LEARNING OBJECTIVES

To be able to

- Explain each of the 15 principles of questionnaire construction.
- Know when open-ended questions and closed-ended questions should be used.
- Give multiple examples of response categories used for completely anchored rating scales.
- Explain how the different parts of a questionnaire are organized into a smoothly operating whole.
- List and explain the five major steps in questionnaire construction.
- Summarize and explain the content in the checklist for questionnaire development.
Rachel, a second-grade teacher, was excited by the prospect of offering her opinion by filling out a survey instrument (i.e., a questionnaire). The questionnaire was designed by a team of researchers investigating what needed to be done to help improve student learning outcomes at her school. Rachel was happy to be asked to fill out the questionnaire. She knew that her opinions were valuable and could hardly wait to write them down. She had already discussed her thoughts with other teachers at her school about some of the topics that the questionnaire would probably address.

But as Rachel sat trying to fill out the questionnaire, she was dismayed. To begin with, the important questions weren’t even asked. Nowhere on the questionnaire was a question about funding or after-school programs. Even worse, there was nowhere to talk about any issues not specifically asked. As she worked through the questionnaire, she began to get more and more worried. The ordering of the questions was confusing to the extent that, at times, she wasn’t sure if she was putting her answers in the proper place. Several questions asked about several issues all in one question, and she had different feelings about the issues. Some questions were so confusing that she didn’t know whether she should say that she agreed or disagreed because she simply didn’t know what the question was asking. She sat for ten minutes wondering how to answer this question: “Do you not agree with letting students not do their homework on a daily basis?” She knew that homework was important, but does that mean she agrees or not? Even worse, what if she agreed with homework but not on a daily basis?

Further on in the questionnaire she couldn’t help but feel that the researchers had an agenda. She could tell that they believed in block scheduling of courses just by how the questions were phrased. Rachel worried about how she would look if she disagreed, even though she never liked the block scheduling idea. Finally, at the end of the questionnaire Rachel decided to start randomly marking answers because the jargon used was so hard to understand. As Rachel turned in the questionnaire, she was saddened by the fact that the researchers would never know any of her well-thought-out opinions, and she decided never to waste her time and volunteer for a research project again.

The purpose of this chapter is to help you understand how to construct a questionnaire when you need this type of data-collection instrument for your research study and one is not already available from past research. The questionnaire might be the only data-collection instrument, or it might be used along with other data-collection methods in a research study. You will learn that if you follow some of the simple principles addressed in this chapter your research participants will be less likely to face situations like the one faced by Rachel, our second grade teacher, and your data will be more complete and useful as well.

WHAT IS A QUESTIONNAIRE?

A questionnaire is a self-report data-collection instrument that each research participant fills out as part of a research study. Researchers use questionnaires so that they can obtain information about the thoughts, feelings, attitudes, beliefs, values, perceptions, personality, and behavioral intentions of research participants. In other words, researchers attempt to measure many different kinds of characteristics using questionnaires.

We view the term questionnaire broadly, meaning that questionnaires are not restricted to a single research method. Questionnaires can be used to collect data with multiple research methods (quantitative, qualitative, mixed, etc.). Furthermore, the content and organization of a questionnaire will correspond to the researcher’s research objectives. The key point is that the questionnaire is a versatile tool available to you and other educational researchers.

Questionnaires typically include many questions and statements. For example, a researcher might ask a question about the present (Do you support the use of corporal punishment in elementary schools?), the past (Have you ever used corporal punishment with one of your students?), or the future (Do you think that you will use corporal punishment sometime in the
future?). See Table 7.1 for more examples. Questionnaires can also include statements that participants consider and respond to. For example, when filling out the Rosenberg Self-Esteem Scale in Figure 7.1, research participants must indicate their degree of agreement or disagreement with ten statements measuring their attitudes toward themselves.

### TABLE 7.1 Type of Question Matrix With Examples

<table>
<thead>
<tr>
<th>Question Focus</th>
<th>Time Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Behavior</strong></td>
<td><strong>Past (Retrospective)</strong></td>
</tr>
<tr>
<td>When you were a teenager, did you use any illicit drug?</td>
<td>Do you currently watch educational television?</td>
</tr>
<tr>
<td><strong>Experiences</strong></td>
<td><strong>Present (Current)</strong></td>
</tr>
<tr>
<td>What was it like taking a class from your favorite teacher?</td>
<td>What is it like being interviewed about your childhood?</td>
</tr>
<tr>
<td><strong>Attitudes, opinions, beliefs, and values</strong></td>
<td><strong>Future (Prospective)</strong></td>
</tr>
<tr>
<td>When you were a child, did you like school or church more?</td>
<td>Do you support school vouchers?</td>
</tr>
<tr>
<td><strong>Knowledge</strong></td>
<td></td>
</tr>
<tr>
<td>Did you know the definition of tabula rasa when you first started college?</td>
<td>What is the definition of tabula rasa?</td>
</tr>
<tr>
<td><strong>Background and demographic</strong></td>
<td></td>
</tr>
<tr>
<td>How old were you when you entered the first grade?</td>
<td>How old are you?</td>
</tr>
</tbody>
</table>

### FIGURE 7.1

The Rosenberg Self-Esteem Scale

Circle one response for each of the following ten items.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I feel that I am a person of worth, at least on an equal basis with others.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2. I feel that I have a number of good qualities.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3. All in all, I am inclined to feel that I am a failure.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4. I am able to do things as well as most other people.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5. I feel I do not have much to be proud of.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6. I take a positive attitude toward myself.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7. On the whole, I am satisfied with myself.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8. I wish I could have more respect for myself.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>9. I certainly feel useless at times.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>10. At times I think I am no good at all.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

*Items marked with an asterisk have reversed wording. The numbers on items with reversed wording should be reversed before summing the responses for the ten items. For example, on item 3, "strongly agree" becomes 4, "agree" becomes 3, "disagree" becomes 2, and "strongly disagree" becomes 1.

PRINCIPLES OF QUESTIONNAIRE CONSTRUCTION

The most important principles of questionnaire construction are shown in Table 7.2. Take a moment to examine this list of 15 principles so that you will have an overview of what is important to consider when constructing a questionnaire. We explain each of these principles in more detail. Remember that the goal of the questionnaire is to tap into and understand the opinions of your participants about variables related to your research objectives. As you construct the questionnaire, you must constantly ask yourself if your questions will provide clear data about what your participants think or feel.

### Table 7.2 Principles of Questionnaire Construction

<table>
<thead>
<tr>
<th>Principle</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Make sure the questionnaire items match your research objectives.</td>
</tr>
<tr>
<td>2</td>
<td>Understand your research participants.</td>
</tr>
<tr>
<td>3</td>
<td>Use natural and familiar language.</td>
</tr>
<tr>
<td>4</td>
<td>Write items that are clear, precise, and relatively short.</td>
</tr>
<tr>
<td>5</td>
<td>Do not use &quot;leading&quot; or &quot;loaded&quot; questions.</td>
</tr>
<tr>
<td>6</td>
<td>Avoid double-barreled questions.</td>
</tr>
<tr>
<td>7</td>
<td>Avoid double negatives.</td>
</tr>
<tr>
<td>8</td>
<td>Determine whether an open-ended or a closed-ended question is needed.</td>
</tr>
<tr>
<td>9</td>
<td>Use mutually exclusive and exhaustive response categories for closed-ended questions.</td>
</tr>
<tr>
<td>10</td>
<td>Consider the different types of response categories available for closed-ended questionnaire items.</td>
</tr>
<tr>
<td>11</td>
<td>Use multiple items to measure abstract constructs.</td>
</tr>
<tr>
<td>12</td>
<td>Consider using multiple methods when measuring abstract constructs.</td>
</tr>
<tr>
<td>13</td>
<td>Use caution if you reverse the wording in some of the items to prevent response sets in multi-item scales.</td>
</tr>
<tr>
<td>14</td>
<td>Develop a questionnaire that is easy for the participant to use.</td>
</tr>
<tr>
<td>15</td>
<td>Always pilot test your questionnaire.</td>
</tr>
</tbody>
</table>

**Principle 1. Make sure the questionnaire items match your research objectives**

This cardinal principle should be obvious. You must always determine why you intend to conduct your research study before you can write a questionnaire. If you plan on conducting an exploratory research study (i.e., you want to collect some preliminary information about a research problem), your questionnaire will usually not need to be as detailed as when you plan on conducting a confirmatory research study (i.e., when you intend on collecting data that will enable you to test research hypotheses). That is, when your primary goal is to explore the topic, you want to be broad in your questions so that you do not miss an important concept that your research participants feel is relevant. In both exploratory and confirmatory research, you should carefully review the existing research literature as well as any related instruments that have already been used for your research objectives before deciding to construct your own questionnaire. One of the worse things that can happen in questionnaire-based research is realizing that you should have asked a question or included a variable after your data have been collected.
Think back to Rachel, our second-grade teacher. She was upset that a question was not asked about after-school programs. After-school programs are clearly important, yet there was no mention of them on the questionnaire. This lack could be an indication that the designers of the questionnaire did not carefully consider the research on the topic before designing the questionnaire. As a result, a likely important variable was not measured fully, which will affect the research results as well as the researchers’ understanding of Rachel’s true opinion on the topic.

Principle 2. Understand your research participants

A key to effective questionnaire construction is understanding your research participants. Remember that it is they, not you, who will be filling out the questionnaire. A very important strategy when you write a questionnaire is to develop an empathetic understanding or an ability to “think like” your potential research participants. If the questionnaire does not “make sense” to your participants, it will not work.

Principle 3. Use natural and familiar language

You should use language that is understandable to the people who are going to fill out your questionnaire. Try to avoid the use of jargon or technical terms. This principle is related to the above principle of understanding your research participants. You must know enough about your participants to use language familiar to those filling out the questionnaire. Consider the age of your participants, their educational level, and any of their relevant cultural characteristics when deciding on the kind of language you need to use. Remember that it is very possible that not everyone uses the same everyday language as you. If you are reading this book, you are probably a college graduate and you are also working on a graduate degree. Using natural and familiar language makes it easier for participants to fill out a questionnaire and helps participants feel more relaxed and less threatened by the task of filling it out.

One key issue related to both the principles of understanding your participants and using natural and familiar language is that of reading level. It is important to use the reading level that is natural and appropriate for your research participants. Poorly constructed questionnaires are written at either too high or too low a reading level for the intended participants. If the reading level is too high for your participants, there is the risk that those filling out the questionnaire will skip questions simply because they do not understand what is asked or, in a worse case, they will “guess” an answer that might not reflect their true opinion. Almost as problematic is when the questionnaire is written significantly below the reading level of those for whom it is intended. When this occurs, participants are sometimes insulted by the low level and do not take the questionnaire seriously or refrain from participating in additional research. Further, too low of a reading level can result in a more simplistic and less rich view of the topic than would have been possible if a higher level had been used. If you can effectively consider how your research participants will interpret and react to each item on your questionnaire, then it is likely that you will be able to write items that will provide useful information.
Principle 4. Write items that are clear, precise, and relatively short

Each item on your questionnaire should be understandable to you (the researcher) and to the participants (the people filling out the questionnaire). Because each item is measuring something, it is important for it to be clear and precise. The GIGO principle is relevant here: “garbage in, garbage out.” If the participants are not clear about what is being asked of them, their responses will result in data that cannot or should not be used in a research study. Your goal is for each research participant to interpret the meaning of each item in the questionnaire in exactly the same way. You should also avoid technical terms; if you must use a technical term, remember to define it for the participants. Finally, try to keep most items relatively short because long items can be confusing and stressful for research participants.

Once again consider Rachel, our ill-fated research participant who reported being confused by the questions even though she was clear about the topics being studied. Although she could have offered valuable insights to the researchers on the topics, she got “lost” in the wording of the questions, the jargon of the language, and perhaps even the reading level. As a result, the researchers did not get a clear picture of her opinions and Rachel became frustrated. This situation would have been avoided if the researchers had taken the time to understand their research participants and write clear, precise questions.

Principle 5. Do not use “leading” or “loaded” questions

A leading or loaded question biases the response the participant gives to the question. A loaded question is one that contains emotionally charged words (words that create a positive or negative reaction). For example, the emotionally charged word liberal was often avoided by politicians during the 1980s and 1990s, even by those with more leftist political leanings, because the word created a negative reaction in some people regardless of the content of the statement. Some other examples of loaded words are politician, communist, welfare, drug czar, soccer mom, pro-life, pro-choice, gay, drug abuser, and juvenile. A leading question is one that is phrased in such a way that it suggests a certain answer. Here is an example of a leading question:

Don’t you agree that teachers should earn more money than they currently earn?

- Yes, they should earn more.
- No, they should not earn more.
- Don’t know/no opinion.

The phrase “Don’t you agree” leads the participant here. A more neutral wording of this question would be as follows:

Do you believe teacher salaries are a little lower than they should be, a little higher than they should be, or about right?

- Teacher salaries are a little lower than they should be.
- Teacher salaries are a little higher than they should be.
Teacher salaries are about right.
Don't know/no opinion.

Here is an entertaining example of a question that is leading and has loaded phrases in it (from Bonevac, 1999):

Do you believe that you should keep more of your hard-earned money or that the government should get more of your money for increasing bureaucratic government programs?

Keep more of my hard-earned money.
Give my money to increase bureaucratic government programs.
Don't know/no opinion.

Always remember that your goal is to write questionnaire items that help participants feel free to provide their natural and honest answers. You want to obtain responses that are undistorted by the particular question wording. Recall in our opening example that Rachel felt the researchers had an “agenda” and she was worried that she couldn't appropriately agree or disagree with certain questions. Have you ever felt that way when filling out a questionnaire? If so, you fell prey to leading or loaded questions.

**Principle 6. Avoid double-barreled questions**

A **double-barreled question** combines two or more issues or attitude objects in a single item. Here is an example: Do you think that teachers should have more contact with parents and school administrators? As you can see, this single item asks about two different issues. The question is really asking, Do you think that teachers should have more contact with parents? and Do you think that teachers should have more contact with school administrators? Each of these two issues may elicit a different attitude, and combining them into one question makes it unclear which attitude or opinion is being measured. It is impossible to determine which barrel of the question was answered.

Because it is impossible to know which part of the question the participant addressed or if he or she addressed the union of the two, it is a good rule to avoid double-barreled questions. As a general rule, if the word and appears in a question or statement, you should check to see whether it is double-barreled or rather if the question is just getting at a very specific situation.

**Principle 7. Avoid double negatives**

When participants are asked for their agreement with a statement, double negatives can easily occur. For example,

Do you agree or disagree with the following statement?

Teachers should not be required to supervise their students during library time.

If you disagree with the statement, you must construct a **double negative** (a sentence construction that includes two negatives). If you disagree, you are saying that you do not
think that teachers should not supervise students during library time (Converse & Presser, 1986). In other words, you probably believe that teachers should supervise students during library time.

Here is another example of a double negative:

Teachers should not be able to do the following things:

- Spank children
  - Yes
  - No
- Expel children from school
  - Yes
  - No

If you use a negative item, you should consider underlining the negative word or words to catch the participant's attention.

**Principle 8. Determine whether an open-ended or a closed-ended question is needed**

An open-ended question enables participants to respond in any way that they please. Open-ended questions take you into the natural language and worlds of your research participants, and therefore, open-ended questions provide primarily qualitative data. In contrast, a closed-ended question requires participants to choose from a limited number of responses that are predetermined by the researcher. Closed-ended questions provide primarily quantitative data. Although open-ended questions are typically analyzed qualitatively, the answers sometimes are analyzed quantitatively by counting the number of times a response was provided. Furthermore, an open-ended question can provide quantitative information as in this example: “How many times have you removed a student from your class for disciplinary reasons in the last year?”

To determine someone’s marital status, you could use the question “What is your current marital status?” and leave sufficient space for participants to write in their answers. In this case, the question would be an open-ended question because the participants would have to provide answers in their own words. On the other hand, you could use a closed-ended question to determine someone’s marital status, like this:

What is your current marital status? (Check one box.)

- Single
- Married
- Divorced
- Separated
- Widowed
In the question about marital status, notice that the item stem (the words forming the question or statement) was the same in the open-ended and the closed-ended question examples: both ask, What is your current marital status? In short, the difference between an open-ended question and a closed-ended question is just the way participants are allowed to respond. In open-ended questions, participants must come up with their own answers; in closed-ended questions, participants must select from the predetermined responses provided by the researcher.

Open-ended questions are usually used in exploratory research (i.e., when the researcher knows little about the topic), and closed-ended questions are usually used in confirmatory research (i.e., when the researcher wants to test specific hypotheses). Open-ended questions are valuable when the researcher needs to know what people are thinking and the dimensions of a variable are not well defined. Because the participants respond by writing their answers in their own words, open-ended questions can provide rich information. For example, the following open-ended question would provide some interesting information: What do you think teachers can do to keep students from using illicit drugs? It is more difficult and more time consuming, however, to analyze the data obtained from open-ended questions. Nonetheless, open-ended questions are at the heart of qualitative research where the goal is to understand participants’ inner worlds in their natural languages and categories.

A closed-ended question is appropriate when the dimensions of a variable are already known. Closed-ended questions expose all participants to the same response categories and allow standardized quantitative statistical analysis. Often, researchers will use the responses obtained from open-ended questions to design closed-ended questions for future questionnaires. For example, a researcher might group the suggestions for keeping students off drugs provided by teachers into a set of categories (e.g., education, after-school programs, discipline) and use these categories as response choices in a future closed-ended question.

Questionnaires can be classified by the type of questions that are used. Questionnaires that include mostly open-ended items are called qualitative questionnaires. These questionnaires are often used for exploratory research, such as when the researcher wants to know how participants think or feel or experience a phenomenon or when the researcher wants to know why participants believe something happens. Questionnaires that include mostly closed-ended items are called quantitative questionnaires. These questionnaires are focused on getting participant responses to standardized items for the purpose of confirmatory research in which specific variables are measured and hypotheses are being tested. The principle of standardization is very important in quantitative research; the goal is to provide a common stimulus (item stem, response categories, and any additional information) to each person in the research study (Dillman, 2007). This is done to help ensure maximum comparability of responses. In practice, most questionnaires employ a mixture of open-ended and closed-ended items; these are called mixed questionnaires. Although we have classified questionnaires into three types, note that questionnaires actually fall on a continuum with qualitative and quantitative as end points and mixed in the middle.

Consider the frustration of our teacher in the opening example. A large part of this frustration was caused by the fact that the researchers failed to address a topic that she considered important. This frustration could have been avoided if the researchers had realized and acknowledged that they might not know all the important topics that their participants wanted to discuss. One way to deal with this potential limitation is to include an open-ended question such as “What topics do you feel are important to student learning outcomes?” The use of this
open-ended question would allow participants to express their opinions more fully, especially
opinions the researcher failed to anticipate, and it would provide the researchers with valuable
information for their research studies.

Principle 9. Use mutually exclusive and exhaustive
response categories for closed-ended questions

Categories are mutually exclusive when they do not overlap. For example, the following
categories for a question about the participant’s age are not mutually exclusive:

- 10 or less
- 10 to 20
- 20 to 30
- 30 to 40
- 40 to 50
- 50 to 60
- 60 to 70
- 70 to 80
- 80 or greater

Do you see the problem with these response categories? The problem is that they overlap. For example, a person who is 20 years old could be placed into two categories. In fact, persons aged 10, 20, 30, 40, 50, 60, 70, and 80 can all be placed into more than one category. In short, the response categories are not mutually exclusive. In a moment, we will show you how to fix this problem.

A set of response categories is exhaustive when there is a category available for all legitimate responses. For example, what is the problem with the following categories from a question asking for your current age?

- 1 to 4
- 5 to 9
- 10 to 14

The problem is that these three categories are not exhaustive because there is no category available for anyone over the age of 14 or anyone younger than one year old. A set of categories is not exhaustive unless there is a category available for all potential responses.

Putting the ideas of mutually exclusive and exhaustive categories together, you can see that the following set of response categories is mutually exclusive and exhaustive:

Which of the following categories includes your current age? (Check one box.)

- Less than 18
- 18 to 29
The principle of mutually exclusive categories applies because none of the categories overlap. The principle of exhaustive categories applies because there is a category available for all possible ages. Whenever you write a standard closed-ended question (a question with an item stem and a set of predetermined response categories), remember to make sure that your response categories are mutually exclusive and exhaustive.

**Principle 10. Consider the different types of response categories available for closed-ended questionnaire items**

In this section we introduce several popular types of closed-ended response categories by explaining the ideas of rating scales, rankings, semantic differentials, and checklists.

**Rating Scales**

Researchers often obtain data from research participants by providing them with questions or statements (the item stem) and rating scales (the response choices), with instructions to make judgments about each item stem using the rating scale that is provided. A **rating scale** is a continuum of response choices that participants are told to use in indicating their responses. Rating scales produce numerical (quantitative) data, rather than qualitative data (nominal-level data). Rating scales have been used by researchers for quite a long time. In an early review of the history of rating scales, Guilford (1936) provides examples from as early as 1805 and many other examples shortly after 1900. Some important early developers of rating scales are Sir Francis Galton (1822–1911), Karl Pearson (1857–1936), and, somewhat later, Rensis Likert (1903–1981).

A **numerical rating scale** consists of a set of numbers and “anchored” endpoints. When you **anchor** a point on a rating scale, you label the point with a written descriptor. Here is an example of an item stem and a numerical rating scale with anchored endpoints:

How would you rate the overall job performance of your school principal?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Low</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Very High</td>
</tr>
</tbody>
</table>

As you can see, the first endpoint (“1”) is anchored with the words “Very Low.” The other endpoint (“7”) is anchored with the words “Very High.” This is a seven-point rating scale because there is a total of seven points on the scale. If you use a numerical rating scale that has only the endpoints anchored (as in the above example), we recommend that you use an odd number of
points, rather than an even number of points. When using an even number of points, a respondent might misinterpret one of the two centermost numbers as representing the center or neutral point (Dillman, 2007). If you choose to use an even number of points, you will need to anchor the two centermost numbers or clearly anchor the area between the two centermost numbers.

A similar type of rating scale is called a fully anchored rating scale. A fully anchored rating scale has all points anchored with descriptors. Here is an example of an item stem followed by a fully anchored rating scale:

My principal is an effective leader.

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

This scale is called a five-point rating scale because there are five points on the scale. (It should not be called a "Likert scale," which is discussed later in this chapter.) Some researchers prefer to exclude the numbers and provide just the descriptors in a fully anchored rating scale. Regardless, you should attempt to make the words or anchors used for adjacent points an equal distance apart from each other. The above scale is sometimes called an "agreement" rating scale because the participants report on their agreement or disagreement with the statement provided by the researcher in the item stem.

You must be very careful in your choice of anchors for both fully or partially anchored scales. Anchors are designed to provide reference points that participants will use to direct the expression of their opinions. If the reference points are one sided, not clear, or are not spaced at equal distances, then you will not get an accurate measure of the participants' opinions. Consider the following unbalanced five-point rating scale:

I enjoy my workplace environment.

1 2 3 4 5
Very Strongly Disagree Slightly Agree Strongly Agree Very Strongly Agree

In the above example, there are four anchor or reference points for agreement and only one for disagreement. This looks like a scale that an unethical politician might try to use because he or she wants people to agree with his or her policies. These faulty response categories would make it easy for a respondent to agree but difficult for him or her to disagree. As you construct anchors for rating scales, remember to use a good set of anchors that are balanced and have an equal distance between each pair of adjacent categories.

You might be wondering how many points a rating scale should have. Research suggests that you should use somewhere from 4 to 11 points on a rating scale (e.g., McKelvie, 1978; Nunnally, 1978). Rating scales with fewer than four points are not as reliable as rating scales with more points. On the other hand, rating scales with more than 11 points can be confusing because most participants have limited abilities to make fine discriminations when there are a great number of scale points.

When deciding how many points to include, it is good to consider how different the anchor points truly are. That is, what is the real difference between someone who indicates a 9 and an 11
on an 11-point scale? If you have more points than real differences, then you have too many points. Conversely, you must be sure to include enough points to see those real differences. Consider an extreme example of a two-point scale: agree or disagree. While this might work on a simple issue, how many more examples can you think of where there is a grey area that is neither full agreement nor full disagreement? In those cases you would need more points to get an accurate picture of the issues. One point to remember: you can always collapse categories together during data analysis if you need to, but you cannot add extra points after the fact (after you already have collected the data). As a result, some researchers err on the side of slightly more rather than less points on a scale. On the other hand, Dillman (2007) reports that he has, over the years, encouraged the use of fewer points (i.e., four or five points) because of their simplicity and their easier comprehension by respondents. We recommend starting with the commonly used (i.e., “standard”) rating scales, such as the ones provided in Exhibit 7.1, and adjusting them only if needed.

An empirical factor that can affect the number of response categories needed is how your respondents actually rate the attitudinal objects. For example, when pilot testing a questionnaire designed to measure researchers’ methodological beliefs, I (Burke Johnson) found that the traditional four-point agreement scale (strongly disagree, disagree, agree, strongly agree) was not working well. My participants sometimes complained that they didn’t fully agree; at other times they complained that they didn’t fully disagree. Therefore, I shifted to a six-point scale (strongly disagree, disagree, somewhat disagree, somewhat agree, agree, strongly agree) to provide more choices.

You might also wonder whether you should include a center or middle category in your rating scale. Research suggests that omitting the middle alternative (“neutral,” “about the same,” “average,” “no difference,” etc.) does not appreciably affect the overall pattern of results (Converse & Presser, 1986, pp. 36–37; Schuman & Presser, 1981; see Chapter 6). As a result, some researchers choose to include a middle alternative, and others choose not to include it. Both practices can be defended. You can see in Figure 7.1 that Rosenberg used four-point rating scales (i.e., he omitted the middle alternative) in his Self-Esteem Scale. Some researchers, such as Rosenberg, prefer to omit the middle alternative because it forces research participants to lean one way or the other.

Take a moment now to examine Exhibit 7.1, and you will see some rating scales that researchers and practitioners commonly use. You can use these in your questionnaires. Note that the ordering of the categories does not appear to affect response patterns (Barnette, 1999; Weng & Cheng, 2000). For example, both of these patterns are commonly used and have been shown to work: (a) strongly disagree, disagree, agree, strongly agree and (b) strongly agree, agree, disagree, strongly disagree. Also note that both four-point and five-point rating scales are commonly used by survey research experts. As you can see in Exhibit 7.1, you can construct rating scales for many dimensions, such as agreement, approval, importance, and satisfaction. When you construct your own rating scales, you will find additional dimensions that you are interested in, and you will need to construct similar (i.e., analogous) response categories for those dimensions.

Rankings

Sometimes you might want your research participants to rank order their responses. A ranking indicates the importance or priority assigned by a participant to an attitudinal object. Rankings can be used with open-ended and closed-ended questions. For example, you might first ask an open-ended question such as, In your opinion, who are the three top teachers in your school? Then you could follow up this question with a ranking item such as, Please rank...
### EXHIBIT 7.1  Examples of Commonly Used Response Categories for Rating Scales

*Note:* When you write response categories, make sure that the distance between each pair of anchors or response categories is the same. For example, the “distance” in meaning between strongly agree and agree should be the same as between strongly disagree and disagree. Some researchers prefer to reverse the order of the categories provided below.

**Agreement**
1. Strongly Disagree
2. Disagree
3. Agree
4. Strongly Agree

1. Strongly Disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly Agree

1. Strongly Agree
2. Agree
3. Disagree
4. Strongly Disagree

**Amount**
1. Too Little
2. About the Right Amount
3. Too Much

1. Not Enough
2. About the Right Amount
3. Too Many

**Approval**
1. Strongly Disapprove
2. Disapprove
3. Approve
4. Strongly Approve

1. Strongly Disapprove
2. Disapprove
3. Neutral
4. Approve
5. Strongly Approve

**Belief**
1. Definitely False
2. Probably False
3. Probably True
4. Definitely True

**Comparison**
1. Much Worse
2. Worse
3. About the Same
4. Better
5. Much Better

1. Much Less
2. Somewhat Less
3. About the Same
4. Somewhat More
5. Much More

**Effectiveness**
1. Not at All Effective
2. Not Very Effective
3. Somewhat Effective
4. Very Effective

**Performance**
1. Unsatisfactory
2. Fair
3. Good
4. Very Good

**Evaluation**
1. Excellent
2. Good
3. Fair
4. Poor

1. Very Poor
2. Poor
3. Fair
4. Good
5. Very Good

1. Very Bad
2. Somewhat Bad
3. Somewhat Good
4. Very Good

**Importance**
1. Not at All Important
2. Not Very Important
3. Fairly Important
4. Very Important
5. Extremely Important

**Knowledge**
1. Not at All Familiar
2. Not Very Familiar
3. Somewhat Familiar
4. Very Familiar

**Probability**
1. A Lot Less Likely
2. Somewhat Less Likely
3. No Difference
4. Somewhat More Likely
5. A Lot More Likely

**Satisfaction**
1. Very Dissatisfied
2. Dissatisfied
3. Neutral
4. Satisfied
5. Very Satisfied

1. Very Dissatisfied
2. Somewhat Dissatisfied
3. Somewhat Satisfied
4. Very Satisfied
order the teachers you just mentioned. Rankings can also be used with closed-ended items. For example, you might use the following closed-ended item:

Please rank the importance of the following qualities in a school principal. (Fill in your rank order in the spaces provided using the numbers 1 through 5 with 1 indicating most important and 5 indicating least important.)

- A principal who is sincere
- A principal who gets resources for the school
- A principal who is an advocate for teacher needs
- A principal who is a strong disciplinarian
- A principal who is a good motivator

As you can see, this is a closed-ended item because predetermined response categories are provided. As a general rule, you should not ask participants to rank more than three to five responses or response categories because ranking can be a difficult task for participants. Additionally, rank order items are difficult to statistically analyze and relate to other variables.

The use of a single item asking for a ranking is usually unnecessary. The recommended procedure is to have the participants rate each of the response categories using a rating scale. During data analysis, you can obtain the average rating for each of the response categories and then you can rank order those averages. This way you have data that are more easily analyzed for relationships with other variables and you can obtain a ranking of the response categories.

**Semantic Differential**

The **semantic differential** is a scaling technique that is used to measure the meaning that participants give to various attitudinal objects or concepts (Osgood, Suci, & Tannenbaum, 1957). The participants are asked to rate each object or concept provided in the item stem on a series of seven-point, bipolar rating scales. The scales are “bipolar” because they have contrasting adjectives anchoring the endpoints. The contrasting adjectives are usually antonyms. You can see an example of a semantic differential in Exhibit 7.2.

Semantic differentials are useful when you want to “profile” or describe the multiple characteristics associated with attitudinal objects. In Exhibit 7.2, you are asked to rate your school principal on 20 different bipolar rating scales. If you had all of the teachers in a school use this semantic differential, you could average the teachers’ responses and profile their principal. You might find that different groups produce different profiles. For example, male and female teachers might view the principal differently. If you need to develop a semantic differential, it is helpful to look at a book of antonyms for contrasting word pairs. You can also find some useful lists of commonly used semantic differential word pairs in Isaac and Michael (1995) and in Jenkins, Russell, and Suci (1958). Finally, there now are free antonym dictionaries available on the worldwide web.

**Checklists**

Researchers sometimes provide a list of response categories (a **checklist**) and ask the research participants to check the responses that apply to them. Multiple responses may be allowed. Here is an example of a checklist:
PART III: FOUNDATIONS OF RESEARCH

Where do you get your information about the most recent advances in teaching? (Please check all categories that apply to you.)

- Other teachers
- Professors
- Principal
- Parents
- Superintendent
- Academic journals
- Professional journals
- Magazines
- Television
- Other. Please list: _______________________________________

EXHIBIT 7.2 Example of Semantic Differential Scaling Technique

Please rate your school principal on each of the following descriptive scales. Place a checkmark on one of the blanks between each pair of words that best indicates how you feel.

**Your School Principal**

<table>
<thead>
<tr>
<th>Sociable</th>
<th>Unsociable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kind</td>
<td>Cruel</td>
</tr>
<tr>
<td>Successful</td>
<td>Unsuccessful</td>
</tr>
<tr>
<td>Wise</td>
<td>Foolish</td>
</tr>
<tr>
<td>Severe</td>
<td>Lenient</td>
</tr>
<tr>
<td>Masculine</td>
<td>Feminine</td>
</tr>
<tr>
<td>Active</td>
<td>Passive</td>
</tr>
<tr>
<td>Excitable</td>
<td>Calm</td>
</tr>
<tr>
<td>Fast</td>
<td>Slow</td>
</tr>
<tr>
<td>Predictable</td>
<td>Unpredictable</td>
</tr>
<tr>
<td>Clear</td>
<td>Confusing</td>
</tr>
<tr>
<td>Authoritarian</td>
<td>Democratic</td>
</tr>
<tr>
<td>Flexible</td>
<td>Rigid</td>
</tr>
<tr>
<td>Happy</td>
<td>Sad</td>
</tr>
<tr>
<td>Work</td>
<td>Fun</td>
</tr>
</tbody>
</table>
Checklists are occasionally useful for descriptive purposes. However, as a general rule, you should avoid multiple response items such as checklists because they are difficult to analyze and because of primacy effects (e.g., respondents are more likely to check items placed earlier in the list; Dillman, 2007). The recommended alternative is to have respondents use a response scale for each of the categories (e.g., have used/have not used, applies to me/does not apply to me).

**Principle 11. Use multiple items to measure abstract constructs**

The use of multiple items designed to measure a construct is done to increase the reliability and validity of the measure. One common procedure for the measurement of abstract constructs is called a summed rating scale. A **summed rating scale** (also called a **Likert scale**) is different from the numerical and fully anchored rating scales that we discussed earlier. Rather than being composed of a single item stem and a rating scale, a summed rating scale is composed of multiple items that are designed to measure the same idea or the same construct. Each of the items is rated by each respondent using a rating scale (e.g., a four- or five-point rating scale), and then the ratings on the multiple items are summed by the researcher for each participant, providing a single score for each person.

The Rosenberg Self-Esteem Scale shown in Figure 7.1 is a good example of a summed rating scale. It consists of ten items designed to measure self-esteem. The lowest possible total score on the full scale for a person is 10, and the highest possible total score for a person is 40. Participants will score somewhere in between these two extremes (i.e., between the minimum and the maximum scores).

The summed rating scale procedure was originally developed by the famous social psychologist Rensis Likert (pronounced LICK-ert). Likert published the results of his dissertation, which included the first known summed rating scale, in an article in 1932 (Likert, 1932). Since this time, summed rating scales have been used extensively by researchers.

The key advantages of multiple-item rating scales compared to single-item rating scales are that multiple-item scales provide more reliable (i.e., more consistent or stable) scores and they produce more variability, which helps the researcher make finer distinctions among the respondents. If you want to measure a complex construct (such as self-efficacy, locus of control, risk taking, test anxiety, dogmatism, or temperament), the use of a multiple-item scale is pretty much a necessity. When you want to measure constructs such as these, you should not, however, jump to develop your own scale. Rather, you should conduct a literature search to find already validated measures of your construct of interest. If a measure is not available, only then would you need to consider developing your own measure. The development of a good summed rating scale takes a lot of time and expertise, and it takes extensive validation before it should be used in a research study. The principles of test construction (reliability and validity) discussed in Chapter 6 must be followed when constructing a summed rating scale.

**Principle 12. Consider using multiple methods for measuring abstract constructs**

This principle follows from the long-standing maxim in social research that our measurements are partially an artifact of our method of measurement. In fact, if you use one method
of measurement for all of your variables, it is possible that your variables are correlated simply because you used the same measurement procedure (Cronbach & Meehl, 1955). The relationship between variables that you thought you were interpreting could be nothing but a measurement problem or artifact! Think about this occurring in your own life. Have you found that there is one type of measurement on which you simply do better on than others? Do you always do well on essay tests no matter the topic, but do worse on true/false tests? If you have experienced something like this you have seen why Principle 12 is so important.

To avoid this measurement-based problem, it is recommended that researchers use multiple methods of measurement. The major methods of measurement (or data collection) are questionnaires, interviews, focus groups, observations, tests, and secondary or existing data. Furthermore, within each of these types are myriad additional possibilities. For example, an interview can be done face to face, over the Internet, or on the telephone. Within each of these types, you might decide to use a summated rating scale as well as open-ended items.

The use of multiple measurement methods is so important today that more and more researchers are using “measurement models” based on two or even three measurement methods or procedures. The resulting data are often analyzed by using advanced statistical software such as LISREL or EQS. The point is that the more methods a researcher uses to measure the relevant concepts or constructs, the more confidence you can place in the researcher’s ability to tap into the essential characteristics of the concept, rather than the method.

Principle 13. Use extreme caution if you reverse the wording in some of the items to help prevent response sets in multi-item scales

When participants rate multiple items using the same or similar rating scale, a “response set” may occur. A response set is the tendency for a research participant to respond to a series of items in a specific direction, regardless of the differences in item content. One type of response set is called the acquiescence response set, which is the tendency to say yes rather than no or to agree rather than to disagree on a whole series of items. Another response set, called the social desirability response set, is the tendency to provide answers that are socially desirable. One technique used to help prevent response sets (especially the acquiescence response set) is to reverse the wording (and scoring) in some of the items. This technique is intended to encourage participants to read each item on the questionnaire more carefully. An example of reversed wording is shown in Figure 7.1. You can see that the wording for items 3, 5, 8, 9, and 10 of the Rosenberg Self-Esteem Scale is “reversed.”

Whether one should use the reverse wording technique has been debated in the questionnaire construction and test development literature. One school of thought does not recommend reversing the wording because there is some evidence that this practice can reduce the reliability and validity of multi-item scales (Barnette, 2000; Benson & Hocevar, 1985; Deemer & Minke, 1999; Weems & Onwuegbuzie, 2001; Wright & Masters, 1982). An opposing view believes that this reduction of reliability is attributed to a reduction in response sets, and it believes that the “benefit” of reducing the effects of response sets is greater than the “cost” of lower reliability. Dillman (2007) believes that reversing some items does not reduce response sets and that the reduction in reliability is due to respondents becoming confused because of the wording reversals. It is our recommendation that you use reverse-worded items only when response sets are a major concern. However, it is important for you to examine your data to try to “catch”
when a response set is occurring and eliminate those responses. Finally, you should not use a reverse-worded item if it results in a double negative.

**Principle 14. Develop a questionnaire that is properly organized and easy for the participant to use**

The checklist for questionnaire construction, shown in Table 7.3, gives an overview of what you should consider when designing the questionnaire. The ordering, or sequencing, of questionnaire items is one consideration. For example, Roberson and Sundstrom (1990) found that placing questions that respondents considered most important first and demographic questions (age, gender, etc.) last in an employee attitude survey resulted in the highest return rate. When constructing a questionnaire, you should begin the questionnaire with positive or non-threatening items because doing so helps obtain some commitment from the participants as they fill out the questionnaire. Furthermore, as writers and professionals in survey research have pointed out for many years, demographic questions should generally go last in a questionnaire, with a lead-in such as “To finish this questionnaire, we have a few questions about you.” The questionnaire should also not be overly long for the types of people in your target population. Otherwise, they might not fill out the questionnaire properly, or they might refuse to complete the entire questionnaire.

It is also a good idea to limit the number of contingency questions used in a questionnaire because participants might become confused or agitated. A **contingency question** (also called a filler question) is an item that directs participants to different follow-up questions depending on their response. It allows the researcher to “filter out” participants from questions that these participants cannot or should not attempt to answer. Here is an example of an item operating as a contingency question:

**Question 1:** What is your gender?

- **Male → (IF MALE, GO TO QUESTION 5)**
- **Female → (IF FEMALE, GO TO QUESTION 2)**

The use of contingency questions might not be a problem for **web surveys** (i.e., those in which participants go to a website to complete a questionnaire), because in web surveys, the skip patterns associated with contingency questions can be programmed to take place automatically. The use of contingency questions also is less of a problem in interview protocols because the trained interviewer does the skipping rather than the research participant.

You should include clear instructions throughout your questionnaire, and do not put too many items on a page. If a questionnaire has several different topical sections, you need to provide transitional or “lead-in” statements to orient the participants to each new topic. Other important tips are to give your questionnaire a title (e.g., “School Culture Questionnaire”), number the items consecutively from the beginning to the end, list response categories vertically rather than horizontally, provide an open-ended question at the end of your questionnaire to give the participant a place to add any comments or additional insights (e.g., “Is there anything else that you would like to add?”), provide clear instructions throughout the instrument (e.g., “please check one of the following categories”), and thank the participant for filling out your questionnaire (you can just put a “Thank You for Completing this Questionnaire” at the bottom of the last page). Finally, always try to make your questionnaire look professional because participants are more
PART III: FOUNDATIONS OF RESEARCH

likely to fill it out and they go away with a better impression of you and your organization. By using font sizes and types that are clear and readable, you enhance the clarity of your questionnaire. Additionally, you should maximize the amount of white or blank space in the questionnaire. Novice questionnaire construction is most evident when there is little or no white space. Never sacrifice white space in an effort to reduce page length. It is better to have an extra page in a readable questionnaire than a compact questionnaire that is unclear. Remember that the appearance and quality of your questionnaire also reflect on you and your organization.

**TABLE 7.3** Checklist for Questionnaire Construction

1. Follow all of the 15 principles of questionnaire construction discussed in this chapter.

2. Remember that appearance matters.
   - Make your questionnaire as professional looking as possible. The overall look of your questionnaire should be presentable, readable, and clear. Several of the points below address more specific appearance issues.

3. Use titles.
   - Always put a title on your questionnaire; it informs the participants about the topic of the questionnaire and gives the questionnaire an identity.
   - Also consider using section titles within the questionnaire, especially with longer questionnaires. These will help focus the participant on the proper topic or direction taken in the instrument.
   - Titles give a professional appearance to the overall document and show how it is organized.

4. Use short questions when possible.
   - Balance the length of the questions with that of the information to be gained. Although it is tempting to write long detailed questions, short questions work better. The longer the question the more likely the participant will misinterpret or simply not understand the item.

5. Carefully consider the placement of each question and each set of related questions.
   - Where a question appears is important. Do not put sensitive questions, such as demographic questions, at the beginning of your questionnaire. Always put sensitive questions at the end. Participants are more likely to answer questions that may make them uncomfortable if they have already invested a great deal of time in filling out the other questions first.
   - Make use of warm-up questions, especially questions that participants find interesting, at the beginning of the questionnaire. Just as you do not want to put sensitive questions at the beginning of the questionnaire, it is a good idea not to start out with your most difficult or time-consuming question. This may “scare off” participants and reduce the response rate. Instead, ask interesting, easy, short, nonthreatening (i.e., warm-up) questions first.
   - Vary question types reasonably. Break up large sections of rating scale items with an open-ended question and vice versa. Although you don’t want to jump around too much, by breaking up your types you can reduce participants’ natural inclination to fall into a response set or reduce their fatigue.

6. Number the items consecutively from the beginning to the end.

7. Use plenty of white space.
   - This produces a less crowded, more easily read questionnaire. Do not crowd a questionnaire in an effort to reduce the number of pages.

8. Use a readable font size.
   - Stick to commonly used font types like Times New Roman or Arial. Cursive or calligraphy type fonts take away from the clean look of the questionnaire. Remember your goal is a readable, professional looking questionnaire, not a pretty one.
   - If your questionnaire is web-based, be sure to use a true type font that appears on the Internet properly.
   - When considering the size of your font, be sure to consider your participants; however, it is a good rule of thumb to stick to fonts no smaller than 12 pt. If you cannot read the questionnaire from a distance of five feet, your font is too small.

9. Use different font styles, and remember that “less is more.”
   - Use different styles, such as underlining or bolding, to emphasize different sections and to aid in the flow of the questionnaire. Additionally, different styles can be used to emphasize specific words such as not and always.
   - Remember that “less is more.” Too many style types can have an opposite effect and hinder the readability of the questionnaire. If everything is underlined or placed in a bold font, then the emphasis is lost.
Principle 15. Always pilot test your questionnaire

It is a cardinal rule in research that you must “try out,” or pilot test, your questionnaire to find out whether it operates properly before using it in a research study. You should conduct your pilot test with a minimum of five to ten people. You may want to start with colleagues or friends, asking them to fill out the questionnaire and note any points of confusion. Then you will need to pilot test the questionnaire with several individuals similar to those who will be in your research study.

One useful technique to use during your pilot test is called the think-aloud technique, which requires participants to verbalize their thoughts and perceptions while they engage in an activity. When this technique is used as part of a pilot test, you ask the participants to verbalize their thoughts and perceptions about the questionnaire, including why they chose a particular response choice, while they are filling it out. You must record or carefully write down exactly what they say. It is also helpful to make audiotape or videotape recordings of the pilot test of your questionnaire

Think-aloud technique
Has participants verbalize their thoughts and perceptions while engaged in an activity
test sessions for later review. The think-aloud technique is especially helpful for determining whether participants are interpreting the items the way you intended.

You will want to use the think-aloud technique with some of the participants in your pilot test, but you should have some others in the pilot test fill out the questionnaire under circumstances that are as similar as possible to those of the actual research study. When you conduct a pilot test, you need to think about several issues. For example, be sure to check how long it takes participants to complete the questionnaire under circumstances similar to those of the actual research study. This will help you know whether the questionnaire is too long. You always can think of some additional items that you would like to add, but you must avoid writing overly long questionnaires. Other things being equal, the response rate and quality of responses are better for shorter and medium-length questionnaires than for long questionnaires.

Using the think-aloud technique, you can listen to how the participants think about the instructions and the items in your questionnaire. Try to determine whether any of the questionnaire items are confusing or threatening. Ask your participants to tell you when they reach an item that is difficult to understand, and then ask them to paraphrase what they believe the problem item is intended to measure. Determine whether your participants understand the items in a consistent way. Be sure to also check the veracity of the responses of your participants (i.e., whether their answers are true and accurate). These strategies will help you determine whether the items actually measure what they are intended to measure. When the participants fill out the questionnaire, check to see whether they skip to the correct place if you have contingency questions in your questionnaire.

After the participants finish filling out the questionnaire, you can discuss the questionnaire with them individually or in group sessions. Explain the purpose of your questionnaire to them, and ask them whether they believe anything important was left out, whether the instructions were clear, and whether there were any items that stood out for any reason and probe for an explanation. If the questionnaire has an experimental manipulation embedded in it, be sure to check to see that the manipulation is working as intended. For example, if a statement or a vignette is supposed to increase empathy toward minority groups, ask your participants whether they understood it and whether they felt empathetic afterward. Ask participants to comment on the appearance and clarity of the presentation; for example, were there too many questions on a page? Was there not enough space to write responses? Was the questionnaire easily readable? Finally, check the responses and determine if there are too many “I don’t know” or “Does Not Apply” answers indicated. If so, you may be asking questions that are unclear or not applicable. After completing your pilot test, revise your questionnaire and then pilot test it again. Remember that you do not want to use a questionnaire in a research study until all of the kinks have been worked out.

PUTTING IT ALL TOGETHER

You now have the 15 principles of questionnaire construction at your disposal, and you should feel ready to start the construction of your own questionnaire. One good way to start your first questionnaire is to model it after an existing questionnaire that was properly constructed. Therefore, we have provided an example or model questionnaire in Exhibit 7.3; it is entitled the Research Methods Demonstration Questionnaire. Notice how the principles for questionnaire construction have been employed in this questionnaire. For example, take note of the appearance of the questionnaire and the ordering of the questions. This is an example of how a basic mixed questionnaire should look. Table 7.3 provides a checklist you should use to make sure you don’t forget any important ideas.
CHAPTER 7: HOW TO CONSTRUCT A QUESTIONNAIRE

EXHIBIT 7.3 Example of a Mixed Questionnaire

Research Methods Demonstration Questionnaire

1. Are you a college student who is currently taking a course on research methods?
   - Yes ----> Please Go To Question 2
   - No ----> Please Do Not Complete This Questionnaire Because It Is Focused On College Students Taking a Research Course. Thanks Anyway For Agreeing To Participate.

2. In your research methods class, are you using the textbook entitled: Educational Research: Quantitative, Qualitative, and Mixed Approaches, written by Johnson and Christensen?
   - Yes ----> Please Go To Question 3
   - No ----> Please Do Not Complete This Questionnaire Because It Is Focused On Current Users of the Johnson and Christensen Textbook. Thanks Anyway For Agreeing To Participate.

3. At what college or university are you currently taking this research methods class?

4. Is the Johnson and Christensen textbook the first book you have studied on research methods during the past 5 years?
   - Yes
   - No

5. How difficult do you find learning about research methods to be?
   - Very difficult
   - Somewhat difficult
   - Not very difficult
   - Not at all difficult
   - Don’t know

6. Which course do you think is more difficult, educational psychology or educational research methods?
   - Educational psychology
   - Educational research methods
   - Don’t know

(Continued)
Next, we want to know how interesting you find each of the following research method topics? Please respond with (1) not at all interesting, (2) not very interesting, (3) somewhat interesting, or (4) very interesting, or (9) don’t know. (Circle one response for each item.)

<table>
<thead>
<tr>
<th>Topic</th>
<th>Not At All Interesting</th>
<th>Not Very Interesting</th>
<th>Somewhat Interesting</th>
<th>Very Interesting</th>
<th>Don’t Know</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developing research questions</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Writing proposals</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Research ethics</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Measurement</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Data collection</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Sampling</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Validity of research results</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Data analysis</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Quantitative research</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Qualitative research</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Mixed research</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>9</td>
</tr>
</tbody>
</table>

18. Given sufficient study time, how much anxiety would you feel if you had to take a 100-item multiple-choice test on research methods?
   - □ A great deal of anxiety----- Go To Question 19
   - □ Some anxiety----- Go To Question 19
   - □ A little anxiety----- Go To Question 21
   - □ No anxiety----- Please Skip To Question 21
   - □ Don’t know----- Please Skip To Question 21

19. What do you think are the reasons for your test anxiety?

20. What might be done by your teacher to help reduce your test anxiety?
Next are three questions about the content of your research methods class.

21. Which of the following research terms refers to “a technique for physically obtaining data to be analyzed in a research study”?
   - [ ] Method of data collection
   - [ ] Method of research
   - [ ] Method of measurement
   - [ ] Method of data analysis
   - [ ] Don’t know

22. How many points should there generally be on a rating scale?
   - [ ] 4 points
   - [ ] 5 points
   - [ ] 10 points
   - [ ] Anywhere from 4 to 11 points is usually fine
   - [ ] Don’t know

23. What is the problem with this potential questionnaire item: “Teachers should have extensive contact with parents and school administrators.”
   - [ ] It is too long
   - [ ] It is a double-barreled question
   - [ ] It has no item stem
   - [ ] Don’t know

24. How useful do you think your knowledge of research methods will be in your career?
   - [ ] Very useful
   - [ ] Somewhat useful
   - [ ] Not very useful
   - [ ] Not at all useful
   - [ ] Don’t know

The next three items refer to how you feel about yourself. Please indicate your degree of agreement or disagreement with each item using the following scale: (1) strongly disagree, (2) disagree, (3) agree, or (4) strongly agree. (Circle one response for each item.)

<table>
<thead>
<tr>
<th>Item</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
<th>Don’t Know</th>
</tr>
</thead>
<tbody>
<tr>
<td>25.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>26.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>27.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>28.</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>9</td>
</tr>
</tbody>
</table>

(Continued)
29. Realistically, what final letter grade do you expect to get in your research methods course?
   □ A
   □ B
   □ C
   □ D
   □ F
   □ Don’t know

Last are some demographic questions that will be used for classification purposes only

30. What is your current college status?
   □ Undergraduate
   □ Graduate student
   □ Other (Please Specify): ______________________________

31. What is your gender?
   □ Female
   □ Male

32. Which of the following best describes your political party affiliation?
   □ Democrat
   □ Republican
   □ Independent
   □ Other (Please Specify): ______________________________

33. Approximately what was your personal income last year?
   _____ $

34. Is there anything that you would like to add for us to think about?

THANK YOU FOR COMPLETING OUR QUESTIONNAIRE!
Okay, so now that you have all your information, what is the next step? Figure 7.2 is a flow-chart that should help guide you through the construction of your first—or ten thousandth—questionnaire. Questionnaire construction is not a straight path. It is an iterative process with many twists and turns. Even the most experienced researcher at questionnaire construction will find that he or she has to go back and revise the instrument at some point in the process. Indeed, the more willing and able a researcher is to revise the questionnaire throughout the process, the better the questionnaire will be. Remember, your goal is to design a questionnaire that works well! Questionnaire construction takes time, but when you get it right, your research participants and the readers of your research reports will thank you for it.

**FIGURE 7.2**

Steps in Questionnaire Construction Flowchart

- **STEP 1. Review the relevant literature and begin planning the questionnaire.**
  
  Remember that if a questionnaire that fits your needs is already available, then there is no need to construct a new questionnaire.
  
  Think about:
  - Do you understand the targeted participants?
  - Do you understand the issues to be examined?
  - What variables do you want to measure?
  - What do you want to know in the participants' own words?
  
  Is the questionnaire to be self-administered, sent through the mail, or filled out on the Internet?
  
  Decision:
  - Select the type of questionnaire you need to construct: qualitative, quantitative, or mixed.
  (Continue with step 1 until you are ready to move to step 2.)

- **STEP 2. Write the items for the questionnaire.**
  
  Think about:
  - Have you examined other related questionnaires?
  - Have you examined items on other high-quality questionnaires that will be helpful models as you write yours?
  - Have you asked others (friends, family members, students) if your items are clear?

  Decision:
  - Are my questions simple and clear?
  - Are all of the issues covered?
  - Does my draft questionnaire look good so far?
  - (If the answer is no to any of these questions then continue working on step 2 or go back to step 1 if needed; otherwise, go to step 3.)

- **STEP 3. Design the layout and overall questionnaire.**
  
  Think about:
  - Does the questionnaire have a title, clear directions, section lead-ins, proper section ordering, demographics at the end, and a “thank you” at the conclusion?
  
  Have you asked others (colleagues) to critique your questionnaire?
  
  Decision:
  - Are the items and sections organized logically and clearly?
  - Do the skip patterns in the questionnaire clearly lead users to exactly where they need to go through the questionnaire? (If the answer is no to any of these questions then continue working on step 3 or go back to step 1 or 2 as needed; otherwise, go to step 4.)

- **STEP 4. Conduct a pilot test of the questionnaire.**
  
  Think about:
  - What people can you administer the questionnaire to who are similar to the kinds of people to be used in your research study?
  - Have I collected reliability and validity data?

*(Continued)*
SUMMARY

This chapter explains how to write items and construct a questionnaire to be used in collecting data in a research study. This might seem like a simple task, and it is not overly difficult, but it is imperative that you take this process very seriously and that you follow the appropriate steps and procedures that we have provided. Remember that if your data-collection instrument (i.e., your questionnaire) does not work well, then your results will be meaningless. When developing a good questionnaire, you need to understand and use the 15 principles of questionnaire construction (Table 7.2) discussed in this chapter. We provided a checklist that you should use to make sure you have not forgotten any important points (Table 7.3). We also provided an example of a well-written questionnaire that you can use as a model example when you start developing your own questionnaire (Exhibit 7.3). Over time, you will have your own collection of model questionnaires that you have developed, as you become an expert. Finally, we provided a flowchart showing you the iterative or cyclical steps in developing and continually improving your questionnaire (Figure 7.2).
CHAPTER 7: HOW TO CONSTRUCT A QUESTIONNAIRE

KEY TERMS

acquiescence response set (p. 186)  
anchor (p. 179)  
checklist (p. 183)  
closed-ended question (p. 176)  
contingency question (p. 187)  
double-barreled question (p. 175)  
double negative (p. 175)  
exhaustive (p. 178)  
fully anchored rating scale (p. 180)  
item stem (p. 177)  
leading question (p. 174)  
Likert scale (p. 185)  
loaded question (p. 174)  
mixed questionnaire (p. 177)  
mutually exclusive (p. 178)  
umerical rating scale (p. 179)  
open-ended question (p. 176)  
pilot test (p. 189)  
principle of standardization (p. 177)  
qualitative questionnaire (p. 177)  
quantitative questionnaire (p. 170)  
rating scale (p. 179)  
reverse-worded item (p. 186)  
semantic differential (p. 183)  
social desirability response set (p. 186)  
summated rating scale (p. 185)  
think-aloud technique (p. 189)  
web surveys (p. 187)  
ranking (p. 181)  
response set (p. 186)  

EDUCATIONAL RESEARCH ON THE WEB

Log on to the web-based student study site at http://www.sagepub.com/bjohnsonstudy for additional web sources and study resources.

RESEARCH EXERCISES

1. Fill out the Rosenberg Self-Esteem Scale shown in Figure 7.1. Then sum your responses to the ten items to obtain your overall score (i.e., your summated score). Be sure that you “reverse score” items 3, 5, 8, 9, and 10 (i.e., a 4 becomes a 1, a 3 becomes a 2, a 2 becomes a 3, and a 1 becomes a 4) before you add up your item scores to obtain your overall score. After doing this, you will know how to score a summated scale. Note that the way the Rosenberg Self-Esteem Scale is coded, lower scores are better (i.e., they represent higher self-esteem) and higher scores are worse (i.e., they represent lower self-esteem). Some researchers have their computer program reverse the final numbers so that higher scores represent higher self-esteem. For now, just be sure that you are careful in interpreting your score!

2. Pick a topic and construct a 15-item questionnaire. Collect data from five of your classmates. Have them evaluate your data-collection instrument (i.e., your questionnaire) on the basis of what they have learned in this chapter. Revise your questionnaire.

3. Go to ERIC or SocINDEX and conduct a journal article search using the term questionnaire. List five questionnaires that you found interesting. What was the purpose of each of these?

RELEVANT INTERNET SITES

Free brochures explaining survey research, including data collection and questionnaire construction  
www.amstat.org/sections/srms/whatsurvey.html

Very well done (and free) guide on tips for constructing questionnaires and interview protocols  

Site maintained by Don A. Dillman, a prominent survey research expert, and containing some of his papers  
http://survey.sesrc.wsu.edu/dillman
RECOMMENDED READING

