy across many questions is a particularly strong indicator to respondents that it is okay to skip words, and may result in the rest of the sentence also being unevenly read. Redundancy can be eliminated easily by not including the answer choices as part of the question stem. However, the problematic version would be appropriate in an interview questionnaire when only the question stem would be read to each respondent.

Figure 2.5 Encouraging even-reading through reduction of redundancy while maintaining a complete sentence format.

Problem:

17. Do you strongly favor, somewhat favor, somewhat oppose, strongly oppose, or have no opinion on whether advertisers should be required to have advertising aimed at children approved by a national board?
- ☐ Strongly favor
 - ☐ Somewhat favor
 - ☐ Somewhat oppose
 - ☐ Strongly oppose
 - ☐ No opinion

A revision:

17. Which of the following best describes the extent to which you favor or oppose requiring advertisers to have advertising aimed at children approved by a national review board?
- ☐ Strongly favor
 - ☐ Somewhat favor
 - ☐ Somewhat oppose
 - ☐ Strongly oppose
 - ☐ No opinion
-

PRINCIPLE 2.3: USE COMPLETE SENTENCES TO ASK QUESTIONS.

It is tempting to meet the goal of minimizing words by using incomplete sentences for surveys. It is true that few people will misunderstand, “Your name,” or even “Age.” However, the series of questions in Figure 2.6 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 “U.S.” for *county*, a word which is only one letter different from *country*. Writing each question as a complete sentence would have helped solve both problems. In addition, county is changed to Idaho County in order to minimize the possibility of listing the United States as the respondent’s county of residence.

Government-sponsored business surveys are some of the most frequent violators of this principle, a topic discussed in more detail in Chapter 10.

PRINCIPLE 2.4: AVOID VAGUE QUANTIFIERS WHEN MORE PRECISE ESTIMATES CAN BE OBTAINED.

The ease of using the same categories for many different kinds of questions has led to a tendency to use vague quantifiers when more precise an-

Figure 2.6 Ask questions as complete sentences.**Problem:****30. Number of years lived in Idaho**

_____ Years

31. Your city or town

_____ City or Town

32. Your county

_____ County

A revision:**30. How many years have you lived in Idaho?**

_____ Years

31. In what city or town do you live?

_____ City or Town

32. In what Idaho county do you live?

_____ Idaho County

swers can easily be obtained. Consider, for example, the question about attendance at religious services in Figure 2.7. Although it is likely that an answer to this question can easily be provided, enormous variation may exist in what respondents mean by their answers. In some religions, “regularly” implies once a week, whereas in others it may imply several times daily.

Also, the standard for regular attendance within religions may vary among respondents. For example, among adherents to a Christian religion, one person may think of regularly as implying at least once or twice a week, whereas another member of the same religion may think that annual attendance at holiday services, such as Christmas and Easter, implies regular attendance. Changing answer categories to numerical amounts, as shown in the revision, eliminates the possibility of widely varied interpretations for regular or occasional attendance.

Figure 2.7 Avoid vague quantifiers when more precise estimates can be obtained.*Problem:*

26. How often did you attend religious services during the past year?

- ☐ Never
- ☐ Rarely
- ☐ Occasionally
- ☐ Regularly

A revision:

26. How often did you attend religious services during the past year?

- ☐ Not at all
- ☐ A few times
- ☐ About once a month
- ☐ Two to three times a month
- ☐ About once a week
- ☐ More than once a week

PRINCIPLE 2.5: AVOID SPECIFICITY THAT EXCEEDS THE RESPONDENT'S POTENTIAL FOR HAVING AN ACCURATE, READY-MADE ANSWER.

Nearly as troublesome as too much vagueness is the request for too much specificity. Figure 2.8 asks respondents to list the number of books they have read for leisure during the past year. Most people will have a reasonable idea of how many books they have read, but cannot list a precise number for such a long period of time. As a result, some people may simply make a reasonable guess. It is also likely that the open-ended question will elicit a high-item nonresponse because the respondent is not able to offer a precise answer. One solution to this problem is to provide answer categories, as shown in the first revision.

However, because people sometimes respond to the position of the categories as well as the category labels, great care should be taken to determine whether categories are needed and which ones are appropriate. As discussed at the beginning of this chapter, widely varied estimates of the number of hours students reported studying were obtained by varying the size of categories. Thus, people who think they read more than most people might pick the top categories regardless of the labels. Selecting categories on the basis of a pretest and other known characteristics of a population may help the surveyor make the middle categories correspond to the average numbers of books read by the survey population. This example illustrates one of the significant problems associated with simply taking question and answer choices from a survey of one population and using them for another.

Figure 2.8 Avoid specificity that exceeds respondent's potential for having an accurate ready-made answer.

Problem:

12. About how many books have you read for leisure during the past year?

_____ Number of books

A revision:

12. About how many books have you read for leisure during the past year?

- ☐ None
- ☐ 1-2
- ☐ 3-5
- ☐ 6-10
- ☐ 11 or more

Another revision with more appropriate categories for a population of heavy readers:

12. About how many books have you read for leisure during the past year?

- ☐ less than 10
- ☐ 11-25
- ☐ 26-50
- ☐ 51-75
- ☐ 76 or more

PRINCIPLE 2.6: USE EQUAL NUMBERS OF POSITIVE AND NEGATIVE CATEGORIES FOR SCALAR QUESTIONS.

The fact that respondents draw information from the number of categories as well as from labels means that the midpoint for number of categories can easily be interpreted as the neutral point. Figure 2.9 shows an example drawn from a customer satisfaction survey in which three positive (satisfied) categories were used but only one negative (dissatisfied) category was used. Thus, the visual midpoint on the scale became “somewhat satisfied.”

When a recommendation was made to change this format, it was met with objections from the survey sponsor. She pointed out that far more people who responded were satisfied than dissatisfied. Further, she felt that the most important goal of the survey was to obtain significant gradations of satisfaction, a situation similar to that which prompted development of the delighted—terrible scale discussed earlier in this chapter. The solution in this case was simply to add additional categories in a second revision, which then

Figure 2.9 Use equal numbers of positive and negative categories for scalar questions.

Problem:

25. How satisfied were you with the service you received when you bought your air conditioner?

- ☐ Completely satisfied
- ☐ Mostly satisfied
- ☐ Somewhat satisfied
- ☐ Neither satisfied nor dissatisfied
- ☐ Dissatisfied

A revision:

25. How satisfied were you with the service you received when you bought your air conditioner?

- ☐ Completely satisfied
- ☐ Somewhat satisfied
- ☐ Neither satisfied nor dissatisfied
- ☐ Somewhat dissatisfied
- ☐ Completely dissatisfied

Another revision to maintain gradations of satisfaction:

25. How satisfied were you with the service you received when you bought your air conditioner?

- ☐ Completely satisfied
 - ☐ Mostly satisfied
 - ☐ Somewhat satisfied
 - ☐ Neither satisfied nor dissatisfied
 - ☐ Somewhat dissatisfied
 - ☐ Mostly dissatisfied
 - ☐ Completely dissatisfied
-

allowed the three levels of satisfaction to be distinguished from one another while the scale remained balanced.

PRINCIPLE 2.7: DISTINGUISH UNDECIDED FROM NEUTRAL BY PLACEMENT AT THE END OF THE SCALE.

Sometimes attitudinal questions are posed without giving the option of a neutral opinion or no opinion at all. In other cases, researchers prefer to al-

Figure 2.10 Distinguish undecided from neutral by placing this category at end of scale.

Problem:

8. To what extent do you agree or disagree with this statement: “Living in rural places is better for a person’s emotional health.”

- ☐ Strongly agree
- ☐ Somewhat agree
- ☐ Undecided
- ☐ Somewhat disagree
- ☐ Strongly disagree

A revision that distinguishes meaning of undecided from neutral:

8. To what extent do you agree or disagree with this statement: “Living in rural places is better for a person’s emotional health.”

- ☐ Strongly agree
- ☐ Somewhat agree
- ☐ Somewhat disagree
- ☐ Strongly disagree
- ☐ Undecided

Another revision to allow both neutral and no opinion responses:

8. To what extent do you agree or disagree with this statement: “Living in rural places is better for a person’s emotional health.”

- ☐ Strongly agree
 - ☐ Somewhat agree
 - ☐ Neither agree nor disagree
 - ☐ Somewhat disagree
 - ☐ Strongly disagree
 - ☐ No opinion
-

low for no opinion in order to distinguish true opinion holders from those who are being “forced” to choose on a topic to which they have given little or no thought.

If an “undecided” category is offered to respondents, it makes a great deal of difference where that category is placed. An experiment by Willits and Jan-ota (1996) compared the first and second alternatives shown in Figure 2.10 for presenting an undecided choice to respondents. When undecided was placed in the middle, respondents were consistently more likely (across 13 items) to

use that category. On average, the percentage of respondents using it more than doubled, from 5 to 13%. When the answer was placed in the last position (first revision), respondents were more likely to select one of the directional opinion categories. Thus, it appeared that when placed in this end position, “undecided” responses were being separated from neutral opinions. This revision appears to pick up some of the desirable features of both of the other scales, providing less of an invitation to avoid a directional response while still providing an opportunity for people who have no opinion to say so.

Another alternative is provided by the second revision, offered in Figure 2.10. It allows respondents to report being neutral on the issue, but also allows for having no opinion at all. The middle and final categories are carefully worded; “Neither agree nor disagree” and “No opinion” are used to make the meaning of choices as clear as possible.

PRINCIPLE 2.8: AVOID BIAS FROM UNEQUAL COMPARISONS.

Closed-ended questions with unordered categories may become unbalanced, albeit for somewhat different reasons than in the case of scalar questions as discussed under Principle 2.6. In these questions, the topic of the question appears in the answer choice. Consider the wording of the question in Figure 2.11, designed to find out whom respondents think is most responsible for outbreaks of violence in schools. The terms “irresponsible” places a value connotation on the first category that is not present in the other choices. Although it is unclear whether unbalancing questions in this way leads to more or less frequent selection of such categories (Schuman and Presser, 1981), the credibility of responses to such questions is inevitably open to question.

The difficulty of revising such questions is that true balance may be extremely difficult to achieve. The first revision uses less emotionally charged words, “The way children are raised by parents,” but results in a category with many more words than the school and television choices. The last two categories could be made more specific by mentioning school discipline policies and violent television programs, but it is unclear without extensive pretesting whether that would improve or detract from the balance. The challenge of achieving balance on such closed-ended questions often leads to reducing choices to simple nouns (parents, schools, television), a solution that also increases the vagueness of the categories. One might, for example, wonder what aspect of television is being referenced: its use in schools, how much television students watch, or the content of the programming. Still another revision that might be considered is to completely restructure the question, converting to a closed-ended ordered question with a detailed concept presented in the stem of the question, as shown in the final revision offered in Figure 2.11.

Figure 2.11 Avoid bias from unequal comparisons.**Problem:**

19. Which one of the following do you feel is most responsible for recent outbreaks of violence in America's schools?

- ☐ Irresponsible parents
- ☐ School policies
- ☐ Television programs

A revision:

19. Which one of the following do you feel is most responsible for recent outbreaks of violence in America's schools?

- ☐ The way children are raised by parents
- ☐ School policies
- ☐ Television programs

Another revision (simplest form):

19. Which one of the following do you feel is most responsible for recent outbreaks of violence in America's schools?

- ☐ Parents
- ☐ Schools
- ☐ Television

Still another revision (retaining more complex descriptions):

19. To what extent do you feel that the way children are raised by parents is responsible for recent outbreaks of violence in America's schools?

- ☐ Completely responsible
- ☐ Mostly responsible
- ☐ Somewhat responsible
- ☐ Not at all responsible

And so forth for the remaining concepts.

PRINCIPLE 2.9: STATE BOTH SIDES OF ATTITUDE SCALES IN THE QUESTION STEMS.

It is tempting to reduce the number of words in questions by mentioning only one side of an attitude scale when posing a question. For example, the question in Figure 2.12 asks the extent to which people agree with a statement, leaving out any mention of disagreement. Structured in this way, the question may encourage people to think of a scale ranging from "not at all" to "strongly

Figure 2.12 State both sides of attitude scales in question stems.**Problem:**

14. To what extent do you agree with this statement: “It’s easier for people to find work in this community than it was about one year ago.”

- ☐ Strongly agree
- ☐ Somewhat agree
- ☐ Somewhat disagree
- ☐ Strongly disagree

A revision:

14. To what extent do you agree or disagree with this statement: “It’s easier for people to find work in this community than it was about one year ago.”

- ☐ Strongly agree
- ☐ Somewhat agree
- ☐ Somewhat disagree
- ☐ Strongly disagree

Another revision, to avoid acquiescence:

14. Do you feel it is easier, the same, or more difficult for people to find work in this community than it was about one year ago?

- ☐ Easier
- ☐ The same
- ☐ More difficult

agree.” Substituting “agree or disagree” conveys to the respondent that the scale has a greater range. It also lets the respondent know that disagreement is an acceptable answer. Mentioning both sides of a scale is even more important when the terminology of a scale changes from one end to another, as it would for a scale that goes from “strongly favor” to “strongly oppose.”

However, one of the shortcomings of agree/disagree is that there is a cultural tendency in many societies for people to agree rather than disagree, or to acquiesce in their responses. Consequently, the second revision, which maintains the use of closed-ended ordered categories but changes to an “easier, the same, more difficult” format, would seem to be a better solution.

PRINCIPLE 2.10: ELIMINATE CHECK-ALL-THAT-APPLY QUESTION FORMATS TO REDUCE PRIMACY EFFECTS.

In an effort to reduce respondent burden, queries containing unordered response categories are sometimes structured as “check-all-that-apply” items,

Figure 2.13 Eliminate the use of check-all-that-apply question formats.**Problem:**

27. Which of the following characteristics would you like to see your child develop as he/she grows up? Please check all that apply.

- ☐ An interest in sports
- ☐ An interest in music
- ☐ An appreciation of art
- ☐ An interest in science
- ☐ An interest in business

A revision:

27. Is each of the following characteristics one that you would like for your child to develop as he/she grows up?

An interest in sports	Like	Not like
An interest in music	Like	Not like
<i>Etc.</i>		

Another revision:

27. To what extent would you like or not like for your child to develop each of these characteristics as he/she grows up?

An interest	Very Great	To Some	A Small	Not
in sports	Extent	Extent	Extent	At All
An interest	Very Great	To Some	A Small	Not
in music	Extent	Extent	Extent	At All
<i>Etc.</i>				

Still another revision:

27. Which of these qualities would you most like for your child to develop when he/she grows up? Put a 1 in the box for most like, 2 in box for second most like, and go on until you have ranked all five qualities.

An interest in sports	<input type="checkbox"/>
An interest in music	<input type="checkbox"/>
<i>Etc.</i>	

as shown in Figure 2.13. Questions of this nature do not meet our criterion of being a survey question; that is, requiring a response to each stimulus. It has been observed that respondents tend to “satisfice” in answering many types of survey questions (Krosnick, Narayan, and Smith, 1996). This type of question is no exception, as respondents begin checking answers and go down the list until they feel they have provided a satisfactory answer. Although certain respondents will read and consider each answer, others will not. As respon-

dents proceed down the list they may feel that “enough” has been done. Thus, the items listed first are more likely to be checked.

In recent years, considerable research has emerged suggesting that respondents to self-administered surveys are more likely to exhibit a primacy effect; that is, a tendency to select from among the first answers presented. Interview respondents are more likely to exhibit a recency effect; that is, a tendency to select from among the last answers mentioned (Krosnick and Alwin, 1987; Schuman and Presser, 1981). Data to support this hypothesis are inconsistent. In particular, an examination of 84 separate mail and telephone experiments conducted by nine investigators showed no systematic effects in either direction (Dillman et al., 1995). Recency effects were nearly as likely to occur in mail surveys as telephone, and vice versa.

The original argument for the existence of primacy effects was based on the presentation of a list of items given to respondents in an interview survey; respondents were more likely to pick answers from items early in the list. I suspect that primacy effects, to the extent that they occur, are most likely to happen when self-administered respondents exercise a satisficing behavior and stop reading the list, as they seem invited to do in “check-all-that-apply” items. Consequently, our first revision shown in Figure 2.13 simply provides a like/not like choice for each item. Two other revisions are also shown, both of which, in addition to requiring an answer for each item, increase the amount of information obtained from each respondent. The second revision uses a rating scale for each item, and the third revision uses a ranking question. Although the latter option is workable for five items, it is less acceptable as the number of items increases. Any of the three revisions is much preferred over the original format.

PRINCIPLE 2.11: DEVELOP RESPONSE CATEGORIES THAT ARE MUTUALLY EXCLUSIVE.

The image that often comes to mind when one discusses the lack of mutually exclusive categories is that a clerical error has been made, as in age categories like these: 35 or less, 35–50, 50–65, and 65 and over. Although the overlap is minor, it can be very annoying to people who happen to fall on the line.

In some instances we have seen deliberate, though minor, overlap created when attempting to ease the task of the respondent. For example, income categories of less than \$15,000, \$15,000 to \$19,999, \$20,000 to \$29,999, \$30,000 to \$39,999, and greater have sometimes been changed so that the latter categories read \$15,000 to \$20,000, \$20,000 to \$30,000, etc. The latter categories are much easier for interviewers to read than the former ones, and make virtually no difference in the ability of respondents to choose an appropriate category. The sensitivity of reporting income is such that the gain from ease of reading (especially for interviews) probably outweighs the mutual exclusivity concern about the few people whose incomes happen to fall exactly on the dividing line.

Our main concern with mutual exclusivity is when it is used for response categories of survey questions, where there is considerable likelihood that its presence will go unnoticed. Figure 2.14 shows a question in which respondents have been asked to choose one answer to the question of how they learned about a disaster, but choices combine sources as well as location. The revision simply breaks the question into two parts, one about source and the other about location.

Mutual exclusivity provides yet another reason for avoiding check-all-that-apply question formats. If people can check more than one category in the example used in Figure 2.15, it might be reasoned that mutual exclusivity is not a problem. However, Israel and Taylor (1990) have reported experimental results that provide further evidence that check-all-that-apply questions create difficulties for surveys. The check-all-that-apply question about forages fed to cattle in winter months (Figure 2.15) was asked in a survey of Florida beef producers. The answer categories for this question were presented in different orders with quite different results.

Switching the “native range” to fourth on the list decreased the proportion

Figure 2.14 Develop response categories which are mutually exclusive.

Problem:

7. **From which one of these sources did you first learn about the tornado in Derby?**

- ☐ Radio
- ☐ Television
- ☐ Someone at work
- ☐ While at home
- ☐ While traveling to work

A revision:

7. **From which one of these sources did you first hear about the tornado in Derby?**

- ☐ Radio
- ☐ Television
- ☐ Another person

8. **Where were you when you first heard about it?**

- ☐ At work
 - ☐ At home
 - ☐ Traveling to work
 - ☐ Somewhere else
-

Figure 2.15 Results from check-all-that-apply question that stems from the lack of mutual exclusivity (Israel and Taylor, 1990).

Question: Which of the following forages are used during the winter months to feed your cattle? (Check all that apply.)

First Order ▼	Category ▼	Percent Selected ▼	Second Order ▼	Percent Selected ▼
1	Native range	70%	(4)	30%
2	Deferred grazing (save pasture for fall and winter)	37	(2)	48
3	Hay	84	(5)	79
4	Silage	1	(1)	2
5	Winter pasture	29	(3)	36
6	Other	14	(6)	15

checking that answer from 70% to only 30%, more than a 40% drop. Winter pasture and deferred grazing, which were moved to the third and second positions respectively, both experienced increases. The likely reason for this change is that “native range” is such a broad category that when listed in the first position, respondents who would otherwise have chosen winter pasture or deferred grazing picked it and then mentally subtracted that from the answers they would have otherwise given. Silage, on the other hand, which was switched from fourth to first position, obtained the same proportion of answers both times. This type of livestock feed is distinctively different from any of the other pasture choices.

Another question in this survey that was sponsored by the Agricultural Extension Service asked respondents to check information sources on a list to indicate where they obtained information about beef production and management (Israel and Taylor, 1990). In response, 59% checked “extension agent” when it was presented first on the list versus only 41% when it was mentioned in third position behind “other cattlemen” and “fertilizer and other salesmen.” We suspect that the social desirability of indicating Extension (the survey sponsor) as a source of information encouraged this response. It is also possible that the extension agent may have presented information at a meeting attended by other cattlemen and salesmen. The work by Israel and Taylor (1990) underscores the importance of being alert to mutual exclusivity concerns.

PRINCIPLE 2.12: USE COGNITIVE DESIGN TECHNIQUES TO IMPROVE RECALL.

One of the difficulties with many self-administered surveys is that respondents answer questions quite rapidly, spending as little time as possible

deciding what answer to choose. Although it is sometimes suggested that an advantage of self-administered questionnaires is that respondents can take their time in completing them and therefore give more thoughtful answers, we know of no evidence to support that as being the typical response behavior (Dillman and Tarnai, 1991).

In addition, Jobe and Mingay (1989) have reported for interview surveys that recall accuracy on visits to health providers can be improved from 41% to 53% by asking about details of the visit. For example, people can be asked who the health care provider was and how they got to the office. Respondents who are asked to recall such visits over a specified period of time, such as the last six months, can also be asked to reconstruct a calendar of important events in their life (e.g., vacations, weddings, etc.) over this period of time before asking about visits to their doctor.

Figure 2.16 shows a quick recall format and cognitively-designed alternative for reporting seat belt use in a survey conducted shortly after a law was passed in Washington state requiring seat belt use, when it was reasoned that people might forget to wear seat belts when starting on a trip in their vehicle, but buckle-up along the way (Dillman and Tarnai, 1991). Use of this cognitively-designed format produced significantly different responses than did the quick recall format; moving people away from a generalized response about the proportion of time the respondent typically wore a seat belt was therefore thought to improve the accuracy of responses.

Although I believe that a cognitive recall sequence of questions has much potential for improving the accuracy of people's responses to survey questions, its use for all questions would result in increased length of questionnaires, not to mention the greater tediousness of responding. Thus, I tend to reserve its use for the most important survey questions—those for which the most accurate estimates of behavior must be made.

PRINCIPLE 2.13: PROVIDE APPROPRIATE TIME REFERENTS.

Respondents are frequently asked to report whether or how often they have engaged in a particular behavior during a recent period of months or years. There are several distinct problems associated with the use of time referents in surveys. First, memory tends to fade and people usually do not categorize information by precise month or year periods. Even if they try very hard to remember, estimating behaviors over a period of the last three years (as suggested by the first example in Figure 2.17) may be impossible. Some people may end up not being able to make a meaningful distinction. One solution, which may or may not be acceptable for some studies, is to shorten the time period and use a cognitive recall set of questions prior to asking for an answer.

Another type of problem is when something is too regular and mundane in life, such as studying, watching television, or eating away from home, as

Figure 2.16 Use cognitive design techniques to improve respondent recall.**Problem:**

10. Next, we would like to ask about the most recent time you drove or rode anywhere in an automobile or other vehicle such as a pickup or van. During this most recent ride, would you say that your seatbelt was fastened...
- ☐ All the time; that is, every minute the car was moving
 - ☐ Almost all the time
 - ☐ Most of the time
 - ☐ About half the time
 - ☐ Less than half the time
 - ☐ Not at all during the time the vehicle was moving

(Question repeated for second and third most recent rides)

A revision:

10. Next, we would like to ask you to please think about the last three times you drove or rode in a car or vehicle such as a pickup or van.

First, when was the last time you drove or rode in a motor vehicle?

- ☐ Today
 - ☐ Yesterday
 - ☐ Sometime before that
11. Could you tell us generally where this most recent trip began and where it ended?
12. About how long was this trip?
- ☐ Less than a mile
 - ☐ One to five miles
 - ☐ Longer
13. During this trip were you the...
- ☐ Driver
 - ☐ A front seat passenger
 - ☐ A back seat passenger
 - ☐ Other
14. During this trip, would you say your seatbelt was fastened...
- ☐ All the time; that is, every minute the car was moving
 - ☐ Almost all the time
 - ☐ Most of the time
 - ☐ About half the time
 - ☐ Less than half the time
 - ☐ Not at all during the time the vehicle was moving

(Repeat entire sequence for second and third most recent trips)

asked about in the second set of questions. A potential solution in these situations is to switch from asking people for a count to asking for a general estimate of how many times they eat a meal away from home in an average week. Faced with this problem, we have sometimes observed questionnaires in which intervals such as “the last three days” are used. This choice is usually unacceptable inasmuch as some activities, such as eating away from home,

Figure 2.17 Provide appropriate time referents.**Problem:**

4. How many times in the last three years have you gone to see a doctor or other health care practitioner?

_____ Number of visits in last three years

A revision with shorter time referent:

4. How many times in the last six months have you gone to see a doctor or other health care practitioner?

_____ Number of visits in last six months

Problem:

11. How many times in last six months have you eaten away from home?

_____ Number of times in last six months

A revision that uses an estimation strategy:

11. On average, how many times per week do you eat a meal away from home?

_____ Average number of times per week

Problem:

22. How many times have you played golf so far this year?

_____ Number of times played this year

Revision to take into account different return dates for questionnaires:

22. During the last calendar year, 1998, about how many times did you play golf?

_____ Number of times played in 1998

may differ greatly by day of week, so that different responses may be obtained depending upon whether one completes a questionnaire on a Monday, following a weekend, or a Thursday, following three weekdays.

The latter problem has some similarity to the final problem question presented in Figure 2.17, in which respondents are asked how many times they have played golf this year. The problem with this question in a mail survey is that responses are likely to be higher later in the year. The revision for this question asks the number of times played in the previous year. We chose a recall for the previous year rather than a general estimation strategy because

people who play golf often keep track of the number of times they have played, or play with sufficient regularity (every Thursday afternoon from June through August) that they can provide fairly accurate estimates of times played in the previous year. The phrase, “your best estimate is fine” is sometimes added in order to limit item nonresponse. Another possibility is to prelist categories, as shown earlier in Figure 2.8.

PRINCIPLE 2.14: BE SURE EACH QUESTION IS TECHNICALLY ACCURATE.

Few features of a questionnaire reflect as negatively on a surveyor’s credibility as a technical error visible to certain respondents or a survey sponsor. Such errors range from the misspelling of a political candidate’s name to providing an incorrect name for an organization asked about in a survey question. Such errors are very easy to make and few can claim never to have made them.

Some errors are more subtle, such as the one in Figure 2.18, where police are identified as being responsible for catching and fining traffic violators. They only do the former; the judicial system, and not the police, is responsible for determining and administering fines. Sometimes as attitude and opinion questions are drafted and redrafted, technical issues such as this one creep into surveys where they have only a minor role, but nonetheless bring into question the surveyor’s competence.

I mention this concern as a separate issue because most surveyors, at some time in their careers, design surveys on topics about which they know little.

Figure 2.18 Besure that each question is technically accurate.

Problem:

31. Another activity of the police department is the catching and fining of traffic violators. Should this activity receive greater, about the same, or less emphasis than at present?

- ☐ Greater
- ☐ The same
- ☐ Less

A revision:

31. Another activity of the police department is the catching of traffic violators. Should this activity receive greater, about the same, or less emphasis than at present?

- ☐ Greater
 - ☐ The same
 - ☐ Less
-

Checking questionnaires for technical accuracy becomes a very important aspect of pretesting, which I discuss in the next chapter.

PRINCIPLE 2.15: CHOOSE QUESTION WORDINGS THAT ALLOW ESSENTIAL COMPARISONS TO BE MADE WITH PREVIOUSLY COLLECTED DATA.

Comparison of survey results with previously collected data often represents a major survey objective. Sometimes the comparison is being done with a previous survey of the same population in order to measure change. Another likely comparison is with U.S. Census data for the same geographical area. Figure 2.19 illustrates the dilemma one may face in deciding whether to replicate the wording of previous surveys. This question asks whether people rent or own their homes, and is much like the structure we have seen used in hundreds of surveys. The revision is far wordier, specifies type of home to some degree (house, apartment, or mobile home), and is shown here exactly as used in the 2000 U.S. Census.

This example also illustrates one of the difficulties with many questions asked in national surveys. The wording is determined through a long process involving input from many stake-holders. As a result, the wording may be somewhat more cumbersome than many surveyors feel is needed for their particular survey. Similar issues exist with respect to the racial and Spanish/Hispanic/Latino questions to be asked in the 2000 Census, which will

Figure 2.19 Choose question wording that allows essential comparisons to be made with previously collected data.

Problem:

13. Do you own or rent the home in which you live?

- ☐ Own
- ☐ Rent

A revision:

13. Is this house, apartment, or mobile home:

- ☐ Owned by you or someone in this household with a mortgage or loan?
 - ☐ Owned by you or someone in this household free and clear (without a mortgage or loan)?
 - ☐ Rented for cash rent?
 - ☐ Occupied without payment of cash rent?
-

provide comparison information for much of the next decade. The race question provides 14 categories and allows for multiple selections. The ethnicity question provides four Spanish/Hispanic/Latino categories plus an “other” category. For policy surveys in which it is crucial to be able to demonstrate similarity or differences from Census data, use of the same questions may be desirable. For other surveys, a different question structure may be quite acceptable, or even more desirable.

PRINCIPLE 2.16: AVOID ASKING RESPONDENTS TO SAY YES IN ORDER TO MEAN NO.

It seems obvious that questions should not include double negatives, thus requiring a respondent to say yes to mean no, as for this question, “Should the city manager not be directly responsible to the mayor?” Yet, such questions are commonly asked in surveys. One of the reasons such questions are so prevalent in surveys 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 2.20. Surveyors are often re-

Figure 2.20 Avoid asking respondents to say yes in order to mean no.

Problem:

5. Do you favor or oppose not allowing the state to raise taxes without approval of 60% of the voters?
- ☐ Favor
- ☐ Oppose

A revision:

5. Do you favor or oppose requiring 60% approval by voters in order to raise state taxes?
- ☐ Favor requiring 60% approval
- ☐ Oppose requiring 60% approval

Another revision (for specific policy situation):

5. 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
-

luctant to pose the question differently than it would 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 what favor and oppose means for purposes of the question, the answer categories specify what favor or oppose means. This wording would seem appropriate during discussion of an issue before it has reached the ballot measure stage. A second revision, indicating that a vote will be taken, specifies the measure exactly as it will appear on the ballot and asks whether respondents are for or against approval of the measure. The switch of categories from favor–oppose to for–against is also an attempt to bring the language of the question more in line with the voting situation.

PRINCIPLE 2.17: AVOID DOUBLE-BARRELED QUESTIONS.

A somewhat similar problem occurs with questions that contain two components which require one answer, but about which a respondent may feel differently. At first glance the question in Figure 2.21 appears to be simply another case of asking respondents to say yes in order to mean no. However, this question about whether people want a swimming pool to be built with lap lanes, but without a winter enclosure, is really asking two separate questions with no opportunity to respond to each part. This question was justified by its proponents because that seemed to be the policy issue facing voters—whether to approve a larger pool with lap lanes that would be too expensive to enclose or a smaller one that could be enclosed for winter use.

The first revision is a conventional one that allows people to respond to each of the issues and is effective in eliminating the inability of people who wanted both lap lanes and a winter enclosure to express their true opinion.

The second revision seemed more preferable in this case. By writing the question with a closed-ended unordered structure, and placing more wording into the answer choices, five clear alternatives could be presented to respondents. However, if the objective of this survey were to find out whether people would vote for or against the first proposal, then the third revision offered in Figure 2.21 might be the best question structure.

PRINCIPLE 2.18: SOFTEN THE IMPACT OF POTENTIALLY OBJECTIONABLE QUESTIONS.

One of the most worrisome aspects of conducting self-administered surveys is that people can skip questions they object to answering, or decide not to answer any more questions. The sources of objections differ, as do the solutions.

Figure 2.21 Avoid double-barreled questions.**Problem:**

16. Should the city build a new swimming pool that includes lanes for swimming laps that is not enclosed for winter use?

- ☐ Yes
☐ No

A revision:

16. Should the city build a new swimming pool that includes lanes for swimming laps?

- ☐ Yes
☐ No

17. Should the city build a new swimming pool that is enclosed for winter use?

- ☐ Yes
☐ No

Another revision:

16. It has been proposed that the city build a new swimming pool that could include or not include lanes for swimming laps and being enclosed for winter use. Which one of the following do you most prefer?

- ☐ I prefer that no pool be built.
☐ I prefer a pool with lanes for swimming laps and winter enclosure.
☐ I prefer a pool without lanes for swimming laps or winter enclosure
☐ I prefer a pool with lanes for swimming laps that is not enclosed for winter use.
☐ I prefer a pool enclosed for winter use without lanes for swimming laps.

Still another revision (for specific policy situation):

16. It has been proposed that the city build a new swimming pool that includes lanes for swimming laps, but will not be enclosed for winter use. If the election were held today, would you vote for or against this proposal?

- ☐ For
☐ Against

Perhaps the one question most likely to elicit negative reactions from respondents is the request for income. The open-ended format presented first in Figure 2.22 is the format most likely to be left unanswered. Not only does it require respondents to record something they are likely to consider no one else's business, but for some people the question is difficult to answer. That is the case for people who do not know their exact income on a yearly basis, and this format provides an easy justification for skipping. Switching from the open-ended format to broad categories, as shown in the revision in Figure 2.22, will reduce item nonresponse. For some studies these categories may be

Figure 2.22 Procedures for softening the impact of potentially objectional questions.

Problem:

38. What was this person's total income from all sources in 1999?

_____ Total income for 1999

A revision:

38. Which of the following broad categories best describes this person's total income from all sources in 1999?

- ☐ \$10,000 or less
 - ☐ \$10,001 to \$20,000
 - ☐ \$20,001 to \$35,000
 - ☐ \$35,001 to \$50,000
 - ☐ \$50,001 to \$100,000
 - ☐ \$100,001 or more
-

Problem:

26. Have you ever shoplifted something from a store?

- ☐ Yes
- ☐ No

A revision:

26. Have you ever taken anything from a store without paying for it?

- ☐ Yes
 - ☐ No
-

(Continued)

too broad, and anything less than an exact number may be considered unacceptable. The U.S. Decennial Census, for example, asks not only for exact total income, but also asks for exact amounts from eight potential sources of income. Nonetheless, the use of categories helps greatly in overcoming the difficulties of obtaining income.

The second question in Figure 2.22 presents a different problem. Here the respondent is being asked whether she has performed an illegal behavior, such as shoplifting. Sometimes questions of this nature can be softened by changing the wording, as also shown in the revision in Figure 2.22. Another

Figure 2.22 (Continued)

Another revision:

26. The questions which follow are being asked to help us understand things that have happened to people over the years, and how their lives have been affected. We really appreciate your help and that of the thousands of others who have been asked to complete this national survey.

Have you ever hit someone with your fist?

- ☐ Yes
☐ No

27. Have you ever taken anything from a store without paying for it?

- ☐ Yes
☐ No
-

Problem:

30. “Most religions are a parasite on society.” Do you agree or disagree?

- ☐ Agree
☐ Disagree

A revision:

30. Here are many different opinions on religions we have heard from others. Please tell us whether you agree or disagree with each of them.

“Most religions try to help members as well as nonmembers.”

- ☐ Agree
☐ Disagree

31. “Most religions are a parasite on society.”

- ☐ Agree
☐ Disagree
-

possibility shown here is to embed the question in a series of items that starts with a reason for answering, then lists other, less objectionable behaviors before asking about the illegal one.

Some questions elicit objections because of the context in which they are presented. The questionnaire is in many respects a conversation between the surveyor and the respondent, and questions that seem particularly nosy, abrupt, harshly asked, or unclear may motivate people to discontinue the conversation. An example shown in the third question in Figure 2.22 is the abrupt statement, “Most religions are a parasite on society. Do you agree or disagree?” Presented in this form, the statement appears to be the opinion of the

surveyor. The revision embeds the statement in a brief context and uses a buffer item, as was done for the revision of the shoplifting question. I consider it important to examine the flow of written items, even reading them out loud to contemplate whether an introductory explanation or change in context will make questions seem less objectionable to respondents.

PRINCIPLE 2.19: AVOID ASKING RESPONDENTS TO MAKE UNNECESSARY CALCULATIONS.

When respondents are asked to report percentages, they are required to make an implicit calculation. Some will take the time to come up with the specific numbers and do the appropriate mathematical calculation, but others will simply make an estimate. To improve accuracy it is usually desirable to ask for the numerical information and reduce the burden on respondents by not asking them to do the calculations. An example is the question about percentage of nights spent away from home for business reasons, as shown in Figure 2.23. Separation into questions about total number of nights away from home, followed by one about number away on business, solves the problem by leaving the calculations to the survey analyst.

An exception to this principle is when the respondent is unlikely to know the base numbers, but might be able to make a reasonable estimate directly. For example, a worker in a large corporation might be able to offer an opinion of what proportion of men usually wear neckties to work every day. Having to determine first the number of employees and then the number who

Figure 2.23 Avoid asking respondents to make unnecessary calculations.

Problem:

- 17. What percent of the nights spent away from home on trips during 1998 were for business reasons?**

_____ Percent of nights away from home spent on business trips

A revision:

- 17. How many nights did you spend away from home on trips of any kind during 1998?**

_____ Number of nights away from home in 1998

- 18. How many of these nights away from home were because of business trips?**

_____ Number of nights away from home in 1998 on business trips.

wear neck ties would be unnecessarily laborious, and an invitation to item-nonresponse.

CONCLUSION

In this chapter a formidable gauntlet of concerns has been presented that must be addressed by the would-be writer of questions for self-administered surveys. Eight criteria have been posed for assessing potential survey questions. Three fundamentally different ways of structuring questions were presented, with advice on when one or another might be most appropriate. Finally, 19 principles for reconciling question wording with the assessment criteria and structures were described and illustrated.

Blind adherence to these criteria and principles is hardly a guarantee of success in writing questions. We have noted that principles sometimes come into conflict, and different survey situations exhibit different needs. Reconciling the issues discussed in this chapter with one's survey situation has sometimes been described as the art of asking survey questions (e.g., Payne, 1951). The information discussed here provides fundamental tools for successfully accomplishing this reconciliation.

However, the outcome of the question writing process must be viewed as incomplete. The complete stimulus presented by any question in a self-administered survey also depends upon the order in which questions are seen and answered, and whether all of the words are read by each respondent. In contrast to interview surveys, there is no guarantee that respondents will read each question in the manner intended by the questionnaire writer. However, careful design and layout of questionnaires, the topic to which I turn next, can greatly increase that likelihood.

CHAPTER 3

Constructing the Questionnaire

SELF-ADMINISTERED QUESTIONNAIRES can be constructed in ways that make them easy to understand and answer. However, they are sometimes designed with features that result in questions being misread and items, or even pages, being skipped altogether. Often these problems have little to do with question wording, as discussed in Chapter 2. Instead, they are the result of unfortunate decisions about questionnaire format, question order, and the appearance of individual pages. Examples encountered through many years of providing consultation to study sponsors include:

- In a university faculty-sponsored survey, the questions were listed in the order they were developed. Respondents were therefore required to switch back and forth to answer questions on similar topics at different places in the questionnaire.
- A several-part income question was placed on page one of a questionnaire for a study about the purchase of health care services. The sponsor considered it the most important question in the survey, and was concerned that including this question on the last page, as I recommended (to reduce perceived costs), would result in its being skipped.
- For a nationwide survey of the general public, a proposed questionnaire was printed on one sheet of paper and folded first to form six pages, and once again so that it would fit into a half-size envelope. Observation of people who tried to fill out the questionnaire revealed that some of them unknowingly skipped pages and others made refolding mistakes so that the return address on the refolded questionnaire did not show through the window of the return envelope.
- In a well-funded survey of electric utilities a variety of graphical forms were intermingled, including font changes, shaded and unshaded backgrounds, boxes that isolated certain text, and combinations of bold,

italicized and underlined words within individual sentences. This visual variation made it difficult for respondents to comprehend and respond to each question, the opposite of the writer's reason for using these graphical variations.

- For a study of environmental attitudes and behavior, the type font was reduced to 9 points, the size shown here, in order to get the questions to fit onto two pages. In addition, background questions on occupation and religion were inserted without explanation on page one, between two sets of opinion questions, because it was the only space left where they would fit.

Transforming a list of questions into a questionnaire involves much more than the manipulation of words. It requires that decisions be made about paper size and binding. It also necessitates determining which questions will encourage the recipient to start responding and keep going to the end. In addition, layout choices must be made that involve spacing, size, and brightness considerations. The graphical aspects of layout and design represent critical decisions that constitute a large portion of the discussion in this chapter. Although my focus here is mainly on paper self-administered questionnaires, this chapter lays the foundation for a discussion of the visual aspects of web survey design in Chapter 11.

CRITERIA FOR DESIGN

Individual design decisions made separately from other questionnaire considerations often have negative results for other parts of the survey process. A friend was enticed by a printer to purchase a brightly colored paper that was on sale for the same price as basic white. However, the lack of contrast with the black print made the words difficult to read. A graphical designer encouraged the writer of a questionnaire on historical aspects of housing to use a stylized font that reminded one of colonial America, and it, too, made reading difficult. Repeatedly, I have been told of well-intentioned printers who responded to a request for a booklet questionnaire with, "I can save you a lot of money if you let me duplex regular size sheets of paper ($8\frac{1}{2}'' \times 11''$, page printed front and back) and put a staple in the upper left corner instead of using a booklet binding with two staples in the spine." In other cases, the first pages of questionnaires have been expanded in order to include a letter to the respondent (as a way of saving insertion and printing costs), a lengthy list of instructions on how to answer individual questions, and a glossary. As a consequence of these additions, the first question on one government questionnaire I reviewed did not appear until page four!

It is important to think of questionnaire design as an attempt to achieve two

objectives. One objective is to reduce nonresponse. It has been shown that respondent-friendly questionnaire design can improve response rates, but only to modest degrees (Dillman, Sinclair, and Clark, 1993). Implementation procedures, the topic of the next chapter, are far more powerful as inducers of high response. Nonetheless, small improvements in response can be achieved. More importantly, some research has shown that making a questionnaire respondent-friendly is most likely to improve response among people who are least likely to respond to surveys, and thus help reduce nonresponse error. For example, in a test of the U.S. Decennial Census questionnaires, response was improved by only 2.9 percentage points in areas of the United States with high response to the previous Census versus 7.5 percentage points in areas with low response in 1990 (Dillman, Sinclair, and Clark, 1993). Specific design features such as the choice of the first question can also contribute significantly to the reduction of nonresponse error, by getting people to whom the survey questions do not apply to return their uncompleted questionnaires.

The reduction or avoidance of measurement error is the second objective of good questionnaire design. Poor questionnaire layout can cause questions to be overlooked or bias the offered responses. A respondent-friendly questionnaire is attractive and encourages people to read words in the same order as other respondents read them. People are guided by graphical layout features from the cover page through the last question. A well designed layout prevents items or answer categories from being missed because of their location on the page.

Design features, which are intended to motivate people to respond, focus on all three of the social exchange elements discussed in Chapter 1. Design features are used to improve rewards by making the questionnaire appear interesting and important. Costs are reduced by making the questionnaire easy to manipulate and easy to complete. Trust is encouraged through attention to detail that makes the questionnaire look and seem important. All of the issues addressed in this chapter, from the size and shape of the questionnaire to cognitive pretesting, are important to achieving the best possible questionnaire.

It has been suggested that one advantage of self-administered questionnaires is that people can fill them out at their own speed, taking time to comprehend each question and provide a thoughtful answer. Based upon many cognitive interview studies I have conducted (e.g., Dillman, Carley-Baxter, and Jackson, 1999), and analyses of mode comparison experiments (e.g., Dillman and Tarnai, 1991), I am convinced that many respondents do not read the entire content of questionnaires in a thoughtful way. Respondents take clues from the layout about what must be read and what can safely be ignored, and some respondents skip many words, with the frequent result that questions get misinterpreted. Our best opportunity for achieving clear responses to questions is to keep both the wording and visual appearance of questions

simple. The design concepts and principles that follow are based upon this belief.

Discussion of these concepts is divided into four sections. The first section concerns the physical format of the questionnaire and explains why the booklet format is preferred. It is followed by criteria for deciding how questions should be ordered in the section. The third section, and the longest one in this chapter, introduces knowledge about how people see and process information on printed pages; from this information we can derive principles for deciding on the layout and design of individual questionnaires. The fourth section focuses on only two pages of the questionnaire, the front and back covers. This chapter concludes with a discussion of pretesting, the aim of which is to see whether the goals discussed in the first four sections have been accomplished.

ALTERNATIVE QUESTIONNAIRE FORMATS AND WHY A BOOKLET IS PREFERRED

The negative consequences of certain questionnaire formats are great enough for us to reject them as possibilities. Unacceptable formats include:

- Printing on both sides of sheets of paper with a staple to hold the pages together.
- Printing of pages in a landscape (horizontal) rather than portrait (vertical) orientation.
- Unusual folds; for example, a large single-sheet of paper that unfolds like an accordion or in some other way, as does a typical road map.
- Unusual shapes; for example, square or diamond-shaped pages.

The difficulty with these formats is that each of them is unconventional and respondents must figure out how to handle or comprehend each set of materials at the same time they are attempting to answer the questions. The vertical book or booklet, with pages taller than they are wide, is a standard reading format for most western cultures. In focus groups and cognitive interviews, I have observed people struggling with other formats, sometimes turning pages incorrectly or missing them altogether, much as some people struggle to fold maps regardless of how many times they have done it before (Dillman, Reynolds, and Rockwood, 1991; Dillman and Allen, 1995). Conversely, booklet formats are handled more or less automatically and usually without error. People are familiar with starting on page one and then turning to page two, which appears as the left of two facing pages, and so forth. Additionally, the ease of setting up and printing booklets results in my strong preference for this format. A horizontal orientation or landscape printing is also unconventional and usually results in longer lines of prose, another undesirable fea-

ture. Long lines of prose are more likely to be read unevenly, with the result that important words get skipped.

In the first edition of this book I recommended a slight photo reduction (79%) of typed $8\frac{1}{2}'' \times 11''$ pages to fit in a $6\frac{1}{2}'' \times 8\frac{1}{4}''$ space. The questionnaire was printed on $12\frac{1}{4}'' \times 8\frac{3}{4}''$ sheets of paper folded in the middle and stapled (when more than one sheet was used). My goal was to make the response task look easier to accomplish, in social exchange terms, and to make the booklets fit into monarch (slightly smaller than business stationery) envelopes. This allowed us to mail a 12-page (three sheets of folded paper) questionnaire for one first-class postage stamp.

These recommendations are now outmoded. Computers allow the manipulation of font and line sizes. The proliferation of printing capabilities has had an opposite effect on page sizes. There are significant cost savings associated with using standard sizes of paper. Consequently, although the size recommendation made in 1978 still works fine, it is now less practical. More importantly, I am aware of no experimental evidence that the smaller booklets have a significant influence on response rates. In fact, there may be a packaging appearance advantage to larger-sized mailouts (as discussed in Chapter 4).

Three basic formats are considered acceptable for use with multi-page questionnaires. The first format is a booklet, similar to that recommended in the first edition of this book. Conventional legal size ($8\frac{1}{2}'' \times 14''$) paper is used. When folded as a booklet and stapled along the spine, an $8\frac{1}{2}'' \times 7''$ questionnaire is formed. Nearly all copy centers stock legal size paper so that printing is widely accessible at competitive rates. These booklets can be folded lengthwise to fit into a regular size ($4\frac{1}{8}'' \times 9\frac{1}{2}''$) business stationery envelope. These questionnaires are normally printed with one column of questions per page, as shown in Figure 3.1.

A second format consists of full-sized pages, also printed as a booklet. The questionnaire is printed on $11'' \times 17''$ paper so that when folded, individual pages are of conventional $8\frac{1}{2}'' \times 11''$ dimensions. If this size of paper is used, it is desirable to divide each page into columns so that the span of type being read does not extend entirely across the page, except for items in a series (see Figure 3.2). For most questionnaires of this size, a two-column format allows more questions per page and may contribute to more accurate comprehension of each question. Observations of people reading material that spans an entire page, especially when less than 12-point type is used, suggest that words are more likely to be skipped.

A third format is to print pages on one side only and staple them in the upper left corner. This format provides a cheaper alternative for printing some questionnaires, even facilitating within-office printing and finishing on copy machines. This type of binding is more awkward for the respondent than booklet construction and should not be considered if two-sided printing is to be used. Printing of this nature generally costs more for mailing (larger en-

Figure 3.1 Facing pages from a single column, full-page format using folded legal size paper (8.5" × 14") (formatted by Pavlov, 1996).

Q1. The following questions address perceptions about the general work environment and interactions among co-workers. To what extent do you agree or disagree with each of the following statements about the work environment in your unit or department within your college? (Please circle your answer.)

	Strongly Agree ▼	Somewhat Agree ▼	Somewhat Disagree ▼	Strongly Disagree ▼	Don't Know ▼
Co-workers listen to my ideas	1	2	3	4	5
Co-workers appreciate my contributions	1	2	3	4	5
Co-workers treat me with respect . . .	1	2	3	4	5
I am given the opportunity for professional growth and success in my working environment	1	2	3	4	5
My co-workers generally go out of their way to help new workers succeed and excel in their position . .	1	2	3	4	5
Perks and benefits are distributed equally and fairly in my unit	1	2	3	4	5
Hiring practices in my unit/department have promoted equality among workers	1	2	3	4	5
All qualified candidates have an equal chance of being promoted in my unit/department	1	2	3	4	5
Focusing on equity issues has a negative impact upon teamwork . . .	1	2	3	4	5
My co-workers constructively confront problems	1	2	3	4	5
Supervisors show respect to employees	1	2	3	4	5
Decisions in my unit/department are often influenced by social relationships with key persons	1	2	3	4	5
Problems or issues in my unit/department are handled openly rather than covertly	1	2	3	4	5
I am satisfied with the opportunities I have for promotion	1	2	3	4	5

1

Q2. To what extent do you agree or disagree with each of the following statements about the

management of your unit/department? (Please circle your answer.)

	Strongly Agree ▼	Somewhat Agree ▼	Somewhat Disagree ▼	Strongly Disagree ▼
I would be comfortable approaching my immediate supervisor about concerns of discrimination or harassment	1	2	3	4
I would be comfortable approaching higher level supervisors (e.g., Department Chair, Director, Dean) about concerns of discrimination or harassment	1	2	3	4
In my work environment, supervisors have made it clear that they will not tolerate harassment or discrimination	1	2	3	4
I believe managers in my unit give preferential treatment to individuals who are similar to them	1	2	3	4
I believe that my supervisors would actively intervene to stop conduct that constitutes harassment or discrimination in my work environment	1	2	3	4

Q3. In the past five years, have you felt discriminated against within this college?

- 1 Yes, I have experienced a lot of discrimination.
 2 Yes, I have experienced some discrimination.
 3 Yes, I have experienced a little discrimination.
 4 No, I have not experienced discrimination.

GO TO Q5

Q4. (If Yes) During these last five years in this college, have you experienced discrimination based upon: (Please circle yes or no.)

- a. gender Yes No
 b. race/ethnicity Yes No
 c. disability Yes No
 d. age Yes No
 e. sexual orientation Yes No
 f. religion Yes No

2

Figure 3.2 Page from a double-column format using conventional 8.5" × 11" paper.

Please help plan AAPOR's future!

A 1995 survey of interests and concerns of members of the American Association for Public Opinion Research

Your Relationship to AAPOR

1. Please indicate how *valuable* each of these aspects of AAPOR is to you as a member, where 5 means "very valuable" and 1 means "Not at all valuable."

	Value	Not	DK/
	Very	at all	NA
A. <i>Public Opinion Quarterly</i> ..	5.....4.....3.....2.....1.....		9
B. <i>AAPOR News</i> , the newsletter.....	5.....4.....3.....2.....1.....		9
C. The annual conference	5.....4.....3.....2.....1.....		9
D. Your local chapter	5.....4.....3.....2.....1.....		9
E. The <i>Blue Book</i>	5.....4.....3.....2.....1.....		9
F. The code of ethics.....	5.....4.....3.....2.....1.....		9
G. The membership directory.....	5.....4.....3.....2.....1.....		9

2. In what year did you first join AAPOR?

1995.....	1
1994.....	2
1990-1993.....	3
1985-1989.....	4
1980-1984.....	5
1970-1979.....	6
1960-1969.....	7
Before 1960.....	8

3. Who paid your AAPOR dues in 1995?

I did.....	1
My employer did.....	2
Honorary Life Member (no dues)....	3

4. Are you a member of any other professional association(s)?

Yes.....	1
No (SKIP to Q8).....	2

5. (If yes) To which types of professional associations do you belong? (circle all that apply)

Disciplinary associations (e.g., Amer. Statistical Assn., Amer. Sociological Assn.)	1
Professional associations (e.g., Amer. Marketing Assn., Advertising Research Foundation)	2
Your clients' professional interest associations.....	3
Other types of associations	4

6. In a typical year, how many other associations' annual meetings do you attend?

None	0
One	1
Two	2
Three	3
Four or more.....	4

7. Thinking about all the professional associations in which you participate, which do you think of as your primary association?

AAPOR	1
AAPOR and another equally.....	2
Another association	3
Don't know	9

*Formatted by Karen Goldenberg and Don A. Dillman and reprinted by permission of American Association of Public Opinion Research.

85

velopes as well as more sheets of paper). Generally it lacks the professional look of booklet formats and is one I prefer not to use, except under very tight budgets.

CRITERIA FOR ORDERING THE QUESTIONS

Consider for a moment the order of these questions proposed for a self-administered questionnaire, presented here in abbreviated form:

- What was your total family income in 1998?
- Do you like to play golf?
- What is your opinion on global warming?
- Are you married?
- How many times have you gone bowling during the past year?
- What is your political party preference?
- Do you favor or oppose these measures to reduce environmental pollution?
- What is your occupation?
- Please describe your favorite recreational activity.
- How adequate is your present health care?
- Which political party does the best job of promoting economic growth?
- How old are you?
- Has your health gotten better or worse during the past year?

A friend described her frustration with a similar set of disconnected questions in this way: "It was like trying to go shopping with an alphabetized grocery list. The order on the list didn't correspond to how things were organized in the store, so that I had to go back and forth between sections of the store and I became so confused that some things on the list were forgotten." The above list of questions vacillates back and forth among topics and is not organized in the way one's knowledge of topics is likely to be organized. Respondents are asked to describe recreational activities in several different places, with unrelated questions and personal characteristics interspersed among them. In addition, the first item on the list, family income, is perhaps the most threatening question. Asking income first is somewhat like requesting that respondents sign a blank check when as yet they have little idea of why; the respondent is more likely to reveal such private information if other questions have been answered first.

A questionnaire is like a conversation which typically evolves in accordance with societal norms (Schwarz, 1996). Constantly switching topics makes it appear that the questioner (survey sponsor) is not listening to the respondent's answers. Each answer seems to stimulate a response on an unrelated topic, as though the person's answer was not heard. Jumping quickly

between topics also means that answers are less likely to be well thought out, as new topics evoke top-of-the-head responses. Asking all of the recreational questions from the above list together is not only a more efficient way of asking a person to think about a topic, but it also cuts down on the respondent effort required for focusing on each issue and giving a reasoned response. Grouping questions by topic according to the principles that follow is helpful to respondents.

Initially, question topics and questions are grouped in a general way from most salient to least salient to the respondent. This decision is based on the research finding that a major predictor of response rates to mail surveys is the salience of the questionnaire topic (Heberlein and Baumgartner, 1978). Most questionnaires consist of a variety of questions, some of which respondents will find interesting to answer and some of which they will not. Consequently, we attempt to start the questionnaire with some of the most salient questions.

This question order takes into account what the respondent has been told in the cover letter. Chances are that a strong argument is offered in that letter for answering the questionnaire. This theme is likely to be carried onto a well-designed questionnaire cover page. Suppose, for example, that a cover letter has informed recipients that they are being asked to respond to an important questionnaire about preferences for possible change in the health care program to which they belong. The questionnaire cover carries out that theme with a title such as, "What Changes Would You Like to See Made in Your Present Health Maintenance Program?" These words develop respondents' expectations that they are going to be asked for their opinions about desired changes. Thus, it seems important that early in the questionnaire they are asked to do that. A salient beginning to a questionnaire is partly a matter of meeting respondent expectations and partly a matter of identifying questions that the respondent will find interesting. One might reasonably begin by asking how long respondents have participated in the current program and how well they like it. However, one should not begin by asking a series of disjointed demographic questions—education, age, income, home ownership, veteran status, or even prior medical history. These demographics may be highly relevant to the study objectives and easy to answer, but respondents will not see obvious relevance to the topic. The first pages are also not the place to include a long series of abstract attitudinal scales, such as those designed to measure a respondent's orientation towards government versus private sector involvement in meeting health care needs.

Objectionable questions are placed near the end of the questionnaire, where they are likely to be seen after the respondent has had an opportunity to become interested in the questionnaire. A respondent who has spent five or 10 minutes already answering questions is less likely to respond to an objectionable question by quitting. Moreover, some questions may seem less objectionable in light of previous questions already answered. One of the

questions that people often object to answering is the question of income. In addition to reporting income, people are often uncomfortable answering questions about their sexual behavior or any laws they may have broken. Significant, though less universal, objections are often expressed towards questions about religious activities or political party affiliation. Through pretesting we attempt to identify questions which people object to answering, and place those questions near the end of the questionnaire.

An effort is also made to order questions in a way that will be logical to the respondent. When respondents are asked why they left a job, it makes sense to ask such questions after asking why they accepted that job and how long they were employed in that position. In general, this suggestion implies asking people about things in the order that they happened. It is also helpful to ask people descriptive questions about an activity before requesting evaluations of the experience, thus encouraging more complete recall of past events. For example, a question about the main reason a person left a job might best be asked after a series of items about the nature of that job.

Finally, it is also helpful to group together questions that have similar component parts. For example, if within a general topic area there are several yes/no questions, some that require unfavorable-favorable scalar responses, and others that require agree-disagree responses, the task of the respondent can be eased by grouping those requiring the same answer categories together. Our purpose is to ease the cognitive burden of responding. Responding to questionnaires in which each new question means having to think about a new topic, as well as a new type of response category, requires more respondent effort. Simply put, we are attempting to keep respondents focused on the substance of our inquiries rather than having to figure out new response formats for each question as well. However, this aspect of ordering usually involves only minor shifts, with our concerns being focused first on salience, objectionable qualities, and an order that the respondent will find logical.

POSSIBLE ORDER EFFECTS BETWEEN OPINION QUESTIONS

When taking advantage of cognitive ties to ease the task of responding, one must also be aware of how answering one question may influence responses to later queries and, in particular, the next one. There are five distinct situations in which answers to a subsequent question may be dramatically altered because of the question immediately prior to it, as summarized in Figure 3.3. The first of these concerns stems from invoking a *norm of evenhandedness*. A survey of Washington State University students showed that 34% thought a student who had plagiarized should be expelled when the preceding question asked whether a professor who had plagiarized should be fired. However, when the student question was asked first, only 21% indicated that the student should be expelled (Sangster, 1993). This phenomenon of adjusting

Figure 3.3 Sources of order effects in self-administered questionnaires.

<u>Problem</u>		<u>Example</u>
Norm of evenhandedness, a value-based effect.	☞	Yes answers to "Should a Communist reporter be allowed to report on visit to America as they saw it?" increased significantly when respondents were first asked whether American reporters should be allowed to report on the Soviet Union as they saw it; a sense of fairness had been invoked. (Hyman and Sheatsley, 1950)
Anchoring, a cognitive-based effect.	☞	Student agreement with general statement, "Cheating at universities throughout the United States is widespread problem," decreases significantly when asked after registering extent of agreement or disagreement with a similar statement about their university, because of increase in "No opinion" to second question. (Sangster, 1993)
Addition (carryover) effect.	☞	Percent choosing "Very Happy" for answer to question, "How would you say things are these days?" increases significantly if asked after question, "How would you describe your marriage?" using the same categories; because of carryover of marriage evaluation to the second, more general question. (Schuman and Presser, 1981)
Subtraction effect.	☞	Percent indicating they feel the economic situation is getting better in the state was significantly higher when that question was asked prior to question about economic situation in their community; because of subtracting out information on which first question was based. (Mason, Carlson, and Tourangeau, 1994)
Increased positiveness of summary items when asked after specific items on same subject.	☞	Respondents give higher ratings to the question, "How would you rate the overall quality of life in your community?" when asked after several rating questions about more specific domains of community life. (Willits and Saltiel, 1995)

answers to succeeding questions based on the previous answer was first noticed in interview surveys, with a question asked in 1948 about whether a communist reporter should be allowed to report on a visit to America as he saw it. When this question was asked before a comparable question about whether an American reporter should be able to report on a visit to the Soviet Union as he saw it, 37% said yes, compared to 73% when the question was positioned second (Hyman and Sheatsley, 1950).

The explanation for this order effect is *value-based*; that is, the similarity of the questions is seen as invoking a norm of fairness or evenhandedness which results in one's answer to the second question taking into account one's answer to the first. It has been argued, based on a classroom experiment, that the norm of evenhandedness is less likely to be invoked in mail than telephone surveys (Bishop, Hippler, Schwarz, and Strack, 1988). Bishop et al. argue that the respondents to a self-administered questionnaire can look ahead to see what is coming and therefore adjust their answers to earlier questions. However, other experiments from field surveys, including a replication of the communist reporter questions by Ayida and McClendon (1990), as well as results reported by Sangster (1993) and Sangster, Lorenz, and Saltiel (1995), provide evidence that effects in mail and telephone surveys are similar.

An *anchoring* or *cognitive-based* order effect has also been observed. Another experiment reported by Sangster (1993) asked a random sample of students whether they agree or disagree that "cheating is a widespread problem at this university." The next question asked whether "cheating at universities throughout the United States is a widespread problem." A comparable random sample was asked the questions in reverse order. The percentage that agreed to either question is similar when that question is answered first (65% United States versus 60% their university), and much smaller (49% versus 44%) when answered second. We conclude that students are much more likely to see cheating as a problem at their university if asked about it after the question about other U.S. universities. These differences are accounted for almost entirely by a change in the use of the no opinion category. When the U.S. question is asked first 15% offer no opinion for that question followed by a similar 16% for the student's university. However, when asked first about their own university with which they should have greater familiarity, 29% offer no opinion, which then rises to 38% for universities throughout the United States. In essence, the response to the first of these opinion questions serves as an anchor for the second answer.

Two other order effects that have been identified through past research are *addition* or *carryover effects* and *subtraction effects*. Schuman and Presser (1981) found that when a specific question, "How would you describe your marriage?" was asked just prior to a general question, "How would you say things are these days?" the percentage responding to the general question varied greatly depending upon which question was asked first. Whereas 52% said very happy (compared to pretty or not too happy) to the general question when it was asked after the marriage question, only 38% said very happy when the order was reversed. Their explanation of this addition effect was that people continued to think about the marriage question (to which 70% said very happy) when answering the general question.

A seemingly opposite effect is illustrated by these questions from Mason, Carlson, and Tourangeau (1994), asked in this order:

"How would you describe the economic situation in your community over the next five years? Do you feel it will Get Better, Get Worse, or Stay the Same?"

"How do you feel about the economic situation in (your state) over the next five years? Do you feel it will Get Better, Get Worse, or Stay the Same?"

Mason and his colleagues found that more people (7–10%) said the state economy would get better when the state economy was presented first. The reason for this difference is that once people have answered the first question they tend to "subtract" out reasons used to justify their first answer (e.g., new industry) and therefore base their answer to the second question in the series on other considerations.

Finally, a body of research has emerged that summarizes the effects of asking general or summary items, such as, "How would you rate the *overall* quality of life in your community?" prior to and after asking about a number of specific domains from streets and roads to education. An example of a specific item is: "How would you rate police protection in your community?" Studies have found that the summary question tends to be scored lower by respondents when asked before a list of specific domain questions than when asked afterwards (Willits and Saltiel, 1995).

Schwarz (1996) has detailed how, in the normal give-and-take of regular conversations, people tend to give answers that take into account things they have already said. Thus, answers to individual questions are less complete, or less able to stand alone, than the framers of those questions intended. Although the carryover is probably most extensive from questions that immediately follow one another, there is limited evidence that effects also occur when questions are widely separated.

Consequently, it is important to recognize early on that a questionnaire cannot be viewed as a compilation of completely independent questions that have no effects on one another. Not only must each question be evaluated on the basis of its individual content, but also with regard to the larger context that often adds or subtracts meaning. It is also important to note that all of the order effects reported here as illustrations of the norm of evenhandedness, the anchoring effect, addition effect, subtraction effect, and summary item effect are opinion questions that require the respondent to formulate an answer on the spot. Thus, we should not be surprised that respondents identify the questions as related to one another and adjust their answers to the second question based on answers to the first one.

The practical conclusion we draw from these various studies is that there are quite different kinds of order effects likely to occur on topically-related opinion questions that use commonly employed but vague quantifiers (agree-disagree, excellent to poor), and these effects cannot be ignored. Although there is some evidence to suggest that the effects may be stronger in telephone

than mail surveys (Lorenz and Ryan, 1996; Bishop et al., 1988), the evidence that they persist in self-administered surveys is also compelling (Sangster, 1993). When one anticipates including questions like those described in Figure 3.3, it may be judicious to consider doing half the questionnaires with one order and half with another, or at a minimum recognize in reports of findings the possibility of question order influence on respondent answers. This topic remains an active area of research about which much remains to be learned.

CHOOSE THE FIRST QUESTION CAREFULLY

No single question is more crucial than the first one. It is more likely than any other questionnaire item to determine whether that questionnaire is destined for the mailbox or the garbage. Thus, it warrants special attention, with some questions serving this purpose better than others, as shown in Figure 3.4.

First, the question should clearly *apply to everyone*. A questionnaire that begins by asking, "From what company do you obtain your health insurance?" assumes that the respondent has health insurance. Although a category, "I don't have health insurance" can be added, the damage is already done. Respondents who read only the query are likely to conclude that the questionnaire really does not apply to them and need not be filled out.

Second, the question should be *easy*, so that all respondents will need only a few seconds to read, comprehend, and respond to it. In telephone interviewing, a common technique for overcoming refusals from people who do not think they are able to respond is to read them the first question, which is easy to understand and answer. Success in understanding and answering the first question encourages people to continue. This is *not* the place for a long question with many response categories or an open-ended question. A test of this idea was done at the U.S. Bureau of the Census, when one questionnaire began with the request that the respondent list the names of all members of the household according to a list of eligibility criteria. The other questionnaire was identical except that it began with the simple question, "How many people live in this household?" The questionnaire that began with the latter item achieved a response rate that was nearly four percentage points higher than the other form (Leslie, 1996). Surveyors must be especially concerned with the ease of answering the first question because it seems likely to produce nonresponse error by discouraging people for whom filling out a written questionnaire is difficult.

Third, the first item in a questionnaire needs to be *interesting*. The well-known role of salience in improving mail survey response suggests that questions that are interesting are more likely to be answered. I was once faced with providing advice for a particularly difficult survey in which a low response rate was expected. The population consisted of licensed commercial salmon fishermen and concerned mostly the size of their boats and investment in their

Figure 3.4 Selecting the first question for a questionnaire.**Unacceptable first questions:**

1. Please think about all the things that make this community a pleasant place to live. Then please write down the five most important aspects of community that make this community a pleasant place to live, and rank them from 1 (meaning most important) to 5 (the least important of the five attributes).


1. Please describe in your own words what you consider good about living in this community?

1. What year were you born?
 _____ year born

Better first questions:

1. Thinking about this community, how would you rate it as a place to live?
 - ☐ Excellent
 - ☐ Good
 - ☐ Fair
 - ☐ Poor

1. How long have you lived in this community?
 - ☐ More than six months
 - ☐ Less than six months



If less than six months, it is not necessary for you to complete the remainder of this questionnaire. However, please return it so that we can check your name off of the mailing list. That will help us a great deal.

equipment. After much discussion with the survey sponsor, the sponsor decided to add an introductory section about their views of the future of salmon fishing, whether they would advise young people to enter it, and other issues that were described to me as the “gut” issues facing their business. The entire first page of the questionnaire was devoted to these questions, with the first one being a simple question about whether the respondent felt that the benefits of being a salmon fisherman were getting better, worse, or staying about the same. The response rate for this study of individuals, many of whom had not completed high school, was well over 50%.

Frequently, when we have advised questionnaire writers to keep in mind questions that are interesting, easy, and applicable to everyone, the response has been, “We’ll start with age or education. It applies to everyone and is really easy.” Although these questions are easy to answer, they are not particularly interesting. Most importantly, they do not meet a fourth criterion of *connectedness* between the respondent and survey purpose as understood by that person. A well-designed survey will include a cover letter which describes what the survey is about and why it is important to respond. Respondents who have been told that their response is important on a survey about health care policies, and who are perhaps looking forward to responding, are likely to be unpleasantly surprised to turn to the first question(s) and discover that they are about age or education. From their perspective these questions have little to do with evaluating health care possibilities. For this reason, we seldom begin a questionnaire with demographic questions.

Special situations sometimes override our search for a first question that meets all of the above criteria. Most common is the situation in which eligibility requirements must be met in order to fill out a questionnaire. In these instances we use the first question (supported by appropriate wording in the cover letter) to determine eligibility and to instruct both those eligible and those not eligible to return the questionnaire, as shown at the bottom of Figure 3.4. One of the major reasons for nonresponse is that questionnaires are discarded because they do not apply to the recipient, and hearing from such respondents is important for understanding coverage and nonresponse issues.

PRINCIPLES FOR CONSTRUCTING QUESTIONNAIRE PAGES

CREATING A COMMON STIMULUS FOR EVERY PERSON

Telephone interviewers are typically instructed to read the same words in the same order to each respondent, so that each respondent receives exactly the same stimulus. Making sure each respondent receives the same stimulus in the same way is considered essential for obtaining quality survey data. Consequently, great care is taken to train interviewers so that the order in which words are delivered does not change from one respondent to the next. Therefore, questions are read verbatim and no words are skipped.

My goal for creating the visual layout of a self-administered questionnaire is the same. Respondents should see and comprehend every word of every question in a particular order, just as it would be delivered by an interviewer. Words construct phrases and phrases build sentences. It is important not only to ensure that people read and comprehend words in a prescribed order but to facilitate comprehension in appropriate *groupings*. Each question and its set

of answer choices comprise a set of words and phrases that need to be viewed as a group—the first part that states the question, and the second part that provides the alternatives for responding.

Concern about getting each respondent to read every word of each question in the same order and with the same grouping (phrasing) of ideas stems partly from the fact that missed words, misunderstood questions, and unseen answer categories constitute major sources of error in self-administered surveys. This concern also stems from the increased prevalence of mixed-mode surveys, in which some data are collected by interview and other portions by self-administered form, and from the frequent reporting of response differences between interview and self-administered surveys (e.g., Dillman, Sangster, Tarnai, and Rockwood, 1996). Achieving the same stimulus in both modes is important for minimizing mode differences. In addition, I am concerned that by not guiding people effectively in how to fill out self-administered questionnaires, frustration is produced that becomes expressed through less thoughtful answers, higher item nonresponse, and occasionally, no response at all.

The challenge of achieving delivery of the same stimulus for all respondents is far more difficult for self-administered than for telephone surveys. Telephone respondents receive the interview stimulus only through hearing, and the order in which words are delivered is controlled by the interviewer. Self-administered questionnaires mostly involve a different human sense, seeing. However, unlike the interview situation, the comprehension order is controlled by the respondent. It is for this reason that the visual aspects of design take on paramount importance in self-administered questionnaire design.

Computers now have marvelous capabilities that allow most individuals to design creative questionnaire layouts. For example, it is possible to insert a variety of symbols **+**☺✓, increase **font size**, and use **bold print**, *italics*, or **shaded backgrounds** all in a single sentence, as done here. Used separately and consistently, such capabilities can ease the task of comprehending and responding to a questionnaire. Used carelessly, and especially together in one sentence, this variation can make a questionnaire far more difficult to complete than necessary, and even lead to important words being misread or ignored altogether.

The content of each questionnaire page is comprised of stimuli presented in two languages, each of which provides meaning and direction to respondents, as described by Jenkins and Dillman (1995, 1997). One of those languages consists of written words and is the traditional one considered important in the design of questionnaires. The second language consists of graphical symbols and arrangements which also give direction to respondents with regard to what they should read and in what order. The meaning of this second language is communicated through a number of visual clues.

Six visual elements, to be described in detail later, are especially important for determining how people divide what they see into separate groupings, such as one complete question or a complete set of response categories, and how they proceed to read and answer a page of survey questions. They include *location* (or spacing between elements), *shape, size, brightness* (shading or color), *simplicity and regularity*, and a *consistent figure-ground* format. Together they influence which words are read in what order and to some extent, the meaning of those words.

The task of constructing a questionnaire page is to bring the verbal language and the graphical language into concert with one another in order to communicate questions and obtain answers. In this section, I describe several key steps for combining written language with the visual language of questionnaire design.

Accomplishing the goal of a common stimulus for all respondents involves adhering to many construction principles for presenting survey questions to respondents. They encompass decisions about the order in which all printed information is to be presented on each page and the use of graphic design elements to guide the respondent through the questions in the desired order. This process can conveniently be divided into three sequential steps:

- Step 1: Defining a desired navigational path for reading all information presented on each page of the questionnaire.
- Step 2: Creating visual navigational guides that will assist respondents in adhering to the prescribed navigational path and correctly interpreting the written information.
- Step 3: Developing additional visual navigational guides, the aim of which is to interrupt established navigation behavior and redirect respondents, for example, through skip patterns.

Each of these steps is described below in conjunction with principles essential to its implementation.

STEP 1: DEFINE A DESIRED NAVIGATIONAL PATH FOR READING ALL INFORMATION PRESENTED ON EACH PAGE OF THE QUESTIONNAIRE.

Defining a desired navigational path consists of nothing more than determining the order in which every stimulus on the questionnaire page, including all words and graphical symbols, should be processed by the respondent. I want to organize this information in a way that makes obvious the right way to proceed through the questionnaire and that, consistent with exchange theory, will minimize the effort needed to comprehend and respond to each question. This concern leads to the development of several principles for information organization.

Principle 3.1: Write each question in a way that minimizes the need to reread portions in order to comprehend the response task.

Consider the organization of words and phrases of the first question in Figure 3.5. Not only are the answer categories presented first (which most people would never do), but the wording of the question presents answering instructions and choices to the respondent before presenting the substance of the question. This organization of words and phrases would invariably require that the respondent retrace a portion of the navigational path in order to understand and answer the question. The revision illustrates the more conventional way of organizing the information in a question.

A more complex example of inefficient word organization is a question taken from the 1993 U.S. Census of Agriculture (Figure 3.6). In this example it is not clear exactly what is being asked until the response space where “Number of acres owned” appears. Respondents who have been following the conventional left to right and top to bottom reading sequence through these separated phrases are reading words without knowing what the question is. The inevitable result is that it will be necessary to reread the question in order to understand it. The drawback to poor information organization is that respondents may become frustrated and/or unwilling to retrace those steps, and therefore may give a wrong answer. In this case the problem is compounded by a visual layout that makes it somewhat unclear what navigational path is to be followed; that is, what information is to be read in what order.

A more effective organization of the information is shown in the revision in

Figure 3.5 Building a more desirable information organization for the navigational path.

Poor information organization:

12. ☐ Very Satisfied
☐ Somewhat Satisfied
☐ Not At All Satisfied

Please check the appropriate category above to indicate whether you were Very Satisfied, Somewhat Satisfied, or Not At All Satisfied with the quality of your meal the last time you visited our restaurant.

Better information organization:

12. The last time you visited our restaurant, how satisfied were you with the quality of your meal? (Check one box.)
- ☐ Very Satisfied
☐ Somewhat Satisfied
☐ Not At All Satisfied
-

Figure 3.6 Poor information organizations with unclear navigational path, and a revision.***Poor information organization and lack of navigational path:***

CENSUS USE ONLY	035	036	037	038	039	040	041	042
SECTION 1 ACREAGE IN 1992 – Report land owned, rented, or used by you, your spouse, or by the partnership, corporation, or organization for which you are reporting. <i>Include ALL LAND, REGARDLESS OF LOCATION OR USE – cropland, pastureland, rangeland, woodland, idle land, house lots, etc.</i>								
S1 <i>If the acres you operated in 1992 changed during the year, refer to the INFORMATION SHEET, section 1.</i>								None Number of acres 043
1. All land owned								<input type="checkbox"/>

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, pasture land, rangeland, woodland, idle land, house lots, etc.), regardless of location, owned by you, your spouse, or by the partnership, corporation or organization for which you are reporting. *(If the acres you operated in 1990 changed during the year, refer to the information sheet, Section 1.)*

_____ Number of acres owned

From 1993 Census of Agriculture conducted by U.S. Bureau of the Census.

Figure 3.6. This allows respondents to know at the beginning that they are being asked to report the number of acres they own, and 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 organized information, which, once read, leaves the respondent unclear about precisely what to do.

Principle 3.2: Place instructions exactly where that information is needed and not at the beginning of the questionnaire.

The information on how to complete a questionnaire (shown in Figure 3.7) was provided in an “instructions” section on page one of a questionnaire sent to nearly 200,000 Americans. At first glance, these instructions seem rather ordinary and reasonable. Upon closer examination, one should ask, “Why tell people where the directions are going to be? Do they need to know up front that not every question applies to everyone?” This information should be revealed at the appropriate time. The direction to refer to the week of April 15th seems useful, except the respondent is also informed that it will not apply to some questions. Thus, the information about the week of April 15th is not self-standing and immediately useful. The benefit of knowing this information has been undone. The direction on what to do if no “Skip” instruction is pro-

Figure 3.7 Place instructions exactly where that information is needed and not in a separate section at the beginning of the questionnaire.

Problem: Instructions placed in a separate section at the beginning:

- Thank you for taking the time to complete this important questionnaire. The directions for filling it out are provided with each question. Because not all questions will apply to everyone, you may be asked to skip certain questions.
- In order to get comparable data, we will be asking you to refer to the week of April 15, 1993, when answering most questions.
- If no “Skip” instruction is provided, you should continue to the NEXT question.
- Either a pen or pencil may be used.
- When answering questions that require marking a box, please use an “X.”
- If you need to change an answer, please make sure that your old answer is either completely erased or clearly crossed out.

A revision: Placing instructions exactly where they are needed:

1. Were you working for pay (or profit) during the week of April 15, 1993? This includes being self-employed or temporarily absent from a job (e.g., illness, vacation, or parental leave), even if unpaid. Mark your answer in the box ☐ with a pen or pencil.

☐ Yes ☐ No
- ◆ 2. (If No) Did you look for work at anytime during the four weeks preceding April 15?

☐ Yes
☐ No

vided is completely unnecessary, and the practical implication of telling people that either a pen or pencil may be used is probably minimal.

Instructions should be provided at the point that respondents are ready to act on them. Thus, the place to instruct people on the type of mark to make, what date to use, etc., is when a question has been asked and the respondent is ready to provide an answer that uses the instructions. In the proposed revision of these instructions, the middle four items are provided with the first two questions and the remaining instructions have been eliminated because they are unnecessary.

Besides getting better compliance with instructions by placing them exactly where they are needed, this placement serves two other important functions. First, it encourages people to immediately start answering questions, so

they can feel rewarded by a sense of immediate progress. Second, providing instructions that are not immediately useful encourages people to scan information rather than to read it carefully. Encouraging respondents to skip words and phrases early in the questionnaire is likely to cause them to read questions in that way later, an outcome that should be avoided.

Even worse than the above example is the tendency for some questionnaires to compile separate instruction booklets or place instructions in a different place in the questionnaire. People are unlikely to go to the instruction booklet before answering each question. At best, instruction books are used unevenly by respondents, resulting in some respondents being subjected to different stimuli than are others.

Principle 3.3: Place items with the same response categories into an item-in-a-series format, but do it carefully.

When one has to ask a series of questions that use the same answer categories, it is convenient to combine them into a format similar to the one shown in the revision example in Figure 3.8. Here respondents are asked to indicate the extent to which each of the listed concerns is a problem in their community. Combining them into this item-in-a-series format, with a common introduction that defines the general question and response format, eliminates considerable redundancy with regard to stating questions; it also saves considerable space.

However, it also changes the nature of the questions. That is, the items are to some extent placed in a comparative framework, whereby the visual structure encourages respondents to think of them as a unit. As a result, when answering one item respondents think about it in relation to other items. Often this is precisely what the questioner wants, and the structure seems ideal for that purpose. However, if for some reason the sponsor wants individuals to contemplate each item separately, it is advisable to present each of them as individual items.

Principle 3.4: Ask one question at a time.

Questionnaires are sometimes constructed in a way that encourages the respondent to answer two questions at once, as shown in Figure 3.9. The problem is that the person who answers in this way must toggle her mind between the questions as she proceeds down the page. The usual justification is that this approach cuts the amount of space required in half. However, this format is exceedingly difficult for many respondents as they try to focus on the basic items of information on the left side of the page and answer both questions for each item.

Whereas some respondents answer the first question for each item and then proceed back over the same items to answer the second question, others try to do both at once. It is undesirable to lose control of the answering pro-

Figure 3.8 Place items with the same response categories into an item-in-a-series format.

An inefficient structure:

7. To what extent do you consider a lack of rental housing to be a problem in this community?

- ☐ Not a Problem
- ☐ Small Problem
- ☐ Moderate Problem
- ☐ Serious Problem

8. To what extent do you consider poor road and street repair to be a problem in this community?

- ☐ Not a Problem
- ☐ Small Problem
- ☐ Moderate Problem
- ☐ Serious Problem

Etc.

A revision that places questions into an item-in-a-series format:

7. Do you consider each of the following to be a Serious Problem, Moderate Problem, Small Problem, or Not a Problem in this community? (Please circle one answer for each.)

Extent to which situation is a problem in this community

A lack of rental housing	Serious	Moderate	Small	Not a Problem
Poor road and street repair	Serious	Moderate	Small	Not a Problem

Etc.

cess in this way so that people think in different ways about the items. An additional problem with double, or sometimes triple, questions is that the respondent who tries to answer both questions and toggles her mind between items may have difficulty remembering exactly what she is supposed to do. Still another problem is that, as a practical matter, setting up the questions often requires abbreviations, so people must constantly refer from one place to another as they try to remember all of the elements of the question. For example, use in the “last six months” seems especially likely to be forgotten as a respondent proceeds through the long list of items. For all of these reasons, this format should not be used for self-administered questionnaires. The usual justification that it will make the questionnaire look shorter and easier is *not* one for which we have seen explicit evidence that it improves response

Figure 3.9 Ask one question at a time.**A Problem:**

6. Please indicate the extent to which each of the following services of our organization are important to you, and for those that are very important, please indicate how frequently you have used that service during the past six months.

<u>How Important?</u>			<u>Frequency of use in last six months</u>
Customer credit	Not	Somewhat Very	_____ Times Used
Next-day delivery	Not	Somewhat Very	_____ Times Used
“No questions asked” return policy	Not	Somewhat Very	_____ Times Used

A revision:

6. How often have you used each of these services from our organization during the last six months?

<u>Number of times used</u>
Customer credit _____
Next-day delivery _____
“No questions asked” return policy _____

7. Thinking about each of these same services, how important is each of them to you?

<u>How Important?</u>		
Customer credit	Not	Somewhat Very
Next-day delivery	Not	Somewhat Very
“No questions asked” return policy	Not	Somewhat Very

rates. Finally, some evidence exists that item nonresponse is significantly higher for these types of double-question formats (Keil and Gaertner, 1998).

Principle 3.5: Minimize the use of matrices.

A closely related type of question is the matrix question, in which one prints a series of individual questions down the left side of the page and then lists a series of units across the top of the page for which respondents are to answer these questions (Figure 3.10). An example of such a matrix is the household composition question asked in the 1990 Decennial Census, in which respon-

Figure 3.10 Minimize the use of matrices.

A matrix structure used in 1990 Decennial Census for seven people (only 3 spaces shown):

Page 2 **PLEASE ALSO ANSWER HOUSING QUESTIONS ON PAGE 3**

	PERSON 1	PERSON 2
Please fill one column for each person listed in Question 1a on page 1.	Last name First name Middle initial	Last name First name Middle initial
2. How is this person related to PERSON 1? Fill ONE circle for each person. If Other relative of person in column 1, fill circle and print exact relationship, such as mother-in-law, grandparent, son-in-law, niece, cousin, and so on.	START in this column with the household member (or one of the members) in whose name the home is owned, being bought, or rented. If there is no such person, start in this column with any adult household member. ■	If a RELATIVE of Person 1: <input type="radio"/> Husband/wife <input type="radio"/> Brother/sister <input type="radio"/> Natural-born or adopted son/daughter <input type="radio"/> Father/mother <input type="radio"/> Grandchild <input type="radio"/> Stepson/stepdaughter If NOT RELATED to Person 1: <input type="radio"/> Roomer, boarder, or foster child <input type="radio"/> Unmarried partner <input type="radio"/> Housemate, roommate <input type="radio"/> Other nonrelative ■
3. Sex Fill ONE circle for each person.	<input type="radio"/> Male <input type="radio"/> Female	<input type="radio"/> Male <input type="radio"/> Female
4. Race Fill ONE circle for the race that the person considers himself/herself to be. If Indian (Amer.) , print the name of the enrolled or principal tribe. If Other Asian or Pacific Islander (API) , print one group, for example: Hmong, Fijian, Laotian, Thai, Tongan, Pakistani, Cambodian, and so on. If Other race , print race.	<input type="radio"/> White <input type="radio"/> Black or Negro <input type="radio"/> Indian (Amer.) (Print the name of the enrolled or principal tribe.) <input type="radio"/> Eskimo <input type="radio"/> Aleut <input type="radio"/> Asian or Pacific Islander (API) <input type="radio"/> Chinese <input type="radio"/> Japanese <input type="radio"/> Filipino <input type="radio"/> Asian Indian <input type="radio"/> Hawaiian <input type="radio"/> Samoan <input type="radio"/> Korean <input type="radio"/> Guamanian <input type="radio"/> Vietnamese <input type="radio"/> Other API <input type="radio"/> Other race (Print race) ■	<input type="radio"/> White <input type="radio"/> Black or Negro <input type="radio"/> Indian (Amer.) (Print the name of the enrolled or principal tribe.) <input type="radio"/> Eskimo <input type="radio"/> Aleut <input type="radio"/> Asian or Pacific Islander (API) <input type="radio"/> Chinese <input type="radio"/> Japanese <input type="radio"/> Filipino <input type="radio"/> Asian Indian <input type="radio"/> Hawaiian <input type="radio"/> Samoan <input type="radio"/> Korean <input type="radio"/> Guamanian <input type="radio"/> Vietnamese <input type="radio"/> Other API <input type="radio"/> Other race (Print race) ■
5. Age and year of birth a. Print each person's age at last birthday. Fill in the matching circle below each box. b. Print each person's year of birth and fill the matching circle below each box.	a. Age 0 0 0 0 0 1 0 1 0 1 2 0 2 0 3 0 3 0 4 0 4 0 5 0 5 0 6 0 6 0 7 0 7 0 8 0 8 0 9 0 9 0 b. Year of birth 1 8 0 0 0 0 0 9 0 1 0 1 0 2 0 2 0 3 0 3 0 4 0 4 0 5 0 5 0 6 0 6 0 7 0 7 0 8 0 8 0 9 0 9 0 ■	a. Age 0 0 0 0 0 1 0 1 0 1 2 0 2 0 3 0 3 0 4 0 4 0 5 0 5 0 6 0 6 0 7 0 7 0 8 0 8 0 9 0 9 0 b. Year of birth 1 8 0 0 0 0 0 9 0 1 0 1 0 2 0 2 0 3 0 3 0 4 0 4 0 5 0 5 0 6 0 6 0 7 0 7 0 8 0 8 0 9 0 9 0 ■
6. Marital status Fill ONE circle for each person.	<input type="radio"/> Now married <input type="radio"/> Separated <input type="radio"/> Widowed <input type="radio"/> Never married <input type="radio"/> Divorced	<input type="radio"/> Now married <input type="radio"/> Separated <input type="radio"/> Widowed <input type="radio"/> Never married <input type="radio"/> Divorced
7. Is this person of Spanish/Hispanic origin? Fill ONE circle for each person. If Yes, other Spanish/Hispanic , print one group.	<input type="radio"/> No (not Spanish/Hispanic) <input type="radio"/> Yes, Mexican, Mexican-Am., Chicano <input type="radio"/> Yes, Puerto Rican <input type="radio"/> Yes, Cuban <input type="radio"/> Yes, other Spanish/Hispanic (Print one group, for example: Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.) ■	<input type="radio"/> No (not Spanish/Hispanic) <input type="radio"/> Yes, Mexican, Mexican-Am., Chicano <input type="radio"/> Yes, Puerto Rican <input type="radio"/> Yes, Cuban <input type="radio"/> Yes, other Spanish/Hispanic (Print one group, for example: Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.) ■
FOR CENSUS USE	<input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/>

From 1990 Decennial Census conducted by U.S. Bureau of the Census.

(Continued)

Figure 3.10 (Continued)

Example of an individual person structure that replaced matrix in an experimental test Census questionnaire:

United States Census 2000

U.S. Department of Commerce • Bureau of the Census

This is the official form for all the people at this address. It is quick and easy, and your answers are protected by law. Complete the Census and help your community get what it needs — today and in the future!

Start Here

→ Please use a black or blue pen.

1. How many people were living or staying in this house or apartment on April 18, 1998?

Number of people

BE SURE TO INCLUDE anyone who is:

- a foster child, roomer, or housemate
- staying here on April 18, 1998 and has no other permanent place to stay
- staying here most of the time while working even if he or she has another place to live

DO NOT INCLUDE anyone who:

- is living away while attending college
- was in a correctional facility, nursing home, or mental hospital on April 18, 1998
- is in the Armed Forces and living somewhere else
- lives or stays at another place most of the time

→ If this house or apartment is a vacation or seasonal home or a temporary residence for your household, please call the Census Bureau at 1-888-421-1998 before you fill out this form. The telephone call is free.

2. Is this house, apartment, or mobile home —
Mark ☒ ONE box.

☐ Owned by you or someone in this household with a mortgage or loan?

☐ Owned by you or someone in this household free and clear (without a mortgage or loan)?

☐ Rented for cash rent?

☐ Occupied without payment of cash rent?

3. Please answer the following questions for each person living in this house or apartment. Start with the name of one of the people living here who owns, is buying, or rents this house or apartment. If there is no such person, start with any adult living or staying here. We will refer to this person as Person 1. What is this person's name?

Last Name

First Name

MI

4. What is Person 1's telephone number? We may call this person if we don't understand an answer.

Area Code + Number

5. What is Person 1's sex? Mark ☒ ONE box.

☐ Male

☐ Female

6. What is Person 1's age and what is Person 1's date of birth? Print numbers in boxes.

Age on April 18, 1998

Month Day Year of birth

→ **NOTE: Please answer BOTH Questions 7 and 8.**

7. Is Person 1 Spanish/Hispanic/Latino? Mark ☒ the "No" box if **not Spanish/Hispanic/Latino.**

☐ No, not Spanish/Hispanic/Latino

☐ Yes, Mexican, Mexican Am., Chicano

☐ Yes, Puerto Rican

☐ Yes, Cuban

☐ Yes, other Spanish/Hispanic/Latino — Print group:

8. What is Person 1's race? Mark ☒ one or more races to indicate what this person considers himself/herself to be.

☐ White

☐ Black, African Am., or Negro

☐ American Indian or Alaska Native — Print name of enrolled or principal tribe:

☐ Asian Indian

☐ Japanese

☐ Native Hawaiian

☐ Chinese

☐ Korean

☐ Guamanian or Chamorro

☐ Filipino

☐ Vietnamese

☐ Samoan

☐ Other Asian — Print race:

☐ Other Pacific Islander — Print race:

☐ Some other race — Print race:

→ If more people live here, continue with Person 2.

Form S-683A (4-14-98)
OMB No. 0607-0725; Approval Expires 09-30-98

From Census Test Form developed by U.S. Bureau of the Census.

dents were asked to report for each person living in the household, name, age, gender, relationship to Person 1, race, and ethnicity. An experiment provided evidence that changing from a matrix to individual-space format (whereby the questions are repeated in successive spaces for each person) improved response slightly and also reduced item nonresponse (Dillman, Sinclair, and Clark, 1993; Dillman, Clark, and Treat, 1994). In the second of these tests the change to an individual-person format required an additional eight pages (from 20 to 28), but overall response improved from 3 to 4 percentage points.

My general objection to the use of matrices is based on two considerations. One is that they are difficult. The ability to relate rows to columns is a literacy skill much more difficult than simply reading and answering individual questions. My second objection is that matrices usually result in a loss of control over the basic navigational path, and with it is lost maintenance of a common stimulus as respondents read and answer each question.

STEP 2. CREATE VISUAL NAVIGATIONAL GUIDES AND USE THEM IN A CONSISTENT WAY TO GET RESPONDENTS TO FOLLOW THE PRESCRIBED NAVIGATIONAL PATH AND CORRECTLY INTERPRET THE WRITTEN INFORMATION.

Once information has been organized in an order that is efficient for the respondent, it is possible to add visual components, the aim of which is to guide respondents through the words and sentences in the desired order. In previous works with Cleo Jenkins (Jenkins and Dillman, 1995, 1997), the theoretical process has been detailed by which people visually perceive and give meaning to words and other symbols. This work is summarized briefly below.

In order to understand a question and the answer possibilities, respondents must use their previous knowledge to gather and interpret the stimuli being registered by visual information printed on the questionnaire page. A pattern is recognized and given meaning partly by *bottom-up processing*, or information gained only through one's senses. Additional meaning is arrived at through *top-down processing*, which refers to what the respondent expects to see based upon past experiences and the visual context.

In addition, the visual information on each page may be examined in two quite different ways. One of these ways is through *preattentive processing*, that is, when the respondent looks at the entire page or visual field and identifies enough features to make general sense of what is there. However, when attentive to the task of responding, the field of vision is much more restricted. When focused, vision and reading are limited to about two degrees, enough to have about eight to nine characters in sharp visual focus, versus as much as 210 degrees at the preattentive stage (Kahneman, 1973). It is critical at the preattentive stage that the respondent be able to figure out quickly where to start and how. Then, when the respondent is focused on responding and see-

ing clearly within a limited visual field (*attentive processing*), it must be clear exactly what is to be done, for example, how to mark an answer and how to get from one question to the next. Important information located 15, 20, or more characters away from the answer box (for example, skip instructions) are therefore likely to be missed.

There are two general aspects to the design task of helping people solve the visual challenge of figuring out what to do. The first involves recognizing inherent cultural tendencies. In western cultures, people learn from an early age to start near the upper left corner of each page, move from left to right, and generally go from top to bottom. Any questionnaire that attempts to get people to do otherwise faces formidable problems. The second aspect is to modify the visual appeal of particular questionnaire elements in order to attract or diminish the respondent's visual attention to various features of the questionnaire page. Six specific aspects of words and other symbols can be shaped visually to enhance or diminish the likelihood of their being perceived and comprehended in the manner desired by the questionnaire designer.

Visual Element 1: Increase the size of written elements to attract attention.

Words, numbers, and graphic symbols can be increased in size to attract the respondent's eyes, or decreased in size to avoid commanding the respondent's attention. It follows that selectively increasing the size of certain symbols is a way of drawing the respondent's attention to the start of the navigational path, or away from it. Therefore, it is important not to insert large fonts in the middle of a questionnaire page and draw attention away from where a respondent should begin to read, as illustrated by the first example in Figure 3.11. The revision shown in Figure 3.11 uses the largest font on the page at the beginning in a culturally conventional way, with the words "Start Here."

Visual Element 2: Increase the brightness or color (shading) of visual elements to attract attention and establish appropriate groupings.

Similarly, patches of an additional color, such as green on a page of black letters printed on white paper, can be used to attract respondent attention to a particular area of a page. Alternatively, the brightness of a color can be increased, such as changing from gray to black or from a lighter color (usually done by printing with a screen that produces a specified percentage of full color) to darker color. New colors and variations in brightness can be used on the same page to gain adherence to the prescribed navigational path. This feature is illustrated in Figure 3.11 by the change from light type to bold for the question, the question number (in reverse print), and Start Here instructions. Thus, size and brightness changes have been used in concert with one another to identify the place where one should start reading the question.

However, guiding respondents through a set of questions involves considerably more than simply attracting attention. It also requires that respondents

Figure 3.11 Increase size and brightness of visual elements to emphasize order for reading questionnaire information.

Problem:

Start Here:

1. Which of the following is your main work activity?

- ☐ **Research**
 - ☐ **Teaching**
 - ☐ **Administration**
 - ☐ **Something else. (Please Specify)**
-

A revision:

START HERE:

① Which of the following is your main work activity?

- ☐ Research
 - ☐ Teaching
 - ☐ Administration
 - ☐ Something else (Please specify)
-

see information in appropriate groupings. For example, a question number, the question, special instructions, and the answer choices that comprise each question need to be grouped together as a unit. Within the question each portion needs to be seen as an appropriate subgroup. Figure 3.12 shows an example in which these subgroups cannot immediately be distinguished. The revision uses brightness to distinguish between question stems and answer categories. It is also used to draw attention to the start of each question, by putting the question number in reverse print. Accomplishing the desired grouping effect in this example relies partly on these visual elements plus the two elements to be discussed next.

Visual Element 3: Use spacing to identify appropriate groupings of visual elements.

Another way of signaling to the respondent that a certain set of words, numbers, or other symbols needs to be read and comprehended as a phrase or unit is to locate them close together. Similarly, separating them suggests that they are to be read as separate units. The fact that elements placed close together encourages comprehending them as a unit has been described by Gestalt psychologists as the *Law of Proximity* (Wallschlaeger and Busic-Snyder, 1992).

Figure 3.12 Change spacing and similarity to identify appropriate groupings of visual elements.

Problem:

1. Thinking about the last time you were enrolled in classes, were you primarily interested in obtaining a degree or certificate, or were you primarily interested in learning a new skill or both? Mark one answer: \triangle Obtaining a degree or certificate; \diamond Learning a new skill; \heartsuit Both a degree or certificate and learning a new skill. 2. What year were you last enrolled in classes? _____.

A revision:

- ① Thinking about the last time you were enrolled in classes, were you primarily interested in obtaining a degree or certificate, learning a new skill, or both of these?**

- ☐ Obtaining a degree or certificate
- ☐ Learning a new skill
- ☐ Both of these

- ② What year were you last enrolled in classes?**

_____ Year enrolled

Thus, adding an extra blank line between questions and putting less space between the subelements of each question (e.g., question and answer spaces) conveys to the respondent which answer choices go with which query. The first example in Figure 3.12 illustrates a lack of grouping between questions and among question elements. The revision separates the question number from the question stem and both from the answer choices by adding space. Question 1 is then separated from Question 2 by adding an additional space.

Visual Element 4: Use similarity to identify appropriate groupings of visual elements.

In much the same way that spacing tends to group visual information, use of similar elements (whether based on brightness, color, size, or shape) encourages elements to be seen together. This effect has been described as the *Law of Similarity* (Wallschlaeger and Busic-Snyder, 1992). Thus, switching from light to dark type, changing fonts, or introducing a different color of type signals to respondents that they are moving from one grouping of information to another. An additional factor that conveys a group identity in Figure 3.12 is shape. The first example uses different shapes ($\triangle\diamond\heartsuit$) to identify answer choices for Question 1, whereas the revision uses the same shape for all answer choices. In addition, the question stem and answer choices of the revision each exhibits a consistent rectangular shape which helps them to be seen

as separate but connected entities. Thus, four distinct elements, that is, changes in size, brightness, location, and similarity, have been manipulated to revise the problem question format in Figure 3.12.

Filling out a questionnaire is a cumulative learning experience. Respondents quickly learn that such things as question numbers are a certain size and shape and answer spaces (whether boxes or numbers that must be circled) are the same. Conveying the same type of information in the same way throughout a questionnaire aids respondents by making the response task easier, and can be expected to improve compliance with the prescribed navigational path. The fifth and sixth visual elements are concerned with capitalizing on this learning process to make the completion process as easy as possible for respondents.

Visual Element 5: Maintain a consistent figure/ground format to make the response task easier.

The respondent's ability to make sense of written information on a page depends upon being able to distinguish *figure from ground*. First, the eye must be able to separate letters and symbols from the background on which they are printed, a process that requires contrast. Changing the figure/ground format represents an additional way of grouping or separating information. This paragraph is written as black figure on white ground. Changing from black ink to red ink represents a change in figure/ground format. However, perhaps the most fundamental change is reverse print—going from black letters on a white background to white letters on a black background. For example, a black background is printed on the page so **the unprinted space becomes figure, that is, white letters**, as done here. It is difficult for the eye to quickly make the transition from one to the other. Figure 3.13 illustrates the intermingling of reverse and normal print along with inappropriate use of several other elements. The effect of reverse print is to increase the likelihood that respondents will skip words when reading sentences. The revision eliminates reverse print except in the question numbers, where their “brightness” helps the respondent navigate from one question to the next.

Visual Element 6: Maintain simplicity, regularity, and symmetry to make the response task easier.

An additional Gestalt principle, known as the *Law of Pragnanz*, states that figures with simplicity, regularity, and symmetry are more easily perceived and remembered than are unusual or irregularly shaped figures (Wallschlaeger and Busic-Snyder, 1992). If one chooses to use numbers for circling as a means of providing answers for the first question, it is undesirable to switch to boxes for the second question, filling in ovals or circles for the third question, etc. To do so would require additional, but unnecessary, learning by the respondent.

Figure 3.13 Maintain simplicity, regularity, symmetry, and a consistent figure/ground format to make respondent task easier.

Problem:

- 1) Do you personally drive a car or other vehicle to work? ☐ Yes ☐ No
- 2 Which of the these parking policies do you most prefer:
- () *pay each day*
 () *pay weekly*
 () *pay monthly*
3. Do you prefer a hanging parking sticker or a decal that attaches to the windshield of your car:
- ☐ prefer hanging sticker ☐ prefer decal on windshield

A revision:

- 1 Do you personally drive a car or other vehicle to work?
- ☐ Yes
☐ No
- 2 Which of the these parking policies do you most prefer?
- ☐ Pay each day
☐ Pay weekly
☐ Pay monthly
- 3 Do you prefer a hanging parking sticker or a decal that attaches to the windshield of your car?
- ☐ Prefer hanging sticker
☐ Prefer decal on windshield

The lack of simplicity, regularity, and symmetry is illustrated in several ways in Figure 3.13. Unusual shapes are included in front of Question 1 and again just before the answer choices, but are not used for the other questions. Answers to Questions 1 and 3 are horizontal, whereas those to Question 2 are vertically displayed. Italics and reverse print are also used irregularly. In addition, each question uses a different shape—lines, parentheses, and boxes—to identify answer spaces. The revision establishes simplicity and regularity.

The challenge in formatting questionnaire pages is less the systematic manipulation of any individual visual element than it is the simultaneous manipulation of all six at once. Figure 3.14 provides an example of a questionnaire page that intentionally ignores all six elements. A larger font is

Figure 3.14 A poorly constructed questionnaire page illustrating poor application of six construction elements and a revision.

A poorly constructed questionnaire:

Start Here

- 1) During the past 30 days, how often did pain interfere with your daily activities such as your job, working around the home, or social activities?

- | | | |
|----------------------------|-------------------------|---------------------|
| 1) All of most of the time | 3) A little of the time | 4) None of the time |
| 2) Some of the time | | |

2. During the past 30 days, how often have you had pain?

- 1) Every day
2) Between 4 and 6 days a week
3) Between 1 and 3 days a week
4) **Less than once a week**
5) **Never**

- 3) Did you feel any pain last week?

- ☐ yes
☐ no

- 4) Do you wear eye glasses?
Virtually all of the time.....1
Usually just to read.....2
Rarely.....3
Never.....4

- 5 Is it difficult for you to walk?

___ always ___ sometimes ___ occasionally ___ never

6. What is the furthest you could probably walk without sitting down and resting?

△ once across the room and back; △ several times across the room and back; △ up a flight of stairs; △ up several flights of stairs; △ further

7. When was the last time you were in a doctor's office?

- ☐ this week
☐ last week
☐ before that

8. On average, how often do you visit a doctor's office?

- a every week
b 2-3 times per month
C once a month
D less than once a month

(Continued)

used selectively in later questions and may draw attention there rather than to the beginning of the page (Element 1). The later questions also appear in darker (brighter) type which may draw attention away from the normal starting place on the page (Element 2). The answer categories of some items (e.g., Question 1) are connected to the succeeding question rather than to the one with which they belong (Element 3). Words that should be seen as a unit, such

Figure 3.14 (Continued)A revision:**Pain and How It Influences Daily Activities: A Research Study****START HERE**

- ① During the past 30 days, how often did pain interfere with your daily activities such as your job, working around the home, or social activities?**

- ☐ All of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

- ② During the past 30 days, how often have you had pain?**

- ☐ Every day
- ☐ Between 4 and 6 days a week
- ☐ Between 1 and 3 days a week
- ☐ Less than once a week
- ☐ Never

- ③ Did you feel any pain last week?**

- ☐ Yes
- ☐ No

- ④ Do you wear eye glasses?**

- ☐ Virtually all of the time
- ☐ Usually just to read
- ☐ Rarely
- ☐ Never

- ☐ Always
- ☐ Sometimes
- ☐ Occasionally
- ☐ Never

- ⑥ What is the furthest you could probably walk without sitting down and resting?**

- ☐ Once across the room and back
- ☐ Several times across the room and back
- ☐ Up a flight of stairs
- ☐ Up several flights of stairs
- ☐ Further

- ⑦ When was the last time you were in a doctor's office?**

- ☐ This week
- ☐ Last week
- ☐ Before that

- ⑧ On average how often do you visit a doctor's office?**

- ☐ Every week
- ☐ 2–3 times per month
- ☐ Once a month
- ☐ Less than once a month

as the answer choices to Question 2, are presented in dissimilar fonts so they appear as two groups (Element 4). The ground/figure format changes periodically throughout the page, making reading more difficult and the likelihood of skipping words greater (Element 5). There is virtually no regularity, simplicity (Element 6), or symmetry across questions, as evidenced by different symbols for identifying answer categories, irregular use of horizontal and vertical listing of answer choices, and several other features that can easily be identified.

Few surveyors would prepare a questionnaire page that is as poorly con-


constructed as the first example in Figure 3.14. However, this example can be contrasted with the revision including exactly the same questions. Here, the visual elements have been manipulated to clearly identify the appropriate starting point of the questionnaire. Elements that comprise each question and the total questionnaire have been appropriately grouped through application of spacing and similarity procedures, in addition to size and brightness manipulations. In addition, changes in figure and ground have been eliminated except for question numbers in which the combination of brightness and reverse print makes it clear which question is to be answered and in what order, thus defining the navigational path. Finally, the maintenance of simplicity and regularity give the page a look that allows the respondent to know exactly how to answer each question after only a quick glance.

The six visual elements that I have just described are the critical tools by which the creation of visual navigational guides can proceed. The remaining questionnaire construction principles in this chapter are based upon applications of these very important concepts. Each of these six visual elements is manipulated in a way that will encourage all respondents to adhere to the navigational path and correctly interpret the survey questions. Size and brightness were identified as ways of attracting people's eyes to a particular part of the questionnaire page. Spacing and similarity were presented as means of helping the respondent achieve appropriate groupings that assist with comprehension of the task. Maintaining simplicity and keeping a consistent figure/ground format were offered as means of making the response task easier and helping respondents achieve efficiency. In practice, individual effects of each element are broader. For example, a change in figure/ground may involve a change in brightness and function as a way of grouping information. I attempt to deal in a practical way with many of these cross functions in the construction principles that follow.

Principle 3.6: Begin asking questions in the upper left quadrant; place any information not needed by the respondent in the lower right quadrant.

The ideal place to start asking a question is near the upper left-hand corner of page 1. When presenting this admonition to a class, the immediate response from a student was, "Isn't that obvious?" Perhaps. Culturally, that is the page location people raised in western cultures typically look at when asked to start. Although it may seem obvious, I have seen many questionnaires that fill this space with agency sponsorship information, special instructions, agency processing information, the mailing address for returning the questionnaire, and other information that interferes with getting started. This is illustrated in Figure 3.15, which places agency information in this area to which the eye is naturally drawn. Such information may be important, but presenting it here in place of the information that a respondent is looking for, that is, how to get started, can only frustrate rather than facilitate the completion process.

Figure 3.15 Example of upper left quadrant of page not being used in most effective visual manner.

 U.S. DEPARTMENT OF COMMERCE BUREAU OF THE CENSUS FORM CB-8600		1992 CENSUS OF SERVICE INDUSTRIES MEMBERSHIP ORGANIZATIONS OMB No. 0607-0729; Approval Expires 06/30/94	
DUE DATE: FEBRUARY 15, 1993 If you have questions about completing this report, please call or write the Census Bureau. In any communication, be sure to refer to the 11-digit Census File Number (CFN) printed in the label to the right. Please return your completed report to: BUREAU OF THE CENSUS 1201 East 10th Street Jeffersonville, IN 47134-0001 Toll-free assistance, 8:00 a.m. to 8:00 p.m., eastern time, Monday through Friday: 1-800-233-6136		CB-8600	
Please read the accompanying instructions before answering the questions.			
Census use			
(Please correct any errors in name, address, and ZIP Code.)			
YOUR RESPONSE IS REQUIRED BY LAW. Title 13, United States Code, requires businesses and other organizations that receive this questionnaire to answer the questions and return the report to the Census Bureau. By the same law, YOUR CENSUS REPORT IS CONFIDENTIAL. It may be seen only by Census Bureau employees and may be used only for statistical purposes. Further, copies retained in respondents' files are immune from legal process.			
Item 1. EMPLOYER IDENTIFICATION NUMBER Is the Employer Identification (EI) Number shown in the label the same as the one used for this establishment on its latest 1992 Employer's Quarterly Federal Tax Return, Treasury Form 941? 094 1 <input type="checkbox"/> Yes 2 <input type="checkbox"/> No - Report current EI No. below (9 digits)		Item 4. LEGAL FORM OF ORGANIZATION AND TAX STATUS a. LEGAL FORM OF ORGANIZATION Which of the following best describes this establishment's legal form of organization during 1992? Mark (X) only ONE box. 003 1 <input type="checkbox"/> Individual ownership (sole proprietorship) 2 <input type="checkbox"/> Partnership 5 <input type="checkbox"/> Government - Specify _____ 0 <input type="checkbox"/> Corporation 9 <input type="checkbox"/> Other - Specify _____	
Item 2. PHYSICAL LOCATION a. Is this establishment's physical location the same as the address shown in the label? (P.O. box and rural route addresses are not physical locations) 093 1 <input type="checkbox"/> Yes 2 <input type="checkbox"/> No - Report physical location below Number and street City, town, village, etc. State ZIP Code		b. TAX STATUS (1) Is this establishment operated on a not-for-profit basis? 005 1 <input type="checkbox"/> Yes 2 <input type="checkbox"/> No - Skip to item 5 (2) Was all or part of the income of this establishment or organization exempt from Federal income taxes under section 501, 521, 527, or 528 of the Internal Revenue Code? 004 1 <input type="checkbox"/> Yes 2 <input type="checkbox"/> No	
b. Is this establishment physically located inside the legal boundaries of the city, town, village, etc.? 095 1 <input type="checkbox"/> Yes 3 <input type="checkbox"/> No legal boundaries 2 <input type="checkbox"/> No 4 <input type="checkbox"/> Do not know			

From 1993 Census of Service Industries, conducted by the U.S. Bureau of the Census.

The place that the respondent's eye is least likely to be attracted to on the page is the lower right-hand quadrant. That makes this area the ideal space for "agency use only" information or other material not essential to the response task, in contrast to its prominent position in Figure 3.15.

Principle 3.7: Use the largest and/or brightest symbols to identify the starting point on each page.

Another way of helping to assure that people start at the beginning is to use the visual concepts of largeness and/or brightness to help draw their atten-

tion to the starting point. This can be accomplished by something as simple as printing the question number in a larger font than that used on the rest of the page, and perhaps putting it in reverse print, for example, ❶. Similarly, printing the words **START HERE** in a larger bold print provides an even more dramatic and powerful attraction. Frequently, I have been shown questionnaires in which the writers attempt to keep people from overlooking an instruction in the middle of the page by placing it in particularly large print. The undesirable result is to attract respondents immediately to that portion of the page, with the result that earlier questions may get skipped. Adhering to the principle of starting to list questions where the respondent expects to find them decreases the need for using highly prominent visual guides (e.g., brightness or font size) in order to draw the respondent's attention to a starting point elsewhere on the page.

Principle 3.8: Identify the beginning of each succeeding question in a consistent way.

Observation of many people as they fill out questionnaires suggests a common response behavior. Once an answer box is checked, the respondent immediately searches for the beginning of the next question. Identifying each question in the same way, and somewhat prominently (through spacing and/or the use of reverse print numbers shown for the revision in Figure 3.14), helps respondents navigate accurately through the questionnaire. These numbers are the most important navigational guides on the page. They persist throughout the questionnaire as a way of keeping respondents on task. For these reasons, we typically reserve the use of reverse print, or enlarged numbers, for this purpose.

Principle 3.9: Number questions consecutively and simply, from beginning to end.

People expect questions to be numbered consecutively with numbers or letters, but not both. Yet questionnaires sometimes identify questions by combinations of letters and numbers, for example, P1, P2 . . . H1, H2, and H3. The explanation in this case was that within this agency P referred to population characteristic questions and H referred to housing questions. The question order meant that P came before H. This cumbersome procedure was intended to convey agency division responsibilities. This information was not only irrelevant to respondents, but potentially confusing as well; numbering questions as 1, 2, 3, 4 and so on to the end would have been easier for respondents to follow.

A tendency also exists, especially among government agencies and other large organizations, to divide questionnaires into sections. One result that is particularly confusing to respondents is to label sections as 1, 2, 3, 4, etc., and then to renumber questions within each section as 1, 2, 3, 4, etc. The result is that small numbers appear repeatedly throughout the questionnaire and re-

spondents become confused when they lay down a questionnaire and attempt to go back to it later. Another unfortunate practice, and one usually aimed at overcoming the foregoing problem, is to develop number schemes such as IA2a, IA2b, IA2c, etc., so that each question is preceded by several symbols instead of one. Such numbering schemes violate the principle of maintaining simplicity and are difficult for people to comprehend. Listing questions in the natural sequence understood by respondents, for example, 1, 2, 3, 4, 5, etc., from the first page to the last is helpful to them.

An unfortunate objection I often hear to the use of a straightforward sequential numbering system is, "If the respondents know how many questions are really being asked, they will think the questionnaire is too long and quit filling it out." I know of no evidence to support this viewpoint. In addition, this type of reasoning suggests trying to get response through "trickery," an attitude which runs contrary to the type of respondent perception that should be fostered with the use of a straightforward, social exchange-oriented approach to respondents. In addition, there are many other indicators of questionnaire length and complexity—including number of pages of paper—that make such attempts at deception quite obvious and place the surveyor in a bad light at a time when trust is most needed.

A simple numbering system also helps make skip pattern instructions easier to follow. Recently I was shown a questionnaire with a high proportion of erroneously followed skip patterns. The problem soon became obvious—respondents who said yes to a particular question were instructed to go to Section Four, which started several pages later. Most turned the page, saw Question 4 of Section 3, and proceeded to try to answer it. Thus, even when sponsors deem it necessary to divide a questionnaire into labeled sections, it is desirable to number questions consecutively throughout the questionnaire so that, for example, Section 1 might have Questions 1–10 and Section B would then start with Question 11. In addition, using letters instead of number to identify sections (e.g., A for 1) is also helpful for distinguishing sections from question items.

Principle 3.10: Use a consistent figure/ground format to encourage the reading of all words.

As noted earlier, the terms figure and ground are used to refer to a relationship between foreground and background. On this page black letters (figure) appear on a white background. When changes are made in the relationship between the letters one is supposed to read and **the background** on which **THEY ARE PRINTED**, sentences are more difficult to read and comprehend. A consistent figure/ground relationship, a characteristic virtually mandated in the era of typewriters, allows one to concentrate on the meaning of the words rather than having to pause in order to adjust to the change.

Recently I was asked to fill out a questionnaire in which the contractor had

been asked to add color. In one question, six of seven answer choices were printed in black ink (figure) on white paper (ground), while the other answer choice was printed in red ink. The visual impact was for the first six choices to be seen as the range of alternatives, and for the seventh answer choice to be missed altogether. The most difficult ground/figure or change to negotiate, as described earlier, is a complete reversal, that is, “reverse print.” It is very difficult to read words inserted into a phrase or sentence that appear in reverse print. As a result, they are more likely to be skipped. In general, the use of reverse print in sentences or among answer categories should be avoided.

In order to attract attention to notes or special instructions, surveyors sometimes place instructions in reverse print or in boxes. Because people can read prose fastest when it is written in a consistent figure/ground format, and it takes most people a second or so to make the transition, they are inclined to skip over such instructions to the next information that appears in a consistent figure/ground format. For these reasons, and in spite of the fact that reverse print is powerful as a means of grouping words and phrases together, reverse print is not used within a question either to emphasize a portion of it or to separate a question subcomponent from some other part (e.g., answer choice categories).

Principle 3.11: Limit the use of reverse print to section headings and/or question numbers.

Respect for the power of reverse print that stems from the introduction of “brightness” leads me to use it sparingly, in a way that is most helpful for defining the respondent path and maintaining adherence to it. Thus, its use is limited to question numbers (as in Figure 3.14) and occasionally section headings.

If proposed section headings consist of only a few words, reverse print is appropriate for this purpose. The main assistance it provides to respondents when used for headings is at the comprehension stage of preattentive processing, when the respondent is perusing the questionnaire to see what it is about and how it works. When actually filling out a questionnaire, the mental switch required for reading any reverse print means that the words are likely to be skipped. Skipping over the exact question number, on the other hand, is less important inasmuch as a well-designed questionnaire will have questions located in a vertical format so that location and the presence of a bright (black) spot on the page is all the respondent needs to maintain correct navigation.

Principle 3.12: Place more blank space between questions than between the subcomponents of questions.

As noted at the beginning of this section, there are two levels of grouping inherent to all questionnaires. First, questions need to be separated from one another. Second, within questions there are meaningful subcomponents that

need to be viewed and comprehended separately; in particular, the query itself needs to be distinguished from the answer choices. One way of making these distinctions clear is by the amount of spacing. More space should be allowed for separating questions from one another than for separating subcomponents of questions from one another, as was shown in Figure 3.12. Yet a certain amount of spacing between the subcomponents of questions—provided by leaving a blank line and indenting the list of answer categories—helps clarify each subentity of the question. Distinctions between questionnaire subcomponents can also be accomplished by other means, such as the brightness variations discussed below.

Principle 3.13: Use dark print for questions and light print for answer choices.

One method of conveniently separating the actual queries from answer choices is to print the former in bold and the latter in regular print (as shown in Figure 3.12). Doing the opposite, that is, using bold letters for answers, is avoided because of the desire to use the brightest symbols for the part of the question that should be read first. In this regard the sequence of reverse print for numbers, bold for query, and normal print for answer choices provides a desirable hierarchical gradient of most to least brightness that is consistent with the flow of each question. In the first edition of this book, I recommended accomplishing this distinction by using capital letters for answer choices. The computer age provides the contrast of bold and regular print as a much better alternative inasmuch as capital letters are more difficult to read than are lower case letters.

Principle 3.14: Place special instructions inside of question numbers and not as free-standing entities.

Frequently it is necessary to provide a special instruction to clarify a question. This leads to the undesirable practice of placing instructions outside of the question number sequence and emphasizing them by boxes or perhaps a different color of ink.

Once people have gotten into the routine of completing a questionnaire, the marking of an answer leads to the immediate search for the next question, which we usually identify by a question number in reverse print. As a result, free-standing instructions tend to be skipped entirely. Further, attempting to bring attention to them with emphases such as a change in ink, boxes, shaded backgrounds, or reverse print increases the likelihood that they will be skipped. To the extent possible, we insert instructions within numbered questions rather than outside (as shown in the revision in Figure 3.16). Such instructions are most likely to be read if they are expressed as part of the query itself rather than being placed in italics, parentheses, or a separate paragraph. However, exceptions are often made, as explained next.

Figure 3.16 Write special instructions as part of question, not as free-standing entities.

Problem:

- ⑤ How many months have you worked in your current job?**

_____ Number of months

Please be as specific as possible in answering the next question, including any area of specialization. Example: High school teacher–Math. If you had more than one job, answer for the job for which you worked the most hours.

- ⑥ What kind of work do you do in your current job?**

_____ Kind of work

A revision:

- 5. How many months have you worked in your current job?**

_____ Number of months

- 6. What kind of work do you do in your current job? Please be as specific as possible in answering. Include any area of specialization, for example: “High school teacher–Math”. If you had more than one job, answer for the job for which you worked the most hours.**

_____ Kind of work

Principle 3.15: Optional or occasionally needed instructions should be separated from the question statement by font or symbol variations.

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. If they are required to read through a great deal of material that does not apply, or can be skipped without negative consequences, the habit of skipping words and phrases is encouraged. For these reasons a distinction is made between words that are essential for every person to read and those that may be needed by only some respondents. There are many different reasons that reading a particular instruction may be optional. Perhaps it is because the instruction, “Put an X in the appropriate box,” is the same instruction used for a previous question and many respondents will remember that. It may also be that only a few respondents need the information: “If you worked last week but have since quit your job, then you should... .” To avoid presenting information that respondents already know, or that applies to relatively few of them, we distinguish this infor-

mation from the query either by the use of italics (as shown for the revision in Figure 3.16) or a symbol variation, such as putting it in parentheses.

Principle 3.16: Do not place instructions in a separate instruction book or in a separate section of the questionnaire.

Sometimes surveyors decide that respondents need detailed instructions and place those instructions in a separate booklet. This decision usually has four significant consequences. First, it greatly increases the likelihood that respondents will ignore the instructions. Second, a detailed review of nearly 20 instruction books, mostly for federal government surveys, led to the conclusion that instruction booklets have a tendency to expand greatly with minute instructions that are rarely, if ever, needed by anyone, thus compounding the respondent's difficulty in finding answers to a particular question. The third consequence is to make the task of responding appear far more difficult than it is so that the questionnaire gets laid aside. Finally, this practice leads to the questionnaire construction tendency of substituting abbreviations, incomplete sentences, and one- or two-word headings for questions in order to make the questionnaire as short as possible. The overall effect of creating separate instruction books is to greatly increase the difficulty of completing the questionnaire.

In many respects, a frustrating cycle is fostered by the decision to create an instruction book. That decision often leads to writing a more cryptic questionnaire that forces more dependence on the instruction booklet. The result is an instruction booklet that is even longer and more difficult to use, a topic I will revisit in the discussion of business surveys in Chapter 10. For these reasons it is important to keep instructions within the questionnaire.

A strong recommendation I made to a surveyor for combining instructions with questions once led to the decision simply to print the instruction book as the last half of the questionnaire booklet. This, too, is undesirable and may be even more frustrating for respondents as they flip pages back and forth without being able to see both of the critical pages at the same time. The solution I favor is to place all of the necessary instructions after the question, where the respondent needs them, as will be discussed in particular for business surveys in Chapter 10. This placement usually results in instructions being shortened. If instructions are quite lengthy and necessary for most questions an alternative is to place all survey questions on the right side of two facing pages and the relevant instructions on the left page, so that a respondent who needs instructions for a particular question can look directly to the left page, at the same level, and find them.

Principle 3.17: Use of lightly shaded colors as background fields on which to write all questions provides an effective navigational guide to respondents.

Many questionnaires, especially those used in large-scale surveys, must include sponsorship information, space for coders, directions on who to call

with questions, and other processing information. This information may be unimportant to most respondents and interfere visually with identifying and answering questions.

Based upon considerable research and experimentation at the U.S. Bureau of the Census (Dillman, Sinclair, and Clark, 1993; Jenkins and Dillman, 1995, 1997), an effective method of helping respondents to find and follow the prescribed navigational path is to use lightly shaded background fields on which are printed every question that must be answered, while information not within the navigational path is printed elsewhere on a white background. One of the main effects is to keep respondents from getting lost or confused by the considerable amount of information deemed necessary for inclusion in the questionnaire by federal law and/or processing needs. Typically these fields are printed in 15–30% of full color and can also be done in gray versions (as illustrated in Figure 3.10). The use of such fields is further justified by a related technique for identifying answer spaces, as discussed next.

Principle 3.18: When shaded background fields are used, identification of all answer spaces in white helps to reduce item nonresponse.

It has also been found through Census Bureau research (Jenkins and Dillman, 1997) that adding white spaces wherever answers are to be provided improves response to questionnaires that use colored background fields. The white spaces provide an additional navigational guide, in effect communicating that wherever there is a set of white spaces an answer should be provided (as illustrated in Figure 3.17). It appears that one of the main beneficial effects of white answer spaces is to reduce item nonresponse, but this effect has not yet been tested experimentally.

There is considerable latitude available for choosing the particular color for use as a background field. There is little evidence to suggest that color by itself has a significant impact on response rates, although it is desirable to avoid colors that test respondents believe tire their eyes, such as a bright and glossy chartreuse or pink. The color that is chosen should have wide tint variation available, and pastel colors should be avoided. Using such a color makes it possible to use the full (or nearly full) color for navigational guides (e.g., number or section heading backgrounds) and a partial screening (e.g., 20% of full color) for the background color against which black letters show clear contrast, but dark enough to allow white answer spaces to be easily discerned. Optical scanning, to be discussed in Chapter 12, may impose further printing requirements.

Principle 3.19: List answer categories vertically instead of horizontally.

Response categories for closed-ended questions can be arranged vertically or horizontally. In general, arranging answer choices vertically is preferred because it allows all choices to be listed in a single line down the page, as shown

Figure 3.17 Example of colored background fields with white answer spaces for guiding respondents.

SECTION 1
IMPLEMENTATION OF THE SCHOOL MEALS INITIATIVE (SMI)

1. How many public School Food Authorities (SFAs) within the state are currently participating in child nutrition programs?
(Record number of SFAs. If none, enter "0".)

Number of public SFAs participating in child nutrition programs

2. Of the total number of public SFAs within the state participating in child nutrition programs, how many are currently using each of the following menu planning options? *(Some SFAs can be using more than one menu planning system. The total number of menu planning options in use might therefore exceed the total number of SFAs in the state; see Glossary, page 11. If none, enter "0".)*

Number of public SFAs currently using:

Nutrient Standard Menu Planning (NuMenus)	<input style="width: 50px;" type="text"/>
Assisted Nutrient Standard Menu Planning (Assisted NuMenus)	<input style="width: 50px;" type="text"/>
Enhanced Food-Based Menu Planning	<input style="width: 50px;" type="text"/>
Traditional Food-Based Menu Planning	<input style="width: 50px;" type="text"/>
Other <i>(Please specify below.)</i>	
<input style="width: 100px;" type="text"/>	<input style="width: 50px;" type="text"/>
<input style="width: 100px;" type="text"/>	<input style="width: 50px;" type="text"/>
<input style="width: 100px;" type="text"/>	<input style="width: 50px;" type="text"/>

3. What role did your Agency play in assisting public SFAs in the selection and implementation of new menu planning systems during the last school year (1997-98)?

Did your Agency, or someone working on its behalf (e.g., contractors), provide public SFAs with:

3a. Assistance in training sessions? *(Mark [x] one box.)*

☐ Yes
☐ No (SKIP TO Q.3b)

→ What level of assistance was provided during the 1997-98 school year? *(Record number for each item. If none, enter "0".)*

3a.1 <input style="width: 50px;" type="text"/>	Number of training sessions assisted
3a.2 <input style="width: 50px;" type="text"/>	Number of public SFAs represented
3a.3 <input style="width: 50px;" type="text"/>	Number of public SFA staff attending

3b. Nutritional expertise either directly or through an outside organization? *(Mark [x] one box.)*

☐ Yes
☐ No

3c. Computer expertise either directly or through an outside organization? *(Mark [x] one box.)*

☐ Yes
☐ No

3d. On-site technical assistance? *(Mark [x] one box.)*

☐ Yes
☐ No (SKIP TO Q.4, PAGE 2)

→ What level of assistance was provided during the 1997-98 school year? *(Record number for each item. If none, enter "0".)*

3d.1 <input style="width: 50px;" type="text"/>	Number of on-site visits
3d.2 <input style="width: 50px;" type="text"/>	Number of SFAs visited

From USDA National Nutrition Administration Survey, conducted by the Gallup Organization.

in most of the figures in this chapter. Placing answer choices horizontally often means that categories and boxes must appear on more than one line, and it is harder for a respondent to get a sense of linear connection, which is important when an extensive set of vague quantifiers (e.g., seven choices ranging from very strongly agree to very strongly disagree) is used. In addition, careful spacing is needed for horizontally-arranged categories so the answer boxes will not be mistakenly associated with the wrong category. This preference for vertically-arranged categories also stems from a related preference for writing questions in narrower columns (to encourage more complete reading of words), and from wanting respondents to continuously move vertically down each page to get a sense of progress towards completion.

In some instances a scale is provided as part of a visual layout (shown under Principle 22 below) and may be difficult to show vertically. Also, in replication studies that have used horizontal scales or category arrangements, it is important to replicate the visual component as well as the word content of the original questionnaire. In these instances it is appropriate to deviate from the normal preference for vertically aligned categories.

Principle 3.20: Place answer spaces consistently to either the left or right of category labels.

A strong argument can be made for placing answer spaces on either side of a vertically listed set of categories. Placement on the left is supported by the recognition that the amount of space needed for writing answers varies considerably. This method also allows all answer boxes to be arranged easily in a vertical column, which helps avoid item nonresponse. This format tends to group answer choices closer to the left side of the page (prime navigational space) and leave considerable space on the right side of the page for skip instructions and the use of an indentation convention for questions that some respondents were instructed to skip (see Principle 3.26). This method was recommended as part of the original TDM.

It has also been observed (Jenkins and Dillman, 1995) that when the box is on the left, respondents accustomed to the typical pattern of checking an answer box and looking for the next question number or navigational guide tend not to see any skip instructions that are placed to the right of the response label. This is understandable because of the typical two degree focus of one's eyes when reading attentively (Kahneman, 1973). Increasingly, I have found this argument for placing answer boxes on the right to be convincing. Placement on the right is also supported by the argument that the respondents' hand does not cover any of the answer choices when marking their answers, so they have a clearer view of the connection between the box and answer. However, doing so requires the use of dotted leaders in order to get vertical alignment of the starting point for each answer choice as well, or uneven left justification as also shown in Figure 3.18. Extensive cognitive interviews have

demonstrated that placement of answer boxes on the right goes mostly unnoticed by respondents, and therefore seems quite workable (Dillman, Carley-Baxter, and Jackson, 1999).

One solution to deciding whether a left or right box method should be used is to preview the kinds of questions to be included in a survey and contemplate the total array of questions to be asked, including the length of answer categories and whether skip instructions are going to be a problem. It is also important to contemplate whether a one- or two-page column format is going to be used. The latter format leaves less room for skip instructions when answer choices are on the right. Often the case for choosing one over the other is not compelling. Definitive research has not yet been done, so personal preference may be an appropriate way to decide. Decisions on this issue, which require bringing together a number of considerations, are characteristic of the many ways in which tailored survey design differs from the original TDM.

Principle 3.21: Use numbers or simple answer boxes for recording of answers.

There are many ways of asking people to record their answers; for example, circle category numbers or letters, check short black lines, fill in small circles or ovals, put an X in boxes, or put a mark between brackets or parentheses. In the first edition of this book, circling was recommended because the technology of the times (typewriters) did not provide an easy method for making boxes, and the nature of circling helped achieved a clearer demarcation of which answer was chosen than did check marks. In addition, numbers provided a convenient precoding system.

The circling of numbers remains an effective way of getting people to record their answers. However, asking people to place X's in boxes is becoming more desirable for most questionnaires. Increasingly, questionnaires are likely to be scanned into computers rather than keypunched, as discussed in Chapter 12. Instructing people to mark X's is more desirable than having boxes marked with a check (✓). Placing an X in a box requires two separate actions (as opposed to a check mark) and is more constrained than a check mark in the way that most people make such marks. Thus, an X is less likely to go as far outside of a box as a check mark, which can then interfere with the optical reading of other boxes. Evidence also suggests that whereas a rectangle encourages respondents to make check marks, boxes encourage individuals to make X's (Caldwell and Dillman, 1998).

Principle 3.22: Vertical alignment of question subcomponents among consecutive questions eases the response task.

In the first edition of this book, considerable emphasis was placed on aligning answer spaces down the page in order to decrease the possibility of answers being missed. Vertical alignment gave a clear navigational signal for where a person was supposed to go on each page in order to complete a questionnaire.

Figure 3.18 Answer categories may be placed to the left or to the right of category labels.

Answer categories in the conventional left-side position:

1. To get to work do you normally ride a bus?
- ☐ Yes
☐ No → (Skip to Question 5)
2. (If Yes) Which one of the following best describes why you normally ride a bus?
- ☐ It's convenient
☐ It costs less
☐ I can't find anyone to car pool with
☐ I do it for environmental reasons
☐ Some other reason (Please explain)

Answer categories in the right-side position:

1. To get to work do you normally ride a bus?
- Yes ☐
No ☐ → (Skip to Question 5)
2. (If Yes) Which one of the following best describes why you normally ride a bus?
- It's convenient ☐
It costs less ☐
I can't find anyone to car pool with ☐
I do it for environmental reasons ☐
Some other reason (please explain) ☐

Answer categories placed in a different style so dotted leaders are not required:

2. (If Yes) Which one of the following best describes why you normally ride a bus?
- It's convenient ☐
It costs less ☐
I can't find anyone to car pool with ☐
I do it for environmental reasons ☐
Some other reason (please explain) ☐

That benefit to vertical alignment remains. Giving attention to how question numbers, the beginning of questions, and presentation of answer choices line up makes it easier for respondents to grasp quickly the entire task before them. Differences in vertical alignment, for example, indenting five to ten spaces from the stem of the question for listing answer categories, are powerful tools for achieving appropriate grouping of question subcomponents (see Visual Element 4).

Principle 3.23: Avoid double or triple banking of answer choices.

Frequently, questionnaire designers must formulate questions that require listing many answer choices and decide to save space by double or even triple banking answer spaces, that is, listing the choices in two or three columns below the query, as shown in Figure 3.19. This practice should be avoided. On closed-ended questions with ordered categories, the scalar nature of the question and the task of placing oneself at the most appropriate point on the scale can easily get lost. For closed-ended questions with unordered categories, the consequences may be worse. Respondents who have been “trained” in earlier questions to pick from among a one-column list of vertically arranged categories are likely to be visually attending to the answer procedures they have become used to and may fail altogether to notice the additional columns of possibilities that lie outside their attentive field of vision (see Figure 3.19).

Figure 3.19 Avoid double or triple banking of answer choices.

Problem:

1. Which one of the following best describes how this land is now used?

- | | | |
|--|-------------------------------------|--|
| <input type="checkbox"/> Farm Land | <input type="checkbox"/> Vacant Lot | <input type="checkbox"/> Pasture |
| <input type="checkbox"/> Play Field | <input type="checkbox"/> Wetland | <input type="checkbox"/> Equipment storage |
| <input type="checkbox"/> Nature Preserve | <input type="checkbox"/> Forest | <input type="checkbox"/> Other |

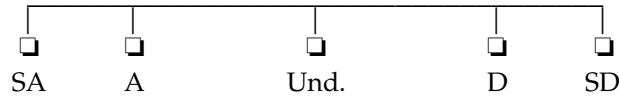
A revision:

1. Which one of the following best describes how this land is now used?

- ☐ Farm land
 - ☐ Vacant lot
 - ☐ Pasture
 - ☐ Play field
 - ☐ Wetland
 - ☐ Equipment storage
 - ☐ Nature preserve
 - ☐ Forest
 - ☐ Other
-

Principle 3.24: Maintain spacing between answer choices that is consistent with measurement intent.

Recently I was shown a questionnaire in which answers for an agree / disagree question were provided in this manner:



Spacing between answer spaces conveys information about what a category means, just as category order and response labels convey information. This is especially true for scalar questions that use vague quantifiers for answers. In this case, the spacing suggests that the “undecided” portion of this scale is greater than that of the other categories and thereby encourages more people to use that answer choice.

Sometimes questionnaires have some scales that extend nearly across the page and others that are confined to less than a third of the page. Similarly, I have observed scales that extend well across a page horizontally, whereas later in a questionnaire the same scale appears vertically, using a single-spaced format, so that the distance is significantly reduced. It is well established that people respond to more than the words used to write questions, and visual layout is an important part of that additional information (Schwarz, Hippler, Deutsch, and Strack, 1985; Rockwood et al., 1997).

The likelihood of such variations existing in a questionnaire increases when people are trying to “squeeze” questions onto a page, and when items used in some other study are borrowed for a new study. The latter situation often creates a dilemma between maintaining format and spacing used in another survey to which comparisons are to be made, and maintaining consistency with other items in the new questionnaire. Although few controlled experiments have examined the magnitude of such effects, the safe course of action is to strive for consistency across pages and sections of questionnaires, except when cross-study comparisons require that a particular visual representation of a scale need be replicated.

Web questionnaires, to be discussed in Chapter 10, provide a special challenge in this regard. The distances between horizontally displayed categories may vary depending upon the choice of screen configuration and whether questions are viewed in a partial-screen (tiled) format.

Principle 3.25: Maintain consistency throughout a questionnaire in the direction scales are displayed.

In general, it is advisable to keep scalar answer categories running the same direction throughout a questionnaire. The direction, that is, negative to positive or vice versa, is less important. The reason for being concerned with con-

sistency is that learning goes on with early questions and people attempt to use that learning to ease the task of responding to later questions. Two circumstances make this consistency especially important. One is when the label for the first response categories shares the same modifiers as the last response categories on a symmetrical scale, as shown in the first example of Figure 3.20. The distance between the answer box and the answer category (i.e., the word agree or disagree) is greater than 8–10 characters, which is the typical breadth of vision during attentive processing. This makes it easy to check the wrong answer without being aware of it. The second circumstance is when answer categories are displayed above a series of many rows of boxes so that the labels are printed some distance from the answer boxes, as shown in the second example included in Figure 3.20. Staying consistent in the direction that scales run helps avoid respondents' checking a choice towards one end of the scale when they meant to choose the other.

It is sometimes argued that by randomly reversing the order of answers throughout a questionnaire, as in this sequence—yes to no; strongly oppose

Figure 3.20 Consistently run scales in one direction (e.g., negative to positive) throughout the questionnaire to overcome separation of answer box from descriptor.

Potential Problem:

1. To what extent do you agree or disagree with this statement: "This company pays fair wages to its employees."

- ☐ Strongly Agree
- ☐ Somewhat Agree
- ☐ Neither Agree nor Disagree
- ☐ Somewhat Disagree
- ☐ Strongly Disagree

2. To what extent do you favor or oppose each of these proposals?

	Strongly Favor ▼	Somewhat Favor ▼	Somewhat Oppose ▼	Strongly Oppose ▼
Requiring all employees to complete travel expense forms on-line	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Being allowed to use sick leave when a family member is ill and needs care	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Working four, 10-hour days per week	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

to strongly favor; high priority to low priority; poor to excellent; no to yes; yes to no—that respondents are encouraged to think more carefully about their answers and are less influenced by the direction of the response layout. I have not seen convincing evidence that this is the case. Instead, this practice appears to lead to respondents having to concentrate more on how to respond correctly than on the substance of each question.

Principle 3.26: Use shorter lines to prevent words from being skipped.

People tend to read text unevenly. As noted earlier, the human eye can focus attentively on about eight to ten characters of text at once. In an effort to be efficient, many readers start and finish reading lines inward from the margins and their eyes move across the page quickly, reading only the text that seems necessary to give meaning to the information. Sometimes crucial words that change the entire meaning of a sentence, such as “not,” get missed. Longer lines present a greater challenge to readers because of the need to stay on the correct line as the eye moves across the page, and they seem somewhat harder to comprehend evenly. Thus, it is desirable to use shorter, as opposed to longer, lines for the construction of most questionnaires. This concern undergirds the preference for using two column formats on 8½" × 11" pages (as shown earlier in this chapter) instead of running questions across the entire page.

STEP 3: DEVELOP ADDITIONAL VISUAL NAVIGATIONAL GUIDES, THE AIM OF WHICH IS TO INTERRUPT ESTABLISHED NAVIGATION BEHAVIOR AND REDIRECT RESPONDENTS.

It is now time to turn to the task of partially undoing the consistency in visual navigation that has been established. The foregoing principles of page construction have been aimed at helping respondents get into a pattern of reading and responding that they can count on. Visual elements are used consistently in these principles to make the response task easier, resulting in fewer unintentional response errors. However, as respondents come to depend upon this consistency, the risk of their ignoring critical changes in the expected patterns of response increases. Such changes might include changing a time reference from what has happened in the last month to the last year. Another significant change occurs when, instead of being asked to answer every question, respondents are asked to skip one or several questions based upon which answer they select for a particular question.

The visual tools available for drawing attention to change in pattern are the same as those used for building the design consistency described in the preceding paragraphs. They include the now familiar changes in size, brightness (or color), spacing and location, similarity, a consistent figure/ground relationship, and simplicity. However, interrupting a routine behavior requires

somewhat different uses of these concepts and, as I shall argue, a more intense expression than is needed when a respondent is first trying to figure out what the questionnaire is about and how to answer it. For this reason, the challenge of interrupting and redirecting routine respondent behavior is specified as the third and final step of the page construction process.

Principle 3.27: Major visual changes are essential for gaining compliance with skip patterns.

Few aspects of questionnaire construction are more frustrating than getting respondents to follow all skip patterns correctly. The difficulty of getting correct compliance in self-administered questionnaires has sometimes led to the undesirable practice, described at the beginning of this chapter, of writing as many questions as possible so that they apply to everyone. For example:

Q23. If you work out every day, about how many minutes per day do you do that?

_____ Minutes per day

The problem associated with this type of question format is that the sponsor does not know whether to interpret a nonresponse as indicating that respondents do not work out each day, or whether they simply chose not to answer the question. To avoid this ambiguity, we follow the principle of writing every question in a way that each respondent should provide an answer, or be directed to skip it following their answer to a prior question.

Figure 3.21 shows a skip instruction supplied in a traditional way, as I have seen it used in many questionnaires. Here, respondents who give a “No” answer are instructed to skip four items and go to Question 28, while a “Yes” respondent is expected to continue with Question 24. Unfortunately, the instructions (Go to 28) and (Go to 24) are identical except for one of the eight characters required to write each of these directions, and the tiny distinguishing characteristic of these instructions, 4 versus 8, does not get expressed until the next-to-last character of print.

An improvement to these instructions would be to change them so that the instructions are of different sizes (number of characters), and the idea of skipping versus continuing is introduced as in (Skip to 28) versus (Continue with 24). However, even this change does not assure compliance. A possible explanation is that respondents are so deeply into the routine of reading a question, selecting an answer choice, and then checking the answer box as they start to seek the next question that the instructions that are beyond their narrow field of attentive vision get missed.

Several visual techniques can be considered for use individually or in combination to reinforce the message that continuing with the next listed question

Figure 3.21 Use multiple visual elements in concert to improve skip pattern compliance.

A traditional skip pattern which works poorly:

23. Normally, do you work out every day?
- ☐ Yes (Go to 24)
- ☐ No (Go to 28)
24. About how many minutes per day do you work out?
- _____ Minutes per day

Addition of four visual elements: Differently shaped directional arrows, word changes, larger font, and redundant instruction to define skip pattern:

23. Normally, do you work out every day?
- ☐ No → (Skip to 28)
- ☐ Yes
- ↓
24. (If Yes) How many minutes per day do you usually work out?
- _____ Minutes per day

Addition of alternative visual elements: Change in location of response boxes with spacing change for screened question to define skip pattern:

23. Normally, do you work out every day?
- No . . . ☐ → Skip to 28
- Yes . . . ☐ ↓
24. (If Yes) How many minutes per day do you usually work out?
- _____ Minutes per day

is not automatic for all respondents. As illustrated in Figure 3.21, the concepts and the visual elements being manipulated are as follows:

- Introduce a directional arrow (new shape) aimed at the path the respondent should follow.
- Introduce a different directional arrow (another new shape) to point to another path that respondents can choose to follow.
- Increase the font size of the skip directions to attract the respondent's attention (size change).

- Repeat the qualifying answer in parentheses, such as, “(If Yes)” at the beginning of the question that some respondents are being directed to skip (redundant verbal instruction).
- Reverse the locations answer box and category description so that the box is located nearer to the skip instruction (change in location).
- Indent the screened question so that only questions that are to be answered by everyone start at the left margin of the page (change in location).

Directional arrows have a widely understood cultural meaning indicating the reader should go to the direction the arrow points. They are therefore particularly effective. Increasing font size is, on the other hand, a more extreme measure, competing with other material for attention, and I prefer not to use it unless pretests suggest other procedures are not strong enough. Placing a verbal instruction before the next question as a reminder of who should complete it does not interfere with other aspects of visual design and is highly compatible with instructions contained in the screen question itself.

Changing the placement of answer boxes from the left side of category descriptions to the right side so that the skip instructions are within a few characters of the answer box is quite consistent with our understanding of the narrowness of the span of attentive vision. People are more likely to see the skip instruction as they check the box. However, placing answer boxes after the answer choices and before the directions is also difficult when skip instructions are long, particularly when a two-column page construction technique is being used. Thus, I have sometimes found this method difficult to use.

Recent research aimed at testing the effects of changing from a procedure similar to the traditional format (shown in Figure 3.21) to the format using either right or left boxes revealed that both offer substantial reductions in not skipping when directed to do so (Redline, Dillman, Smiley, Carley-Baxter, and Jackson, 1999). That research project, which used somewhat different combinations of elements than those displayed here, also revealed that the tendency to make skip-pattern errors varies greatly by type of question (Dillman, Redline, and Carley-Baxter, 1999).

Undoing the routine behavior of answering every question is sufficiently difficult that I do not believe using only one of the visual techniques is as effective as using two or more of them simultaneously. In addition, to a considerable extent, the above visual tools lend themselves to being used in combination, as illustrated in Figure 3.21. If a questionnaire contains many questions that involve skips, learning to follow them becomes a part of what a respondent must grasp early on in order to answer the questionnaire. I would probably refrain from using more than two or three of the above techniques (directional arrows, verbal redundancy, and indentation of the

screened question) in combination. However, for isolated skips that break a well-established pattern, increases in font size would also be considered.

Sometimes the substance of questions alone is enough to get people to follow skip patterns correctly. If the first question asks whether a person owns a car, and the follow-up question asks what color the car is, respondents will immediately know if they have not followed a skip direction correctly. In a question like the one listed above, it may be less clear whether a question is to be answered. Also, when many skip patterns, some of which involve several questions, are linked, respondents may get hopelessly lost. Such questionnaires require that careful attention be given to building in redundancy between graphics and words.

A second tendency I have noted in discussions with questionnaire designers is for many of them not to be concerned about whether skip patterns are followed correctly. They argue that the usual error is to answer questions that do not need to be answered. However, the frustration of answering questions that “don’t seem to apply to me” can lead to people stopping in the middle of a questionnaire and deciding not to return it. It can also lead to carelessness in responding, as represented by the attitude: “If the questions don’t quite fit me, then maybe I don’t need to be all that careful with my answers.” Consequently, the design of good skip directions is considered an essential feature of good questionnaires.

Principle 3.28: Words and phrases that introduce important, but easy to miss, changes in respondent expectations should be visually emphasized consistently, but sparingly.

Emphasizing certain words in questions is frequently essential; for example, when a question introduces a new and particularly important qualification such as “the last three weeks” (see Figure 3.22). Words should be emphasized in the same manner rather than using different techniques in the same sentence, as shown in the first example of Figure 3.22. However, too much emphasizing of words may encourage people to focus only on the words that are emphasized and contribute to the frequently observed tendency to read prose unevenly. In the second example, the entire meaning of the question could be changed if someone were to read only the underscored words. The complete meaning of the question would be changed from asking whether respondents had not called a doctor to asking whether they had. The third example is preferred because it brings attention to the time referent, which was deemed the most critical part of this question for the respondent to comprehend.

I have used underlining here to emphasize these few words, assuming that bold printing of questions was already a convention for stating questions. Italics could be used, unless perhaps their use had been accepted as a convention for additional instructions. Reverse print would not be acceptable because of its strong figure/ground contrast, which might result in the emphasized

Figure 3.22 Emphasize words and phrases in questions consistently, but sparingly.

Problem:

8. During the **LAST THREE WEEKS**, did you ever experience pain but decide it wasn't bad enough to *call a doctor*?

Still a problem:

8. During the last three weeks, did you ever experience pain but decide it wasn't bad enough to call a doctor?

Another revision:

8. During the last three weeks, did you ever experience pain but decide it wasn't bad enough to call a doctor?
-

words not getting read. Similarly, I would avoid other figure/ground changes such as printing the words in a new color. I also avoid using different kinds of emphasis in one sentence (as shown in the first example) since that compounds the difficulty of reading. Underlining maintains the same figure/ground format as in the rest of the sentence, but simply adds a new element to it.

Placing emphases on words seems to be a phenomenon that grows as multiple people with different interests become involved in determining the wording of questions. I prefer to use it when an idea changes a pattern from previous questions. In this case, changing the recall period to three weeks from the time period used in previous questions seemed most likely to get missed, and these were the words chosen for emphasis.

FINAL ADVICE ON GETTING FROM PAGE CONSTRUCTION PRINCIPLES TO FINISHED PAGES

Adhering to the 28 construction principles provided here may seem a daunting challenge, if not a confusing one. They tie together a clear set of objectives about what the written page should look like, instructions about how respondents should process the information on those pages, and implicit propositions about what visual elements will best help achieve those objectives. Efforts to apply these principles to dozens of questionnaires suggest that they leave much room for judgment and alternative means of application. The result is that good questionnaires can be constructed in ways that are visually quite different. Some designers never use reverse printing; others use it as the

cornerstone of their navigational system. Some designers use one-column questionnaire designs and others use two-column designs. Some sponsors prefer to use no color and adhere to a fundamental black on white figure/ground design, whereas others have the capability and desire to make extensive use of color. Suffice it to say that there are many different ways that competently constructed questionnaires can be produced while adhering to the objectives and procedures outlined in this book.

An approach I have found quite workable with survey sponsors and questionnaire designers is to create a style sheet. It outlines visual choices that should be followed, ranging from font size variations to column widths. Draft after draft after draft is then constructed until a satisfactory final product is provided, which is ready for printing between two covers, the topic to which we now turn.

DESIGNING THE QUESTIONNAIRE COVER PAGES

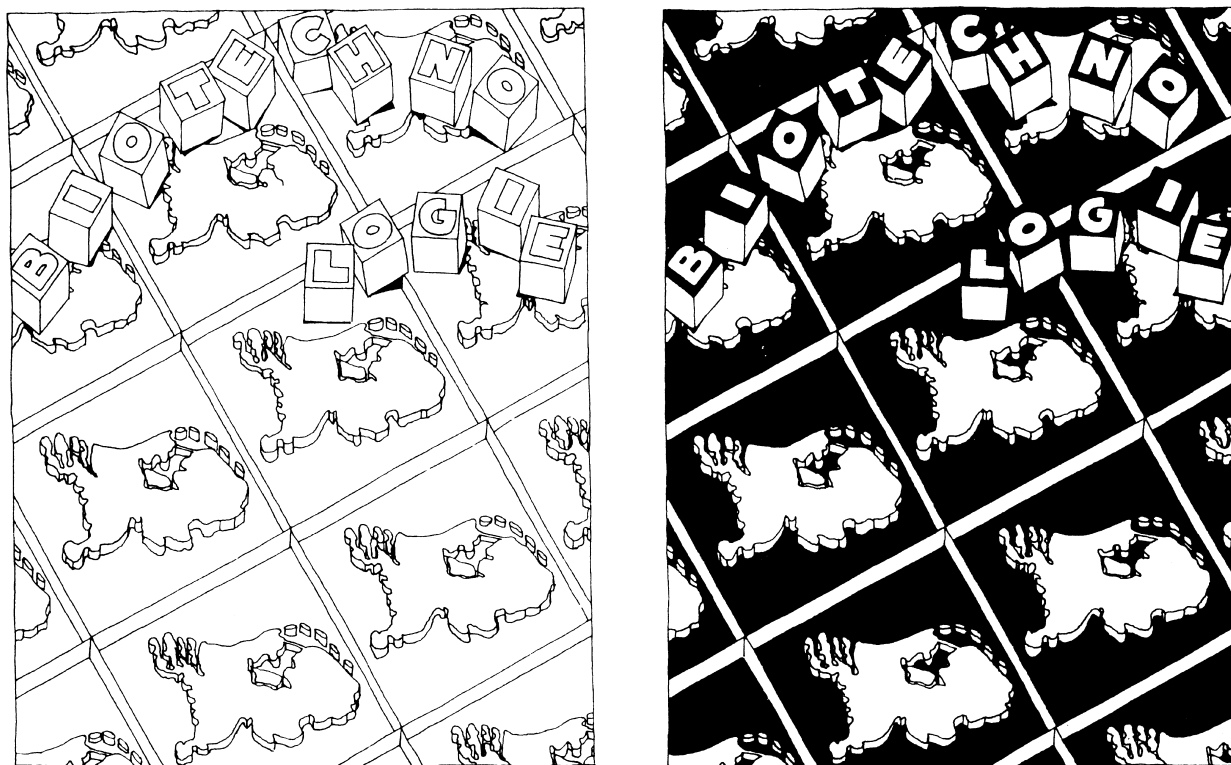
THE FRONT COVER

Questionnaire covers, especially the front cover, are likely to be examined before any other part of the questionnaire. This space should be viewed as an opportunity to motivate respondents rather than a place to put information that doesn't seem to fit anywhere else, such as detailed directions or background on why the survey is being done.

Evidence exists that questionnaire cover designs can improve response rates. Anton Nederhof, a Dutch psychologist, experimented with the two covers shown in Figure 3.23. The first cover is white with little contrast and consists of a repeated map of the Netherlands and a title. The other is the same except for a black emphasis with higher contrast (Nederhof, 1988). The questionnaire with the dominantly black cover achieved a significantly higher response rate than the low contrast white cover (86% compared to 76%) from random samples of Dutch biotechnologists. Of equal importance was the finding of when the improvement in response rate occurred. After an initial mailing, response rates were virtually the same. However, when a postcard reminder was sent, the response rate for the dominantly black questionnaire cover achieved a 10% lead that was maintained through two additional follow-ups. Nederhof reasoned that the black cover was more memorable and when the reminder postcard arrived it was easier to retrieve from previous mail.

In another experiment, Grembowski (1985) found that a graphic design of a young girl sitting in a dental chair with the title of "At What Cost, Dental Care" achieved a significantly lower response rate than a comparable questionnaire with a child by a fountain of water and the caption, "The Effects of Fluoridation on Children's Health." It was unclear why the differences in this response rate occurred, although all of the recipients had state-provided den-

Figure 3.23 High and low contrast covers used by Nederhof (1988) to improve retrievability.



Reprinted by permission of Academic Press.

tal insurance so cost may have not been as salient a concern as the effects of fluoridation on their children.

An experiment on the effects of colored questionnaires versus black and white, which also incorporated variations in word versus graphic designs, and an attempt to replicate the Nederhof experiment failed to produce any differences in response rates (Dillman & Dillman, 1995). However, results indicated a slight favoring of the colored covers (two to four percent) over the others. Although differences were small, this design included tests of four different covers, including a text-only version, all of which were deemed well designed.

An argument can also be made for not having a separate front cover page and starting immediately on page one with the survey questions. If a questionnaire is short and the reasons for completing the questionnaire already known by most respondents, a questionnaire cover may only make the questionnaire look lengthier or make it look like a brochure that does not need to be filled out and returned. In such situations one may consider limiting the cover design to a masthead containing the bare identification essentials and starting the questions immediately below. The concept of starting questions on the first page is supported by extensive cognitive research about people's expectations for the questionnaire to be used in the 2000 Decennial Census (Dillman, Jenkins, Martin, and DeMaio, 1996).

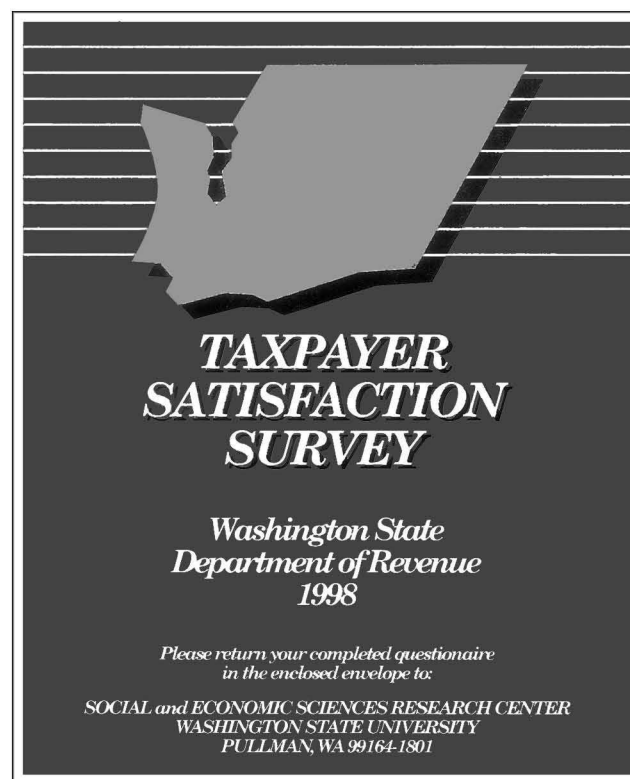
The final word has not yet been written on the effects of questionnaire covers on response rates. However, for reasons of practicality and the need for a place to convey critical information, I follow certain principles for designing separate cover pages for most questionnaires.

First, the questionnaire needs to be immediately distinguishable from all other questionnaires that a respondent might receive, while creating a positive first impression. For this reason a simple and neutral graphical design is often used, complete with a title, as shown in Figure 3.24. Following Nederhof, my goal is to make the questionnaire memorable and therefore retrievable at the time of follow-up. Detailed pictures, especially those selected quickly from clip-art files, should be avoided. I once rushed to find a graphic design for surveying commercial apple growers in Washington state and chose a picture of someone on a ladder picking apples. The first return was from one rancher who circled the cover of his blank questionnaire and simply noted, "We haven't picked apples in this way for years!"

Similarly, a draft cover for a statewide survey on seat belt use was rejected because the graphic designer used an outline of a family in a car, none of whom appeared to be wearing seat belts. In addition, it was a family of four with the father driving, the mother sitting in a seemingly deferential position facing him, and a boy and girl in the back seat of a station wagon that was at least a decade old. A redraft of the cover produced a new car, a unisex driver, one person wearing a seat belt, and the others seated in the back seat where

Figure 3.24 Examples of front and back questionnaire covers used for a state government-sponsored survey.

Front Cover



Back Cover

Thank you for taking the time to complete this questionnaire. Your assistance in providing this information is very much appreciated. If there is anything else you would like to tell us about this survey, or the services and products provided by the Washington State Department of Revenue, please do so in the space provided below.

To request this questionnaire in an alternate format for the visually impaired, or in a language other than English, please call 206-753-3217. For assistance for the hearing impaired, please call (TTY) 1-800-451-7985.

Please return your completed questionnaire in the envelope provided to:
Social & Economic Sciences Research Center
PO Box 641801
Pullman, WA 99164-1801

that determination could not be made. The more detail one attempts to build into covers, the greater the likelihood of presenting something that some respondents will find unacceptable. For this reason, simple yet distinctive graphics aimed at making the questionnaire more retrievable are chosen.

Second, a title is included so that in any conversation with the respondent it is easy to identify which questionnaire the respondent is asking about. The title, like that on a book, should be short and simple and written to the respondent, not to the organization or agency sponsoring the survey. In the mid-1990s the U.S. Bureau of the Census began experimenting with a survey that might be used each year to obtain information usually obtained every ten years (the Decennial Census). The working title for the project was the "Continuous Measurement Survey," which was prominently displayed on page one as the title of the survey. Further consideration resulted in the title being changed to the "American Community Survey," a title which undoubtedly communicates better with the members of the general public who received it in the mail.

Third, the name and address of the study sponsor are included so the respondent will know where to send the questionnaire in the event the return envelope becomes separated or lost. As a visible and frequent conductor of mail surveys in Washington state, the Social & Economic Sciences Research Center has sometimes found that questionnaires it has *not* sent to people were returned to it in error because no name or address was placed anywhere on some other organization's questionnaire. Including the name and address is also important from the standpoint of appropriate disclosure of sponsorship, recognizing that cover letters that explain this information may also get lost.

Finally, the questionnaire cover is viewed as an extension of the cover letter. Identifying a questionnaire clearly as being sent from a well-known and legitimate source is desirable for fostering trust that the survey is legitimate and useful. Rather than identifying the questionnaire as being from an individual person or a relatively unknown entity within an organization, it needs to be recognized as coming from a legitimate and respected sponsor, whenever possible.

DESIGNING THE BACK COVER

The back cover should be kept simple. Normally it consists of an invitation to make additional comments, a thank you, and plenty of white space, as shown in Figure 3.24. This is not the location for adding a few remaining questions or providing detailed instructions.

Our first concern is that this page should not compete for attention with the front cover or detract from it in any way. It is preferred that a respondent first see the front cover and then immediately open the questionnaire. A design that extends from the front to the back cover increases the likelihood that the respondent will examine both covers before starting the questionnaire, and as

a result may start flipping through pages of the questionnaire. If the questionnaire is opened from the back, the first questions to be seen will be the most cost-inducing ones, such as income.

Questions are never included on the back page. If our ordering principles are strictly adhered to, the questions that would appear there, such as income, religion, or politics, are those that respondents are most likely to find objectionable. Not only would their placement on the back cover increase the chance of nonresponse, but experience has shown that the item nonresponse for those who do respond increases as well (Dillman, 1978).

Of course, one solution to the potential problem of drawing attention to the back cover is to leave it blank. However, this page can be used in more positive ways. A request for any additional comments on the topic of the study seeks to overcome one of respondents' most frequent objections to questionnaires—questions written in ways that do not allow a complete answer. The wording of the typical request, "Is there anything else you would like to tell us about . . . (the topic of the survey) to help us understand . . .," clearly implies that this question should be completed last, and thereby refers people to the front of the questionnaire. The solicitation of comments might help in future efforts based on an exchange principle; that is, many people are rewarded by being asked for their advice in a consulting manner. The expression of appreciation is also a small but important attempt to reward respondents for completing the questionnaire.

PRETESTING

Pretesting has always been a highly touted part of questionnaire design. However, in practice it is often done haphazardly, if at all. The term also means different things to different people. Some think of a pretest as an evaluation of procedures which should be done by sending some questionnaires to a small sample of the respondent population and seeing whether any problems come up. Others think of pretesting as finding out if any production mistakes were made in printing the questionnaire by having a few people fill it out. Still others think of pretesting as learning whether people understand the questions. Pretesting consists of all of these things and more, with each actively providing feedback that is not likely to come from other methods in a timely way. We divide this process into four sequential stages.

STAGE 1: REVIEW BY KNOWLEDGEABLE COLLEAGUES AND ANALYSTS

- Have I included all of the necessary questions?
- Can I eliminate some of the questions?
- Did I use categories that will allow me to compare responses to census data or results of other surveys?

- What are the merits of modernizing categories versus keeping categories as they have been used for past studies?

These are the kinds of questions that only knowledgeable people can answer. Few sets of data ever are analyzed without the analyst saying about one or more questions, "I wish we would have asked that question differently." This stage of pretesting, which takes place after all of the questions have been written and ordered, is designed to elicit suggestions based on experience with previous surveys and knowledge of study objectives.

Several types of experience are relevant. Some of it rests with people who have analyzed data and know, for example, that they could not use responses to a particular question because there is virtually no variation in the use of categories. Alternatively, the reviewer might recall that for some reason that question had a particularly high item nonresponse rate in a previous survey. Sometimes the knowledge of policy makers is relevant when they say things such as, "It may be an interesting question, but I don't know what consequence it would have for making a company decision, regardless of which way it gets answered." Another needed type of response is from individuals with survey experience who may notice that "Don't Know" categories are missing, or that a scale is unbalanced and likely to produce an answer with a positive bias. At this stage of pretesting, people need to be consulted who can identify with respondents and determine how likely it is that each of the questions can or will be answered.

It is particularly important to get feedback from people with diverse expertise, which then needs to be reconciled for producing another draft of the questionnaire. The number and types of people vary by study. In some studies feedback is solicited from dozens of individuals and divisions of an organization, representing areas such as marketing, data processing, upper management, and statistical analysis. In other cases, one or two people have been able to provide all of the help that seemed necessary.

My goal with this stage of pretesting is to finalize the substantive content of the questionnaire so the construction process can be undertaken. Once finished, a questionnaire is ready for the second stage of testing.

STAGE 2: INTERVIEWS TO EVALUATE COGNITIVE AND MOTIVATIONAL QUALITIES

- Are all of the words understood?
- Are all of the questions interpreted similarly by all respondents?
- Do all of the questions have an answer that can be marked by every respondent?
- Is each respondent likely to read and answer each question?
- Does the mailing package (envelope, cover letter, and questionnaire) create a positive impression?

In recent years a technique known as cognitive interviewing has been developed for determining whether respondents comprehend questions as intended by the survey sponsor, and whether questions can be answered accurately (Forsyth and Lessler, 1991). Potential survey respondents are asked, individually, to respond to a questionnaire in the presence of an interviewer who asks them to think out loud as they go through the draft questionnaire and tell the interviewer everything they are thinking. The interviewer probes the respondents in order to get an understanding of how each question is being interpreted and whether the intent of each question is being realized. The technique has mostly been applied to interview questionnaires. However, it and a companion method—the retrospective interview—are quite effective in identifying problems with self-administered questionnaires.

In the think-aloud interview, an interviewer explains to respondents that they will be asked to complete a questionnaire in a special way, which is outlined in Figure 3.25. This includes telling the interviewer everything they are thinking as they complete the questionnaire. Respondents are then asked to complete a practice question in order to learn the technique; for example, “How many windows are in the home where you live?” If a person is silent while appearing to be thinking of an answer, the interviewer gently probes, “Could you tell me what you are thinking now?” Typically respondents will begin to describe a counting process that might include: “Well, let’s see, I’ll start with the kitchen. It has one window, the living room has four, but if you count individual window panes, which I guess I will, there are eight,” etc. One question of this nature is usually sufficient to get across the idea of what is to be done. The interviewer then listens and probes gently whenever the respondent falls silent.

This type of interview is designed to produce information when the respondent is confused or cannot answer a question. Typical responses might include, “Let’s see, where do I go now?” or “I don’t know when this house was built, so I guess I’ll leave it blank.”

A potential shortcoming of this type of interview is that respondents are dividing their attention between the interviewer and the questions, rather than focusing entirely on the questionnaire. It is also possible that respondents will read more of each question and do so more slowly than they would if they were alone at home. As a result, the skipping of critical words that leads to wrong answers may go undetected. For these reasons, the retrospective interviewing technique is also used.

Under the retrospective technique, also shown in Figure 3.25, respondents are asked to complete a questionnaire as if they received it at home, and to complete it in whatever way they would if the interviewer were not there. The interviewer watches while respondents fill out the questionnaire, noting any wrong answers, skipped questions, hesitations, confused expressions, erasures, or other behavior that would seem to indicate a problem with under-

Figure 3.25 Examples of protocols used for testing U.S. Census questionnaires developed by Cleo Redline and Don A. Dillman (reported in Dillman, Jackson, Pavlov, Schaefer, 1998).

Concurrent think-aloud protocol:

I'm going to hand you the first Census questionnaire in an envelope, and I'd like you to fill it out the same way you would if it came to you at home, except I'd like for you to read aloud anything you would normally read to yourself. And I would like you to tell me everything you are thinking and feeling, from the moment you first see the envelope until you finish filling out the questionnaire and put it into the return envelope to send back. That includes your telling me anything you like or don't like about the envelope, the letter, and the questionnaire.

Remember to tell me what you find enjoyable and what you find frustrating from the moment I hand you the envelope until you are finished. It is important that you tell me what you are thinking and feeling while you are actually in the process of coming up with your answers. The reason we are asking you to read and think aloud is to discover how well people work with the Census mailing packages. Having people read and think aloud is a common technique in testing these packages to assess how well they work, and to discover ways to make them better.

Now because some people aren't used to reading, thinking, and expressing their feelings aloud, I'd like to begin today with a very short practice question. Remember to read aloud whatever it is you would normally read to yourself and to express your thoughts and feelings from the moment I hand you the envelope until you are finished filling out the questionnaire.

Question: How many windows are there in your home?

Examples of general probes:

- What are you thinking?
- Remember to read aloud for me.
- Can you tell me more about that?
- What do you mean by that?
- Could you describe that for me?
- Remember to tell me what you are doing.

(*Reinforcement, when done correctly.*) Ok, good. You've told me what you were thinking as you answered the question. That's what will help us out. For the census package I'm about to give you, this is exactly what we would like for you to do. Remember, I want you to read aloud and tell me everything you are reading and thinking from the time I hand you the envelope until you have put the completed questionnaire back into the return envelope for mailing.

(Hand envelope to respondent, address side up.)

(Continued)

standing. Then, after the questionnaire is completed, the interviewer asks questions about each of these potential problems. The retrospective interview may be especially useful in revealing navigational difficulties that stem from the graphical layout or the nonverbal language used in questionnaire construction. A potential shortcoming of this method is that respondents may show no outward evidence of being confused at critical points in the questionnaire. In addition, by the time the interviewer asks questions respondents

Figure 3.25 (Continued)

Retrospective interviews:

In a minute I'm going to hand you a census questionnaire in an envelope and I'd like you to fill it out the same way you would if it came to you at home. I'll stay here in the room while you fill it out, but please don't ask me any questions; just do it like you were sitting at home and I wasn't there. I will be taking some notes while you fill out the form. Please don't let this distract you. When you have finished, put it into the envelope for mailing back to the Census Bureau, let me know and then I would like to ask some questions. Okay?

(Hand envelope to the respondent, address side up.)

All right, here is the first mailing package which we would like you to think of as having just received at the address where you live.

Questions asked after each questionnaire is completed by respondent.

1. Do you have any reactions to this envelope, either positive or negative? What are those reactions?
2. *(Turn envelope over.)* Do you have any reactions to the back side of the envelope, either positive or negative? What are those reactions?
3. *(Turn back to front side.)* Does this envelope look like something that is coming from the federal government, or does it look like it's from somewhere else?
 - A. *(If federal government)* What about the envelope makes it look like it is coming from the federal government?
 - B. *(If somewhere else)* Where does it look like it is coming from and why?
4. If you received this envelope in the mail would you open it? Why or why not? Does it look official? Does it look like junk mail?
5. Did you notice a letter inside the packet?
6. Did you read the letter? Was there anything in the letter that you liked or didn't like?
7. On a scale of 1 to 5, where 1 means very easy and 5 means very difficult, how easy or difficult was it for you to figure out where to begin on the form?

Examples of post-interview probes to fulfill specific objectives of both type of interviews:

- A. Specific questions about the forms.
 1. In comparing the first form you worked with against the second one, which form did you like more? Can you provide some reasons for liking this form better? Can you provide reasons for disliking this form more?
 2. Overall, which form was easier to complete and which form was more difficult? How was this form easier and the other form more difficult?
 3. Next, I'd like to ask about how the forms were folded. Was there anything confusing or difficult about how to unfold or refold the first form you worked with? Was there anything confusing or difficult about how to unfold or refold the second form you worked with? Which form was easier to work with in terms of unfolding and refolding?

Figure 3.25 (Continued)

-
4. Suppose in the next census we can only use one of these forms. Which one of these forms do you think we should use?

B. Questions based on observation of potential problems.

Now I'd like to ask a few other questions. I noticed when you were filling out the form that ... (*What goes here depends upon what happened in the interview.*)

1. Suppose they skipped an item.... "I'd like to ask about this item (*point to it*). I see you left it blank. Was there a particular reason for that?"
2. Suppose they frowned.... "I noticed here that you seemed to be thinking really hard, or was there something about this question you were trying to figure out?"
3. Suppose they scanned ahead.... "I noticed that when you got here you stopped for a minute and looked ahead and turned over the form. Could you tell me what you might have been thinking about here?"

C. Wrap-up question.

1. Do you have anything else you would like to tell us that you haven't had a chance to mention?
-

may have forgotten something they would have articulated using the think-aloud method.

Both of these techniques can be supplemented with previously formed questions at the end of the interview to learn about motivational features of the questionnaires. Such questions may include, "Was it interesting? Was there any time that you wanted to stop answering? Did any of the questions offend you? Would you have filled out this questionnaire if it had come to you at home?"

The test questionnaire can be presented to the respondent in the mailout envelope with a cover letter. Questions are then asked to gain insight into how the person views the entire mailing package. This procedure allows insights to be gained into how well the mailing package and cover letter connect to the questionnaire.

Both the concurrent and retrospective techniques are useful and both are preferred for evaluating questionnaires and making revision decisions. I have also found them useful in comparing alternative mailing packages. In one test conducted at the U.S. Bureau of the Census on alternative mailing packages, 54 interviews were conducted in which people responded to all three of the mailing packages, using both think-aloud and retrospective interviews (Dillman, Jenkins, Martin, and DeMaio, 1996). Results of these interviews were quite consistent with response differences observed in a simultaneous national field test of these forms (Leslie, 1996). Respondents to cognitive interviews said they were more likely to open and respond to a questionnaire that emphasized an official government approach than they were to respond to others that emulated mass marketing techniques with extensive use of bright colors. The latter forms obtained response rates five to ten percentage

points less than the government-oriented forms, thus confirming the identification of problems with the marketing appeal materials identified in the cognitive/motivational interviews (Dillman, Jenkins, Martin, and DeMaio, 1996; Leslie, 1997).

There can be little doubt that greater use of cognitive/motivational interviews would improve most questionnaires. However, such interviews are labor intensive and are usually done in small numbers of a dozen or so, unless a large-scale survey of much importance is being undertaken. Doing the interviews in such small numbers means one should be cautious about thinking that all problems can be caught by such interviews. Problems that exist for only five to 10 percent of the survey population are likely to be missed entirely as a result. Nonetheless, the use of cognitive interviews is indispensable and a natural lead-in to a third type of pretesting.

STAGE 3: A SMALL PILOT STUDY

Some questions cannot be answered by cognitive interviews or the other pretest activities I have discussed. Presumably, the knowledgeable person review and the cognitive/motivational interviews have revealed ways of improving the questionnaire. The next pretest step is to do a pilot study that emulates procedures proposed for the main study. Generally, this means that a considerable amount of resources are being invested in the survey and there is a need to see if all of the parts work.

- Have I constructed the response categories for scalar questions so people distribute themselves across categories rather than being concentrated in only one or two of them?
- Do any items from which I hope to build a scale correlate in a way that will allow me to build scales? What kind of response rate is the survey likely to obtain?
- Are some questions generating a high item nonresponse rate?
- Do some variables correlate so highly that for all practical purposes I can eliminate one or more of them?
- Is useful information being obtained from open-ended questions?
- Are entire pages or sections of the questionnaire being skipped?
- What response rate can I expect?

For a pilot study, a sample of 100 to 200 respondents is generally drawn, but it may be larger if resources allow. The respondents receive each of the mailings or other contacts just as they would with the main survey. Such a pilot allows quantitative estimates to be made for response rates, item nonresponse, and variable distributions.

Pilot studies frequently result in substantial revisions being made in the

survey design, from adding additional contacts or an incentive to improve response rates, to eliminating or adding survey questions. Entering data from 100–150 responses allows one to make reasonably precise estimates as to whether respondents are clustering into certain categories of questions. One pretest I observed resulted in changing a number of community satisfaction questions from a four-point (excellent, good, fair, or poor) to a six-point (outstanding, very good, good, fair, poor, or very poor) scale because nearly two-thirds of the respondents were grouping into the “excellent” categories, thus limiting the variation analysts had hoped to have for examining relationships among variables. I have also seen skip instructions made more prominent by means of the visual techniques described in this chapter, because some were being missed in the pilot study.

STAGE 4: A FINAL CHECK. DID WE DO SOMETHING SILLY?

The final step of pretesting is to ask a few people who have had nothing to do with the development or revision of the questionnaire and related materials to sit down by themselves and answer it completely. People who have worked on one revision after another soon lose their ability to detect obvious problems. Asking employees to read a questionnaire again for problems does not always result in catching them.

I once was explaining this last step of pretesting to some students a few minutes after handing a questionnaire to a likely respondent for this final test. Before finishing my explanation the respondent appeared in the doorway with a frown and said, “Did you really mean to do this?” Her finger was pointing to the scale used for six opinion items on the first page which read: “strongly oppose, somewhat oppose, somewhat favor, strongly oppose.” Somehow this error had slipped through several final readers who had read the questionnaire more as proofreaders than as respondents.

CONCLUSION

Too often, constructing the questionnaire is viewed by survey sponsors as an afterthought—the task that someone else does after they have approved the list of questions. Sometimes the task is delegated to an inexperienced employee with the instruction to make it look as short as possible. A more appropriate perspective is that a list of questions is only the starting point, and there is much left to be done that will significantly influence nonresponse, item omissions, measurement error, and overall survey success.

Constructing a self-administered questionnaire is difficult at best. A good questionnaire is almost never drafted in one sitting, or even several. Experienced questionnaire designers often need eight or even ten revisions to get close to a version that meets all of the criteria discussed in this chapter, in-

cluding selecting a format, ordering the questions, and organizing the two information languages—words and graphics—that comprise every page.

Some of the concepts discussed in this chapter are unfamiliar to most well-trained survey methodologists. These may include the simultaneous manipulation of size, brightness, figure/ground, simplicity, location, and regularity to achieve appropriate groupings and subgroupings of the information that comprises each questionnaire page. Yet, application of some of these concepts is not something completely new. It has been done implicitly by many surveyors since self-administered questionnaires were first used. What is new is that modern word processors and copying capabilities have given every questionnaire designer far greater capabilities for this manipulation than were available prior to the 1980s. The options now available require us to achieve a more formal understanding of how people see and process information. We expect these capabilities for manipulating words, shapes, and symbols on questionnaire pages to continue to expand, especially for the development of web questionnaires, as discussed in Chapter 11. Thus, it is important that understanding and working with these concepts be added to the professional skills of questionnaire designers.

The reactions I have received to attempts to teach the integration of visual and word aspects of questionnaire construction to others has been instructive. For some people it seemed only natural to visually shape a questionnaire page, even to the point of their thinking that words are mostly irrelevant in determining how they move from one part of a questionnaire page to another. Others have responded with folded arms and the declaration that they are “readers” and would do whatever the words suggest, regardless of graphical layout. This perspective sometimes leads to the conclusion that for them, visual layout is immaterial. Cognitive interviews done on dozens of questionnaires suggest to me that neither reality is as pure as it is sometimes expressed—self-professed readers often make navigational mistakes and graphically-oriented individuals often become engrossed in the visual flow of questionnaires to the point of missing the intent of a question or instruction. These experiences suggest to me that the science of applying these concepts is still being developed. In my view, these concepts represent exceedingly important research priorities for the future.

The desired outcome of the questionnaire construction process is a document that has been tested and retested, and is ready to be turned over to the respondent. The procedures for doing that in pursuit of a response from most or all members of the survey population is the topic of the next chapter.