Conducting Educational Research
Fifth *Edition*

Bruce W. Tuckman
The Ohio State University

Harcourt Brace College Publishers
Fort Worth Philadelphia San Diego New York Orlando Austin San Antonio
Toronto Montreal London Sydney Tokyo

Publisher Earl McPeek
Acquisitions Editor Carol Wada
Market Strategist Laura Brennan
Project Editor Angela Williams Urquhart
Art Director Sue Hart
Production Manager Andrea A. Johnson

ISBN: O-15-505477-5

Library of Congress Catalog Card Number: 98.86589

Address for Orders
Harcourt Brace College Publishers, 6277 Sea Harbor Drive,
Orlando, FL 32887.6777 l-800-782-4479

***Address for Editorial** Correspondence*
Harcourt Brace College Publishers, 302 Commerce Street, Suite 3700,
Fort Worth, TX 76102

Web Site Address
http://www.hbcollege.com

Harcourt Brace College Publishers will provide complimentary supplements or supplement packages to those adopters qualified under our adoption policy. Please contact your sales representative to learn how you qualify. If as an adopter or potential user you receive supplements you do not need, please return them to your sales representative or send them to: Attn: Returns Department, Troy Warehouse, 465 South Lincoln Drive, Troy, MO 63379

Printed in the United States of America

8 9 0 1 2 3 4 5 6 7 0 3 9 1 0 9 8 7 6 5 4 3 2 1

Harcourt Brace College Publishers

CONTENTS

PREFACE

Many aspects of research design and methodology have changed very little in the quarter-century since this book first appeared. The challenge of formulating a researchable problem and labeling and operation&zing the variables that make up that problem remains today as it was then. The research designs, their graphic representation, and their sources of validity and invalidity, initially popularized in the first edition of this book, remain largely unchanged as well. The bases for evaluating measuring instruments, many of the statistical tests, and the format for reporting research show little difference from the versions presented in the initial appearance of this textbook.

But this fifth edition of Conducting *Educational Research reflects two* dramatic areas of change and a number of other changes of lesser note. While graduate students were learning to do research to complete theses and dissertations and then advance to become themselves the practitioners and researchers of today (and while my children, to whom, incidentally, every edition of this book has been dedicated, were growing up), we have been experiencing a technological revolution. This revolution has had its impact in educational research as in virtually all areas of life. When this book first appeared, statistical tests were mostly laboriously performed on "high-speed" desk calculators, which, like the dinosaur, have become extinct, The computer age, however, was in its infancy and most of us were learning how to run data on mainframe computers using punched cards and program control language.

Today, the availability and easy use of statistical software packages for use on personal computers enable researchers to remain at their desks and run any number of statistical analyses from simple tests to sophisticated analyses. In this edition of CER, instructions and descriptions for running the different statistical tests described in previous editions are now presented for personal computer software. In keeping with my strong belief that instruction in research methods should be as concrete and applications-oriented as possible, the process of carrying out statistical tests by computer is not presented in a general way, but in terms of the specific operations required to perform a number of the most commonly used statistical tests, using the Statistical Package for the Social Sciences (SPSS). This software program has been chosen because of its common familiarity and widespread use throughout the research world.

New technological developments have also transformed the process of literature searching, and new updates have been added to every *CER* edition, this latest one being no exception. The Internet has exploded into the everyday world, and only a new edition every year could keep up with its burgeoning development.

The second major change area in educational research has been the growing interest in and use of qualitative data to study and understand educational phenomena. While the first edition of *CER* described methods for collecting interview and observational data, it was not until the third edition that a full

chapter was added to cover qualitative research. This chapter has been considerably expanded in this edition to include more background on the nature and characteristics of the qualitative approach, more specifics on the "mechanics" of its procedures, and special sections on techniques for interviewing children, analyzing taped interview transcripts, and doing analysis of think-aloud verbal protocols, all of which reflect the increasing emphasis on exactness and objectivity in collection and analysis of qualitative data.

Smaller changes in the text reflect developments and growing and changing emphases in educational research methodology. These alterations cover a wide range of subjects, beginning with an expansion of the detail in the discussion of researchers' ethical requirements, reflecting the growing concern with the welfare of participants in research. Additional new material addresses issues regarding the validity of meta-analysis and controversies about some of the conclusions it has spawned. The text also covers techniques for employing a wider range of research designs, including single-subject, cross-sectional, and longitudinal designs, the latter two specifically intended to study developmental change.

All of the editions, including this one, focus on teaching researchers-to-be and practitioners alike to be astute and critical readers and users of research. As the universe of knowledge expands, critical evaluation of what one can and cannot accept as good research becomes a progressively more necessary yet more challenging task.

I would like to acknowledge the people who have helped me in the preparation of this edition. Foremost among them is Dr. Andy Palmer, one of my former doctoral students, who made a major contribution to the revision of the chapter on statistics. One can learn much, I am often reminded, from the very people one teaches. In this same spirit, my current doctoral student, Dennis Abry, has been most helpful. I am also gratified to be able to include the work on protocol analysis by my colleague, K. Anders Ericsson. Finally, I offer my thanks to those who assisted me by reviewing the fourth edition and offering suggestions for its improvement, including Dr. Michael Pressley, University of Notre Dame; Dr. Alan Klockars, University of Washington; and Dr. John Taccarino, De Paul University.

As always, I am grateful to my colleagues in the field who have used *CER as a text* for teaching educational research methods, and to the many students who have pored over its pages in an effort to discover the "mysteries" of research.

Bruce W. Tuckman

10 Chapter

CONSTRUCTING AND USING QUESTIONNAIRES AND INTERVIEW SCHEDULES

Objectives

- Identify the purposes served by questionnaires and interviews and the shortcomings of each.
- Identify different question formats and response modes, and describe their relative characteristics.
- Describe the bases for choosing between a questionnaire and an interview and for choosing a response mode for a specific set of conditions and purposes.
- Construct a series of items (for example, a questionnaire or interview) designed to answer specific research questions.
- Describe sampling procedures for questionnaire and interview studies.
- Lay out the process for administering a questionnaire, including the preparation of a cover letter.
- Describe procedures for conducting an interview.
- Describe the procedures for coding and scoring interview and questionnaire data.

WHAT DO QUESTIONNAIRES AND INTERVIEWS MEASURE?

Questionnaires and interviews help researchers to convert into data the information they receive directly from people (research subjects). By providing access to what is "inside a person's head," these approaches allow investigators to measure what someone knows (knowledge or information), what someone likes and dislikes (values and preferences), and what someone thinks (attitudes and beliefs). Questionnaires and interviews also provide tools for discovering what experiences have taken place in a person's life (biography) and what is occurring at the present. This information can be transformed into quantitative data by using the attitude or rating scales described in the previous chapter or by counting the number of respondents who give a particular response, which generates frequency data.

Questionnaires and interviews provide methods of gathering data about people by asking them rather than by observing and sampling their behavior. However, the self-report approach incorporated in questionnaires and interviews does present certain problems: (1) Respondents must cooperate to complete a questionnaire or interview. (2) They must tell what is rather than what they think ought to be or what they think the researcher would like to hear. (3) They must know what they feel and think in order to report it. In practice, these techniques measure not what people believe but what they say they believe, not what they like but what they say they like.

In preparing questionnaires and interviews, researchers should exercise caution. They must constantly consider:

- To what extent might a question influence respondents to show themselves in a good light?
- To what extent might a question influence respondents to attempt to anticipate what researchers want to hear or learn?
- To what extent might a question ask for information about respondents that they may not know about themselves?

The validity of questionnaire and interview items is limited by all three of these considerations. However, certain information can be obtained only by asking. Even when an alternative is available, simply asking subjects to respond may be (and often is) the most efficient one. Thus, the advantages and disadvantages of a questionnaire or interview as a source of data must be considered in each specific case before a decision can be made to use it or not to use it.

QUESTION FORMATS: HOW TO ASK THE QUESTIONS

Certain forms of questions and certain response modes are commonly used in questionnaires and interviews. This section deals with question formats and the following section addresses response modes.

Direct Versus *Indirect Questions*

The difference between direct and indirect questions lies in how obviously the questions solicit specific information. A direct question, for instance, might ask someone whether or not she likes her job. An indirect question might ask what she thinks of her job or selected aspects of it, supporting the researcher's attempt to build inferences from patterns of responses. By asking questions without obvious purposes, the indirect approach is the more likely of the two to engender frank and open responses. It may take a greater number of questions to collect information relevant to a single point, though. (Specific administrative procedures may help a researcher to engender frank responses to direct questions, as described later in the chapter.)

Specific Versus *Nonspecific Questions*

A set of specific questions focuses on a particular object, person, or idea about which a researcher desires input regarding an attitude, belief, or concept; nonspecific questions probe more general areas. For example, an interviewer can ask a factory worker (specifically) how he likes operating a lathe or (nonspecifically) how he likes operating machinery or working at manual tasks. An interviewer can ask a student (specifically) how much she likes a particular teacher versus (nonspecifically) how satisfied she feels with a particular class taught by the teacher. Specific questions, like direct ones, may cause respondents to become cautious or guarded and to give less-than-honest answers. Nonspecific questions may lead circuitously to the desired information while provoking less alarm by the respondent.

Questions of Fact Versus *Opinion*

An interviewer may also choose between questions that ask respondents to provide facts and those that request opinions. A factual question might ask a respondent the type of car he or she owns or to specify marital status. An opinion question might ask about preference for Ford or Chevrolet models or reasons why (or why not) a respondent thinks that marriage contributes to a meaningful

relationship between a man and a woman. Because the respondent may have a faulty memory or a conscious desire to create a particular impression, factual questions do not always elicit factual answers. Nor do opinion questions necessarily elicit honest opinions, because they are subject to distortions based on social *desirability;* that is, respondents may reply in ways that show themselves in the most socially acceptable light. With both fact and opinion questions, questionnaires and interviews may be structured and administered to minimize these sources of bias.

Questions Versus Statements

To gather input on many topics, an interviewer can either ask a respondent a direct question or provide a statement and ask for a response. To a question, a respondent provides an appropriate answer. For a statement, the respondent indicates whether he or she agrees or disagrees (or whether the statement is true or false). Applied in this manner, statements offer an alternative to questions as a way of obtaining information. In fact, attitude measurement instruments more commonly present statements than ask questions. Consider an example:

- Do you think that the school day should be lengthened? YES NO

  versus

. The school day should be shortened. AGREE DISAGREE

These two formats are indistinguishable in their potential for eliciting honest responses. Usually, researchers choose between them on the basis of response mode, as discussed in the next section.

Predetermined *Versus Response-Keyed Questions*

Some questionnaires predetermine the number of questions to be answered; they require respondents to complete all items. Others are designed so that subsequent questions may or may not call for answers, depending upon responses to keyed questions. For example, a keyed item may ask a respondent if he is a college graduate. If the response is no, the respondent is instructed to skip the next question. The decision whether or not to answer the question is keyed to the response to the previous question.

Consider another example of response keying. An interviewer asks a school superintendent if her district is using a nationally known curriculum. Two possible questions are keyed to the response. If the superintendent says that the district is using the curriculum, the next question asks about its effectiveness; if the superintendent says the district is not using the curriculum, the next question asks why.

RESPONSE MODES: HOW TO ANSWER THE QUESTIONS

Besides asking questions in a variety of ways, responses can take a multiplicity of forms or modes. This section reviews a number of different response modes.

Unstructured Responses

An *unstructured response*, perhaps more commonly referred to by the term *open-ended question* (although the response, not the question, is open-ended), allows the subject to give a response in whatever form he or she chooses. Open-ended and nonopen-ended questions may target identical information. The difference between an unstructured (open-ended) question from a structured one centers on the type of response that the respondent is allowed to make. For instance, a question might ask if a respondent thinks that schools should not grade assigned work; if the respondent says yes, another question asks why he thinks so. The resulting unstructured response might take several minutes and include a series of arguments, facts, ramblings, and so on. A structured response format would offer, say, five reasons and ask the respondent to choose one.

Here are some examples of questions in the unstructured response mode:

- Why do you think you didn't try harder in high school?

  _____

  _____

  _____

. What led you to go to college?

  _____

  _____

  _____

- Describe your feelings as you think of your mother.

  _____

  _____

  _____

Items II and IV in Figure 10.1 provide additional examples of questions in the unstructured response mode.

Thus, the unstructured response mode is a responsive form over which the researcher attempts to exert little control other than by asking questions and limiting the amount of space (or time) provided for the answers. Once an unstructured question is asked, the response may be stated in the way the respondent chooses, Allowing the respondent such control over the response ensures that the

A SAMPLE QUESTIONNAIRE **FIGURE 10.1**

I. **Suppose** you were offered an opportunity to make a substantial advance in a job or occupation,. Place a check opposite each item in the following list to show how important it would be in stopping you from making that advance.

	Would stop me	Might stop me from making change	Would be a serious consideration but wouldn't stop me	Wouldn't matter at all
Endanger your health				
Leave your amily for some time				
Move around the country a lot				
Leave your community				
Leave your friends				
Give up leisure time				
Keep quiet about political views				
Learn a new routine				
Work harder than you are now				
Take on more responsibility				

II. Looking at your present situation, what do you expect to be doing *5* years from now? _____

III. What are your chances of reaching this goal?
__ excellent _ good ____ fair ____ poor _ very poor

IV. What would you like to be doing *5* years from now? _____

_____

V. What are your chances of reaching this goal?
- excellent _ good ____ fair ____ poor - very poor

respondent will give his or her own answers rather than simply agreeing with one provided by the researcher.

However, the unstructured mode does raise problems in quantification of data and ease of scoring (discussed in detail in the last section of the chapter, which covers coding and scoring procedures). In contrast, more structured response modes simplify quantification.

Fill-In Response

The fill-in response mode can be considered a transitional mode between unstructured and structured forms. Although it requires the subject to generate rather than choose a response, it typically limits the range of possible responses by limiting the answer to a single word or phrase. Consider the following examples:

- What is your father's occupation? _____
- In what school did you do your undergraduate work? _____
- Looking at the above picture, what word best describes the way it makes you feel? _____

Note that the unstructured response mode differs from the structured, fill-in mode in degree. The fill-in mode restricts respondents to a single word or phrase, usually in a request to report factual information (although the third example elicits a response beyond facts). The very wording of such a question restricts the number of possible responses the respondent can make and the number of words that can be used.

Tabular Response

The tabular response mode resembles the fill-in mode, although it imposes somewhat more structure because respondents must fit their responses into a table. Here is an example:

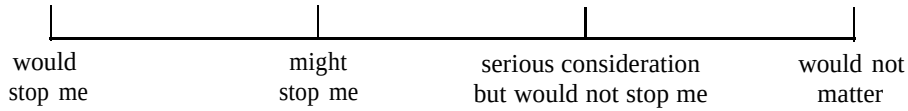
Next Previous Job Title	Specify Type of Work Performed	Name of Employer	Annual Salary	Dates	
				From	To

Typically, a tabular response requires numbers, words, or phrases (often factual information of a personal nature), but it may also allow respondents to reflect their degree of endorsement or agreement along some scale, as shown in Item I in Figure 10.1. (This use of the tabular mode is described in more detail in the following section on scaled response.)

A table is a convenient way of organizing a complex response, that is, a response that includes a variety of information rather than a single element. However, it is otherwise not a distinct response mode. The tabular form organizes either fill-in responses (as in the example) or scaled responses (as in Item I, Figure 10.11.

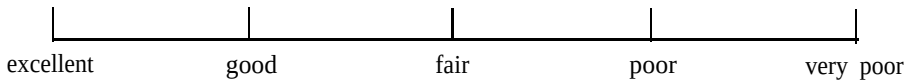
Scaled Response

A commonly used structured response mode establishes a scale (that is, a series of gradations) on which respondents express endorsement or rejection of an attitude statement or describe some aspect of themselves. Item I in Figure 10.1 (which uses the tabular form of organization) illustrates the scaled response mode. Note that the question asks the respondent to consider each potential obstacle to job advancement and indicate on the scale the effect of that concern on his or her acceptance of a new job:



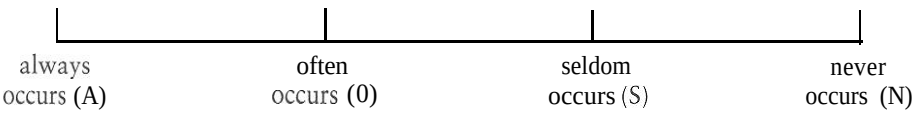
This example illustrates a four-point scale of degree of influence, from total influence at the left to no influence at the right.

Consider also Items III and V in Figure 10.1. Identical in wording but referring to different goals, they ask the respondent to assess his or her likelihood of reaching a goal, using the following five-point scale:



By choosing one of these five categories, the respondent indicates the degree to which he or she sees goal attainment as a likely prospect.

The Career Awareness Scale is an example of a questionnaire that uses a scale to indicate *frequency*. *(See* Figure 10.2.) The instrument presents a descriptive statement about career-seeking behavior to a respondent, a high school student, and asks for an indication of the frequency with which this behavior occurs, using the following four-point scale:



The scale is used primarily to assess whether a high school student has engaged in behaviors intended to learn about careers.

All scaled responses measure degree or frequency of agreement or occurrence (although a variety of response words may indicate these quantities). They all assume that a response on a scale is a quantitative measure of judgment or feeling. (Recall that Chapter 9 discussed priorities for constructing such a scale.) Unlike an unstructured response, which requires coding to generate useful data, a structured, scaled response collects data directly in a usable and analyzable form. Moreover, in some research situations, scaled responses can yield interval data.[1]

[1] *See* the early part of Chapter *9* for a discussion of the types of measurement scales.

FIGURE 10.2 A FREQUENCY QUESTIONNAIRE: THE CAREER AWARENESS SCALE

Instructions: All of the questions below are about what you actually do. If you "Always" do what the statement says, circle the 1 for A. If you "Often" do what the statement says, circle the **2** for 0. If you "Seldom" do what the statement says, circle the 3 for S. If you "Never" do what the statement says, circle the 4 for N.

There are no right or wrong answers for these questions. We are interested only in what you **actually** *do.*

1. I think about what I will do when I finish school.	1. A 2. 0 3. S 4. N
2. I read occupational information.	1. A 2. 0 3. S 4. N
3. I visit my guidance counselor to talk about my future.	1. A 2. 0 3. S 4. N
4. I attend "career days" held in school.	1. A 2. 0 3. S 4. N
5. I think about what it will take to be successful in my occupation.	1. A 2. 0 3. S 4. N
6. I talk to workers to learn about their jobs.	1. A 2. 0 3. S 4. N
7. Before I go on a field trip, I read whatever information is available about the place I am going to visit.	1. A 2. 0 3. S 4. N
8. I look at the "Want Ads" in order to find out about jobs.	1. A 2. 0 3. S 4. N
9. I visit factories, offices, and other places of work to learn about different kinds of jobs.	1. A 2. 0 3. S 4. N
10. I take advantage of opportunities to do different things so that I'll learn about my strengths and weaknesses.	1. A 2. 0 3. S 4. N
11. I keep myself prepared for immediate employment should the necessity arise.	1. A 2. 0 3. S 4. N
12. I talk with my parents about my choice of career.	1. A 2. 0 3. S 4. N
13. I work at different kinds of part-time jobs.	1. A 2. 0 3. S 4. N
14. When the school gives an interest or career aptitude test, I take it seriously.	1. A 2. 0 3. S 4. N
15. I consider my own values, my own abilities, and the needs of the job market when I plan my career.	1. A 2. 0 3. S 4. N

For example, the difference in frequency between N and *S* on the Career Awareness Scale would be considered equivalent to the differences between *S* and 0 and between 0 and A. Provided other requirements are met, such interval data can be analyzed using powerful parametric statistical tests. (These statistical procedures are described in Chapter 11.)

Ranking Response

If a researcher presents a series of statements and asks the respondent to rank order them in terms of a particular criterion, the question will generate ordinally arranged results. Consider an example:

- Rank the following activities in terms of their usefulness to you as you learn how to write behavioral objectives. (Assign the numbers 1 through 5, with 5 indicating the most useful activity and 1 indicating the least useful one. If any activity gave you no help at all, indicate this by a 0.)
  _ Initial presentation by consultants
  _ Initial small-group activity
  _ Weekly faculty sessions
  _ Mailed instructions and examples of behavioral objectives
  _ Individual sessions with consultant

Ranking forces respondents to choose between alternatives. If respondents were asked to rate (that is, scale) each alternative or to accept or reject each one, they could assign them all equal value. A request for a ranking response forces them to give critical estimates of the values of the alternatives.

Typically, ranked data are analyzed by summing the ranks that subjects assign to each response, giving an overall or group rank order of alternatives. Such an overall ranking generated by one group (for example, teachers) can be compared to that generated by a second group (for example, administrators) using nonparametric statistical techniques. (See Chapter 11.)

Checklist Response

A respondent replies to a checklist item by selecting one of the possible choices offered. This form of response does not, however, represent a scale, because the answers do not represent points on a continuum; rather they are *nominal categories.* Consider two examples:

- The kind of job that I would most prefer would be:
  Check one:
  __ (1) A job where I am almost always certain of my ability to perform well.
  _ (2) A job where I am usually pressed to the limit of my abilities.
- I get most of my professional and intellectual stimulation from:
  Check one of the following blanks:
  _ A. Teachers in the system
  _ B. Principal
  _ C. Superintendent
  _ D. Other professional personnel in the system

_ E. Other professional personnel elsewhere
_ F. Periodicals, books, and other publications

Respondents often find the nominal judgments required by a checklist easier to make than scalar judgments, and they take less time to give such responses. At the same time, those responses yield less information for the researcher. Nominal data are usually analyzed by means of the chi-square statistical analysis (described in Chapter 11).

Categorical Response

The categorical response mode, similar to the checklist but simpler, offers a respondent only two possibilities for each item. (In practice, checklist items also usually offer only two responses: check or no check on each of a series of choices, but they may offer more possibilities.) However, the checklist evokes more complex responses, since the choices cannot be considered independently, as can categorical responses. Also, after checking a response, the remaining choices in the list leave no further option.)

A yes-no dichotomy is often used in the categorical response mode:

- Are you a veteran? Yes _ No ___

Attitude-related items may give true-false alternatives:

- Guidance counseling does not begin early enough.
  True _ False _

Analysis can render true-false data into interval form by using the number of true responses (or the number of responses indicating a favorable attitude) as the respondent's score. The cumulative number of *true* responses by an individual *S* on a questionnaire then becomes an indication of the degree (or frequency) of agreement by that S-an interval measure. Counting the number of *Ss* who indicate agreement on a single item provides a nominal measure. (See the section on coding and scoring at the end of this chapter to see how to score this and the other types of response modes.)

CONSTRUCTING A QUESTIONNAIRE OR INTERVIEW SCHEDULE

How do you construct a questionnaire or interview schedule? What questions should you ask and in what formats? What response modes should you employ? To answer, begin by asking, "What am I trying to find out?"

Specifying the Variables to Measure

The questions you should ask on a questionnaire or in an interview reflect the information you are trying to find, that is, your hypotheses or research questions. To determine what to measure, you need only write down the names of all the variables you are studying. One study might attempt to relate source of occupational training (that is, high school, junior college, or on-the-job instruction) to degree of geographic mobility; it would have to measure where respondents were trained for their jobs and the places where they have lived. A study might compare 8th graders and 12th graders to determine how favorably they perceive the high school climate; it would have to ask respondents to indicate their grade levels (8th or 12th) and to react to statements about the high school climate in a way that indicates whether they see it as favorable or not. A study concerned with the relative incomes of academic and vocational high school graduates 5 years after graduation would have to ask respondents to indicate whether they focused on academic or vocational subjects in high school and how much money they were presently earning.

Thus, the first step in constructing questionnaire or interview questions is to *specify your variables by name. Your* variables designate what you are trying to measure. They tell you where to begin.

Choosing the Question Format

The first decision you must make about question format is whether to present items in a written questionnaire or an oral interview. Because it is a more convenient and economical choice, the questionnaire is more commonly used, although it does limit the kinds of questions that can be asked and the kinds of answers that can be obtained. A questionnaire may present difficulties in obtaining personally sensitive and revealing information. Also, it may not yield useful answers to indirect, nonspecific questions. Further, preparation of a questionnaire must detail all questions in advance. Despite the possibility of including some limited response-keyed questions, you must ask all respondents the same questions. Interviews offer the best possibilities for gathering meaningful data from response-keyed questions.

Table 10.1 summarizes the relative merits of interviews and questionnaires. Ordinarily, a researcher opts for the additional cost and unreliability of interviewing only when the study addresses sensitive subjects and/or when personalized questioning is desired. (Interviews are subject to unreliability, because the researcher must depend on interviewers to elicit and record the responses and often to code them, as well.) In general, when a researcher chooses to use the unstructured response mode, interviewing tends to be the better choice because people find it easier to talk than write; consequently, interviews generate more information of this type.

The choice of question format depends on whether you are attempting to measure facts, attitudes, preferences, and so on. In constructing a questionnaire,

TABLE 10.1

SUMMARY OF THE RELATIVE MERITS OF INTERVIEWS VERSUS QUESTIONNAIRES

CONSIDERATION	INTERVIEW	QUESTIONNAIRE
Personnel needed to collect data	Interviewers	Clerks
Major expense categories	Payments to interviewers	Postage and printing
Opportunities for response keying (personalization)	Extensive	Limited
Opportunities for asking	Extensive	Limited
Opportunities for probing (following leads)	Possible	Difficult
Relative magnitude of data reduction	Great (because of coding)	Mainly limited to rostering
Number of respondents typically reached	Limited	Extensive
Rate of return	Good	Poor
Sources of error	Interviewer, instrument, coding, sample	Limited to instrument and sample
Overall reliability	Quite limited	Fair
Emphasis on writing skill	Limited	Extensive

use direct, specific, clearly worded questions, and keep response keying to a minimum. In constructing an interview schedule, you may sacrifice specificity for depth and use indirect, subtle probes to work into an area of questioning. Response-keyed questions-those whose answers guide the choices of subsequent questions, if any, to ask-are also recommended as a labor-saving shortcut.

Choosing the Response Mode

No specific rules govern selection of response modes. In some cases, the kind of information you seek will determine the most suitable response mode, but often you must choose between equally acceptable forms. You can, for instance, provide respondents with a blank space and ask them to fill in their ages, or you can present a series of age groupings (for example, 20–29, 30–39, and so on) and ask them to check the one that fits them.

The choice of response mode should be based on the manner in which the data will be treated; unfortunately, however, researchers do not always make this decision before collecting data. It is recommended that data analysis decisions be made in conjunction with the selection of response modes. In this way, the researcher **(1) gains assurance** that the data will serve the intended purposes and (2) can begin to construct data rosters and to prepare for the analyses. (See Chapter 11.) If analytical procedures will group age data into ranges to provide nominal data for a chi-square statistical analysis, the researcher would want

CONSIDERATIONS IN SELECTING A RESPONSE MODE **TABLE 10.2**

RESPONSE MODE	TYPE OF DATA	CHIEF ADVANTAGES	CHIEF DISADVANTAGES
Fill-in	Nominal	Limited bias; expanded response flexibility	Difficult to score
Scaled	Interval	Easy to score	Time-consuming; potential for bias
Ranking	Ordinal	Easy to score; forces discrimination	Difficult to complete
Checklist or categorical	Nominal (may be interval when totaled)	Easy to score; easy to respond	Limited data and options

Note: The tabular mode is just a way of organizing till-in or scaled responses, so this table omits it as a distinct category.

to design the appropriate questionnaire item to collect these data in grouped form.

Scaled responses lend themselves most readily to parametric statistical analysis, because they often can be considered interval data. Ranking procedures may provide less information, because they generate ordinal data. Fill-in and checklist responses usually provide nominal data, suitable, unless otherwise coded, for chi-square analysis. Thus, the ultimate criterion in choosing a response mode is the nature of your variables and your intentions for statistically testing your hypotheses.[2] If the statistical tests for data analysis are not determined in advance, the best rule of thumb is to use the scaled response mode, because the interval data so collected can always be transformed into ordinal or nominal data. (See Chapter 11.)

Certain other practical considerations also influence the choice of response modes. Respondents may need more time to provide scaled responses than they would take to give true-false responses (and the researcher may spend more time scoring scaled responses). If your questionnaire is already lengthy, you may prefer the true-false response mode for additional questions in order to limit the burden upon the respondent. Fill-ins offer the advantage of not biasing the respondent's judgment as much as the other types, but they carry the disadvantage of difficulty in scoring or coding. Response-keyed questions provide respondents with response flexibility, but, like the fill-ins, they may be more difficult than other alternatives to score and do not provide parallel data for all respondents. Some of these considerations are summarized in Table 10.2.

Thus, selection between response modes requires consideration of several criteria:

[2] Conversely, should the response mode be specified first, it should be the criterion for choosing statistical tests.

1. Type *of* ***data*** *desired for analysis.* If you seek interval data to allow some type of statistical analysis, scaled and checklist responses are the best choices. (Checklist items must be coded to yield interval data, and responses must be pooled across items. An individual checklist item yields only nominal data.) Ranking provides ordinal data, and fill-in and some checklist responses provide nominal data.
2. *Response flexibility.* Fill-ins allow respondents the widest range of choice; yes-no and true-false items, the least.
3. *Time to complete.* Ranking procedures generally take the most time to complete, although scaled items may impose equally tedious burdens on respondents.
4. Potential response bias. Scaled responses and checklist responses offer the greatest potential for bias. Respondents may be biased not only by social desirability considerations but also by a variety of other factors, such as the tendencies to overuse the *true* or yes answer and to select one point on the scale as the standard response to every item. Other respondents may avoid the extremes of a rating scale, thus shrinking its range. These troublesome tendencies on the part of respondents are strongest on long questionnaires, which provoke fatigue and annoyance. Ranking and fill-in responses are less susceptible than other choices to such difficulties. In particular, ranking forces respondents to discriminate between response alternatives.
5. *Ease ofscoring.* Fill-in responses usually must be coded, making them considerably more difficult than other response types to score. The other types of responses discussed in this chapter are approximately equally easy to score.

Preparing Interview Items

As pointed out earlier, the first step in preparing items for an interview schedule is to specify the variables that you want to measure; then construct questions that focus on these variables. If, for example, one variable in a study is openness of school climate, an obvious question might ask classroom teachers, "How open is the climate here?" Less direct but perhaps more concrete questions might ask, "Do you feel free to take your problems to the principal? Do you feel free to adopt new classroom practices and materials?" Note that the questions are based on the operational definition of the variable *openness,* which has been operationally defined as freedom to change, freedom to approach superiors, and so on. In writing questions, make sure they incorporate the properties set forth in the operational definitions of your variables. (Recall from Chapter 6 that these properties may be either dynamic or static, depending on which type of operational definition you employ.)

A single interview schedule or questionnaire may well employ more than one question format accommodating more than one response mode. The sample interview schedule in Figure 10.3 seeks to measure the attitudes of the general

A PORTION OF A TELEPHONE INTERVIEW SCHEDULE

FIGURE 10.3

Reprinted by permission of the Eagleton Institute of Politics, Rutgers University, New Brunswick, N.J.

Now, I've got a few questions about public education in New Jersey.

21. Students are often given the grades A, B, C, D, and FAIL to denote the quality of their work. Suppose the public schools, themselves, in your community were graded in the same way. What grade would you give the public schools in your community—A, B, C, D, or FAIL?

135—1. A
2. B
3. C
4. D
5. FAIL
9. DON'T KNOW

22. In your opinion, is enough money being spent on public schools in your community?

136—1. YES → SKIP TO Q. 24, PAGE 6
2. NO
9. DON'T KNOW → SKIP TO Q. 24, PAGE 6

(IF "NO" TO Q. 22, ASK:

23. In order to spend more money on public schools, would you be willing to see an increase in local taxes?

137— 1. YES
2. NO
9. DON'T KNOW

24. At the present level of spending for public schools in your community, do you feel the citizens of your community are getting their money's worth?

138— 1. YES
2. NO
9. DON'T KNOW

25. Do you think that students should be required to pass a state examination in order to graduate from high school?

139— 1. YES
2. NO
9. DON'T KNOW

26. Which of the following three areas do you think is most important for the schools to concentrate on? [START AT DESIGNATED POINT AND READ CHOICES]

140— 1. TEACHING MATHEMATICS AND READING
2. PREPARING STUDENTS FOR COLLEGE OR A CAREER
OR 3. TEACHING A CODE OF VALUES AND MORAL BEHAVIOR
9. DON'T KNOW

27. Which is second most important? [READ REMAINING TWO CHOICES]

141— 1. TEACHING MATHEMATICS AND READING
2. PREPARING STUDENTS FOR COLLEGE OR A CAREER
OR 3. TEACHING A CODE OF VALUES AND MORAL BEHAVIOR
9. DON'T KNOW

28. How much attention do you think the schools pay to people like you when they decide what to teach—a great deal of attention, some attention, or no attention at all?

142— 1. A GREAT DEAL OF ATTENTION
2. SOME ATTENTION
3. NO ATTENTION AT ALL
9. DON'T KNOW

29. Turning to the matter of discipline, are the schools in your community too strict, not strict enough, or just about right?

143— 1. TOO STRICT
2. JUST ABOUT RIGHT
3. NOT STRICT ENOUGH
9. DON'T KNOW

public toward some current issues in public education such as cost, quality, curriculum emphasis, and standards. The interview schedule is highly structured to maximize information obtained in minimal telephone time. One of the questions is response keyed. All of them are specific, and all responses are precoded in scaled, categorical, or checklist form.

FIGURE 10.4 A FOLLOW-UP QUESTIONNAIRE

From Butler, 1977. Reprinted by permission of the author.

1. What is the title of your present job? _____
   From _____ TO _____
2. What is the title of your next previous job? _____
   From _____ TO _____
3. Check one of the following to show how you think you compare with other people.
   A - I like my work much better than most people like theirs.
   B _____ I like my work better than most people like theirs.
   C _____ I like my work about as well as most people like theirs.
   D _____ I dislike my work more than most people dislike theirs.
   E _____ I dislike my work much more than most people dislike theirs.
4. Check one of the following to show how much of the time you feel satisfied with your job.
   A _____ Most of the time B _____ A good deal of the time
   C __ About half of the time D _ _ Occasionally E _____ Seldom
5. Put a check before the category which most accurately describes your total, personal income in 1975 before taxes.
   A - Less than $5,000.00 B _____ Less than $10,000.00
   C _____ Less than $15,000.00 D _____ Less than $20,000.00
   E _____ $20,000.00 or more
6. Was there anything unusual (e.g., sickness, layoffs, promotions, unemployment) about your income in 1975 as reported in question #5 above?
   CIRCLE ONE: YES NO If YES, please explain _____
7. If you answered YES to question 6 above, put a check before the category which most accurately describes your total, personal income in 1974 before taxes.
   A - Less than $5,000.00 D _____ Less than $20,000.00
   B _____ Less than $10,000.00 E _____ $20,000.00 or more
   C _____ Less than $15,000.00
8. Are you a high school graduate? CIRCLE ONE: YES NO
   What high school? _____
9. Have you successfully completed one, two, or three college courses as a part-time student? YES NO
10. Have you successfully completed more than three college courses as a part-time student? YES NO
11. Have you attended a Z-year college as a full-time student without graduating? YES NO
12. Have you earned a 2-year college diploma? YES NO
    If YES to 9, 10, 11, or 12, what college? _____

Preparing ***Questionnaire Items***

The procedures for preparing questionnaire items parallel those for preparing interview schedule items. Again, maintain the critical relationship between the items and the study's operationally defined variables. Constantly ask about your items: Is this what I want to measure? Three sample questionnaires appear in Figures 10.4, 10.5, and 10.6.

(continued)

13. Have you enrolled in a 4-year college and successfully completed one, two, or three years? YES NO
    If YES, how many years and what college? _____
14. Have you earned the bachelor's degree? YES NO
    If YES, what college? _____
15. Have you earned a degree beyond the bachelor's? YES NO
    If YES, what degree and what college? _____
16. What was your father's job title at the time you graduated from high school? _____
17. What was your mother's job title at the time you graduated from high school? _____
18. How many brothers and sisters do you have? CIRCLE ONE:
    1 2 3 4 5 6 7 8 9 10 11 12 13 14
19. How many brothers and sisters are older than you? CIRCLE ONE:
    1 2 3 4 5 6 7 8 9 10 11 12 13 14
20. Are you a veteran? CIRCLE ONE: YES NO
    Dates of service: from _____ to _____
21. Describe your health since graduation from high school. CIRCLE ONE:
    EXCELLENT GOOD AVERAGE FAIR POOR
22. How many months have you been out of work because of illness since graduation from high school?
    CIRCLEONE: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
    17 18 19 NONE OTHER _____
23. Did you receive your training for the job in which you are now employed: (CHECK ONE BELOW)
    __ high school armed forces _____
    __ technical institute in-plant training _____
    __ L-year college apprenticeship _ _
    _____ 4-year college other _____ please explain _____
24. Marital status. CHECK ONE (or more):
    _____ single separated _____
    _____ married remarried _____
    _____ divorced other _____ please explain _____
25. Do you consider yourself a member of a minority group? YES NO
    CIRCLE ONE:
    If YES, check one: Black _____ American Indian __
    Chicano _____ other _____ please explain _____

The questionnaire in Figure 10.4 was used in a follow-up study of community college graduates and high school graduates who did not attend college. The researcher was interested in determining whether the community college graduates subsequently obtained higher socioeconomic status (that is, earnings and job status) and job satisfaction than a matched group of people who did not attend college. The items on the questionnaire were designed to determine (1) earnings, job title, and job satisfaction (the dependent variables) [Items 1–7]; (2) subsequent educational experiences, in order to eliminate or reclassify subjects pursuing additional education (a control variable) as well as to verify the educational status distinction of 2-year college students versus those who completed high school only (the independent variable) [Items 8-15, 23]; (3) background characteristics, in order to match samples [Items 16-20, 24, 25]; and (4) health, in order to eliminate those whose job success chances were impaired [Items 21, 22].

The researcher intended for all respondents to complete all of the items except Item 7, which was response keyed to the preceding item. (Items 12 to 15 also have response-keyed parts.) The result is a reasonably simple, easy-to-complete instrument.

The sample questionnaire in Figure 10.5 employs scaled responses in an attempt to measure students' attitudes toward school achievement based on the value they place on going to school and on their own achievement. This questionnaire actually measures the following six topics related to a student's perceived importance or value of school achievement:

1. Quality of school performance (Items 2, 3)
2. Importance of school (Items 1, 4, 8, 18)
3. Enjoyment of school (Items 5, 6, 7)
4. Pride taken in school performance (Items 9, 10, 14, 19)
5. Enjoyment of class participation (Items 11, 12, 13)
6. Importance of performing well (Items 15, 16, 17)

Note that for each of the 19 items, the questionnaire provides a 4-point scale for responses employing the statement format. (This sample resembles the standard Likert scale shown in Chapter 9, except that it omits the middle or "undecided" response.) Note further that some of the items have been *reversed (Items* 2, 5, 6, 9, 13, 14. These questions have been written so that *disagreement* or strong *disagreement* indicates an attitude favoring the importance of school achievement; on all the other items, agreement or strong *agreement* indicates such an attitude. *Agreement* with Item 10 for example, indicates that the respondent takes pride in school progress and performance, a reflection of a positive attitude toward school achievement. *Disagreement* with Item 9 indicates that the respondent does not feel that grades are unimportant, another reflection of a positive attitude toward school achievement.

Reversing direction in some items is a protection against the form of response bias caused when an individual simply selects exactly the same response choice

for each item. This tendency to mark a single choice for all questions out of boredom, disinterest, or hostility is referred to as *acquiescence response bias. Item reversal* guards against respondents creating erroneous impressions of extremely positive or extremely negative attitudes, because responses to items written in one direction cancel out or neutralize items written in the other.

FIGURE 10.5

A QUESTIONNAIRE ON STUDENTS' ATTITUDES TOWARD SCHOOL ACHIEVEMENT

Instructions: All questions are statements to which WC seek your agreement or disagreement. If you "Strongly Agree" with any statement, circle the 1. If you "Agree," but not strongly, with any statement, circle the 2. If you "Disagree," but not strongly, circle the 3. If you "Strongly Disagree" with any statement, circle the 4.

There are no right or wrong answers for these questions. We are interested only in *how you feel* about the statements.

1. I believe it is important for me to participate in school activities.	1. SA 2. A 3. D 4. SD
2. I do poorly in school.	1. SA 2. A 3. D 4. SD
3. I think I am a good student.	1. SA 2. A 3. D 4. SD
4. I believe education can offer many achievements	1. SA 2. A 3. D 4. SD
5. Schoolwork is uninteresting.	1. SA 2. A 3. D 4. SD
6. Schoolwork bores me.	1. SA 2. A 3. D 4. SD
7. I am happy to be a student.	1. SA 2. A 3. D 4. SD
8. I believe school is challenging.	1. SA 2. A 3. D 4. SD
9. Grades are not important to me.	1. SA 2. A 3. D 4. SD
10. I take pride in my progress and performance in school.	1. SA 2. A 3. D 4. SD
11. I enjoy volunteering answers to teachers' questions.	1. SA 2. A 3. D 4. SD
12. I feel good when I give an oral report.	1. SA 2. A 3. D 4. SD
13. I dislike answering questions in school.	1. SA 2. A 3. D 4. SD
14. Success in extra-curricular activities means very little.	1. SA 2. A 3. D 4. SD
15. I feel depressed when I don't complete an assignment.	1. SA 2. A 3. D 4. SD
16. I feel good when I am able to finish my assigned homework.	1. SA 2. A 3. D 4. SD
17. I believe it is my responsibility to make the honor roll.	1. SA 2. A 3. D 4. SD
18. School offers me an opportunity to expand my knowledge.	1. SA 2. A 3. D 4. SD
19. I do well in school so that my parents can be proud of me.	1. SA 2. A 3. D 4. SD

To maximize the effectiveness of this safeguard, half of the items should be written in each direction. Note that in the sample questionnaire in Figure 10.5, only 6 of the 19 items have been reversed.

The likelihood of this form of response bias is lessened also by the elimination of the "undecided" response alternative from the basic Likert scale format. The possibility of noninvolvement or "fence sitting" is avoided by omitting the "undecided" response choice.

Note also that the sample questionnaire obscures its true purpose somewhat by measuring multiple features of the topic in question rather than a single one. As a questionnaire's purpose becomes more transparent or obvious, the likelihood increases that respondents will provide the answers they want others to hear about themselves rather than the truth. This tendency to respond in a way that shows oneself in the best possible light is referred to as *social desirability response bias. It can* be minimized by not revealing the true name or purpose of the questionnaire prior to its completion, by including items that measure a variety of topics or aspects of a single topic, or by including filler items-questions that deal with areas unrelated to the one being measured. The sample questionnaire combines the first approach (carrying the title "Questionnaire" when actually administered) and the second (including multiple topics). No filler items appear in this version of the questionnaire, but some may appear in longer versions. However, a questionnaire about attitudes toward school achievement may not benefit from efforts to disguise the nature of the attitude being measured, so some responses can be expected to reflect social desirability bias rather than true feelings.

A third sample questionnaire is shown in Figure 10.6. The Satisfaction Scale is used to determine the degree of students' satisfaction with a course. Items are scaled using a S-point scale as the response mode. Note the format, which features a question followed by the response choices stated in both numbers and words.

No attempt has been made to counteract response bias by reversing the direction of some of the items or by disguising their meanings; each item has been written so that a *1* indicates a positive response on a single topic. Obviously, these items are susceptible to biased response based on considerations other than the respondent's judgment.

Pilot Testing and Evaluating a Questionnaire

Most studies benefit substantially from the precaution of running pilot tests on their questionnaires, leading to revisions based on the results of the tests. A pilot test administers a questionnaire to a group of respondents who are part of the intended test population but who will not be part of the sample. In this way, the researcher attempts to determine whether questionnaire items achieve the desired qualities of measurement and discrimination.

If a series of items is intended to measure the same variable (as the eight items in Figure 10.6 are), an evaluation should determine whether these items are mea-

A COURSE SATISFACTION QUESTIONNAIRE **FIGURE 10.6**

Name _____ **Course Name** _____

Teacher _____

SATISFACTION SCALE

1. Do you ever feel like skipping this class?

1	2	3	4	5
never	rarely	sometimes	often	always

2. Do you like this class?

1	2	3	4	5
very much	quite a bit	it's all right	not much	hate it

3. How much do you feel you have learned in this class?

1	2	3	4	5
a great deal	quite a bit	a fair amount	nor much	nothing

4. Are you glad you chose or were assigned to be in this class?

1	2	3	4	5
very glad	most of the time	sometimes	not too often	not at all

5. Do you always do your best in this class?

1	2	3	4	5
all the time	most of the time	sometimes	usually not	never

6. Do you like the way this class is taught?

1	2	3	4	5
very much	quite a bit	a fair amount	not much	not at all

7. Does the teacher give you help when you need it?

1	2	3	4	5
always	most of the time	usually	sometimes	never

8. Do you find the time you spend in this class to be interesting?

1	2	3	4	5
very much	quite	fairly	not too	not at all

suing something in common. Such an analysis would require administering the scale to a pilot sample and running correlations between response scores obtained by each person on each item and the scores obtained by each person across the whole scale. (See the discussion on item analysis in the previous chapter.) As the correlation between an item score and the total score rises, it indicates a stronger relationship between what the item is *measuring* and what the total scale is measuring. Following the completion of this item analysis, the researcher can select the items with the highest correlations with the total score and incorporate them in the final scale. For example, consider 10 items to measure a person's attitude toward some object, giving the following correlations between each item score and the mean score across all 10 items:

ITEM NUMBER	CORRELATION
1	.89
2	.75
3	.27
4	.81
5	.19
6	.53
7	.58
8	.72
9	.63
10	.60

Based on these data, the researcher should decide to eliminate Items 3 and 5 which fall below .50, and to place the other eight items in the final scale, confident that the remaining items measure something in common.

Item analysis of questions intended to measure the same variable in the same way is one important use of the data collected from a pilot test. However, item analyses are not as critical for refining questionnaires as they are for refining tests. Responses to questionnaire items are usually reviewed by eye for clarity and distribution without necessarily running an item analysis.

A pilot test can uncover a variety of failings in a questionnaire. For example, if all respondents reply identically to any one item, that item probably lacks discrimination. If you receive a preponderance of inappropriate responses to an item, examine it for ambiguity or otherwise poor wording. Poor instructions and other administration problems become apparent on a pilot test, as do areas of extreme sensitivity. If respondents refuse to answer certain items, try to desensitize them by rewording. Thus, pilot tests enable researchers to debug their questionnaires by diagnosing and correcting these failings.

SAMPLING PROCEDURES

Random Sampling

A researcher administers a questionnaire or interview to gain information about a particular group of respondents, such as high school graduates, school administrators in New England, or home economics teachers in New Jersey. This target group is the study's *population,* and the first step in sampling is to define the population. The researcher then selects a sample or representative group from this population to serve as respondents. As one way to ensure that this sample is representative of the larger population, a researcher might draw a random sample, because random selection limits the probability of choosing a

biased sample.' For example, you are interested in obtaining information about presidents of 2-year colleges. The population is 2,800 presidents, from which you want a sample of 300. Which 300 should you choose? To draw a random sample, you might write the names of all the 2-year colleges in alphabetical order giving each a number in the sequence.[4] You could then select 300 numbers by matching schools' assigned numbers against a table of random numbers like the one in Appendix C. The resulting list of 300 colleges, each with one president, is a random sample of the population from which it was drawn. Systematic biases in selection or selectees can be minimized by this procedure.[5] However, when certain sample variables are of special interest to the researcher (for example, age) stratified sampling should be employed, defining variables of interest as *sampling parameters. (See* the section on stratified random sampling later in the chapter.)

Defining the Population

The *population (or target* group) for a questionnaire or interview study is the group about which the researcher wants to gain information and draw conclusions. A researcher interested in the educational aspirations of teachers, for example, would focus on teachers as the population of the study. The term defining *the population* refers to a process of establishing boundary conditions that specify who shall be included in or excluded from the population. In the example study, the population could be defined as elementary school teachers, or public school teachers, or all teachers, or some other choice.

Specifying the group that will constitute a study's population is an early step in the sampling process, and it affects the nature of the conclusions that may be drawn from a study. A broadly defined population (like "all teachers") maximizes external validity or generality, although such a broad definition may create difficulties in obtaining a representative sample, and it may require a large sample size. Conversely, defining the population narrowly (for example, as "female, elementary school teachers") may facilitate the selection of a suitable sample, but it will restrict conclusions and generalizations to the specific population used, which may be inconsistent with the intent of the study.

---

[3] The process for random assignment of subjects to groups was described in Chapter 7. The description here concerns the random selection of subjects from a population. Although random assignment is a strategy for controlling threats to internal validity, random selection is a strategy for controlling threats to external validity. A single study may combine both procedures to define experimental and control groups. Where groups are to be obtained only by sampling and not by assignment, random sampling or some variant would he used alone.

[4] Actually, you can write the names in any order. Alphabetizing is merely a convenience. Indeed, this example study might refer to a directory that lists the colleges in alphabetical order.

[5] A statistical test allows a researcher to determine how closely sample characteristics approximate population characteristics and thus the extent to which the sample is representative of the population from which it was drawn (as would be expected of a random sample). For a discussion of these statistics, see Ferguson (1981).

The most reasonable criteria for defining a study's target population reflect the independent, moderator, and control variables specified in the study design along with practical considerations such as availability of subjects or respondents. When a control variable in a study deals with a population characteristic, the researcher must systematically include or exclude individuals with this characteristic in defining the population. (Chapter 7 discusses priorities for limiting a study's population.] For example, a researcher might want to make a comparison between academic high school graduates and vocational high school graduates, but only for students who had attended one or the other school for 3 consecutive years; the population for this study would be defined to exclude all graduates who had switched from one school to the other. In studying school superintendents in urban and rural settings, a researcher might define the population to exclude all superintendents who have not yet completed their second year in their districts, thus controlling for longevity, a potentially important control variable. (Longevity might also be controlled through stratified sampling, as discussed later in the chapter.)

In addition to design considerations, practical considerations affect the definition of the population. In Slavin and Karweit's (1984) study of mastery learning, for example, the availability of inner-city children might have influenced the researchers to define their population as "urban" children. However, because of the variables of interest, both urban children and suburban children might have been included in the definition of the population, had both been available, allowing the researchers to evaluate residence as a moderator variable. In general, independent and moderator variables require that a population include individuals with certain characteristics, whereas control variables require exclusion of particular groups.

Thus, an early step in sampling is to define the population from which to draw the sample. By referring to the variables of interest and by taking into account practical considerations, the researcher chooses characteristics to be included in and excluded from the target population. A concrete sampling plan illustrating the process of exclusion in defining the population appears in Figure 10.7.

Establishing Specifications for Stratified Random Sampling

Techniques of *stratified* random sampling permit researchers to include parameters of special interest and to control for internal validity related to selection factors through applications of moderator or control variables. In addition, stratification represents a good operational strategy for screening members of the population into and out of the study and for reducing the variability of the sample.

The first step in stratified sampling is to identify the stratification parameters, or variables. Each stratification parameter represents a control variable, that is, a

A **SAMPLING** PLAN **FOR** **SAMPLING** 2-YEAR COLLEGES FIGURE 10.7

Population: All 2-year colleges in the U.S.A.

Variables controlled by exclusion:

(1) College must have graduated a minimum of one class

(2) President must have held position for a minimum of one year

Variables controlled by stratification:

(1) Private-Public

25% private 75% public

(2) Urban-Rural

15% urban 10% rural 60% urban 15% rural

(3) Size of Student Body*

5% large	1% large	48% large	3% large
10% small	9% small	12% small	12% small

If the sample size were to be 300, it would be broken down as follows:

		Sample	***Population***
private, urban, large	***5%***	15	140
private, urban, small	10%	30	280
private, rural, large	1%	3	28
private, rural, small	9%	27	252
public, urban, large	48%	144	1,344
public, urban, small	12%	36	336
public, rural, large	3%	9	84
public, rural, small	12%	36	336
	100%	300	2,800

*Large = more than 2,000 students; small = fewer than 2.000 students.

potential source of error or extraneous influence that may provide an alternative explanation for a study's outcome. Assume that you want to contrast the teaching techniques of male and female elementary school teachers. The study would restrict the population to elementary school teachers, because that is a specified control variable, and it would sample across male and female teachers, because gender is the independent variable. You are concerned, however, that teaching experience may be an extraneous influence on your results. To offset this potential source of error, first you would determine the distribution of years of experience for male and for female elementary school teachers; then you would select the sample in proportion to these distributions. (The ***selection*** of specific subjects

within each stratum or proportion would be done randomly.) The other control variables would be treated in a similar way.

Consider sampling procedures for national political polls. Results are usually reported separately for different age groups and for different sections of the country. The studies treat age and geography as moderator variables and define separate samples according to them. However, within each age and geographical group, such a study may control for gender, race, religion, socioeconomic status, and specific location by proportional stratification. If half of the young people in the northeastern United States are male, then males should constitute half of the sample of northeastern young people. If 65 percent of the southeastern middle-aged group is poor, then poor people should make up 65 percent of the sample of this group. (Of course, terms like *middle-aged* and poor must be operationally defined.) The pollsters then consider these subpopulation differences in evaluating the outcomes of their studies.

Consider the example on sampling 300 presidents of 2-year colleges. Some bias may still affect results in spite of this random selection due to overrepresentation of private colleges. To control for this factor, use it as a variable or parameter for stratified sampling. Suppose one-quarter of the Z-year colleges are private schools and three-quarters are public institutions. In proportional stratified sampling, you would embody these percentages in your sample. In a sample of 300 college presidents, you would want 75 from private, 2-year colleges and 225 from public ones (the specific individuals in each stratum being randomly chosen). These specifications ensure creation of a sample systematically representative of the population.

To accomplish this stratified sampling method, you would make two separate alphabetical lists, one of private colleges, the other of public schools. You would then use your table of random numbers to select 75 private and 225 public colleges from the two lists, respectively. Of course, you could go further and control also for factors such as urban versus rural setting or large versus small colleges. However, in considering stratification, remember that each additional control variable complicates the sampling procedure and reduces the population per category from which each part of the sample is drawn. The sampling plan for this study is shown in Figure 10.7.

Random choice is the key to overcoming selection bias in sampling; *stratification* adds precision in ensuring that the sample contains the same proportional distribution of respondents on selected parameters as the population. Where stratified sampling is used, *within each stratum, researchers must choose sample respondents by random methods to* increase the likelihood of eliminating sources of invalidity due to selection other than those controlled through stratification. The combination of stratification and random selection increases the likelihood that the sample will be representative of the population. Because it controls for selection invalidity based on preselected variables in a systematic way, stratification is recommended for use with the variables identified as representing the greatest potential sources of selection bias. For information about determining sample size, see Chapter 11.

PROCEDURES FOR ADMINISTERING A QUESTIONNAIRE

This section focuses on procedures for mailing out a questionnaire, following it up, and sampling from among those in the sample who do not respond (hereafter called *nonrespondents).*

Initial Mailing

The initial mailing of a questionnaire to a sample of respondents typically includes a cover letter, the questionnaire itself, and a stamped, return-addressed envelope.

The cover letter is a critical part of the initial mailing, because it must establish the legitimacy of the study and the respectability of the researcher. The cover letter should briefly make its case for participation, focusing on the following points:

1. *The purpose of the study.* To satisfy the intellectual curiosity of potential respondents and to allay any doubts that participation will threaten their privacy or reputations, the researcher should disclose the ultimate uses of the data. Therefore, the cover letter should indicate the purposes and intentions of the study. It is often impossible, however, to give respondents complete details about the purposes of the study, because such knowledge might bias their responses.
2 *The protection afforded the respondent.* Respondents are entitled to know how a researcher will treat their privacy and confidentiality; thus the letter should indicate whether respondents must identify themselves and, if so, how their identities and responses will be protected. If questionnaires will be destroyed after rostering, and if rostering will be done by number rather than name (both recommended practices), the cover letter should include this information.
3 *Endorsements of the study.* Because respondents will feel secure about participating if they know that recognized institutions are behind the study, the cover letter should appear on university or agency letterhead. If a study will evaluate respondents as part of a professional group, then the cooperation and endorsement of this group should be obtained and mentioned in the letter. If the study is undertaken as a doctoral dissertation, mention the dissertation advisor by name and/or ask the dean of the school to sign or countersign the letter. If any agency or organization is providing financial support for the study, then this connection should be acknowledged.
4 *Legitimacy of the researcher. Say* who and what you are. Identify yourself by both name and position.

5. *Opportunities for debriefing.* If respondents can obtain the results of the study or additional explanations of its purpose at some later date, tell them so.
6. *Request for cooperation.* The letter constitutes an appeal from you for the respondent's help. If you have identified any special reasons why they should help (for example, the importance of the study for their profession) be sure to mention them.
7. Special *instructions.* The questionnaire should be self-administering and self-contained, although general instructions may be contained in the cover letter. Be sure to set a deadline for returning completed instruments, and caution against omitting answers to any items.

These seven points are important considerations in any research administration conducted by mail or in person. A personal interview should begin, in effect, with an oral cover letter. Figure 10.8 is an example of a cover letter.

The initial mailing may include more than one cover letter. For example, a letter of endorsement from a funding agency or from an organization to which the respondent belongs may help to gain the cooperation of prospective participants. A wise researcher does not print each respondent's name on his or her copy of the questionnaire to avoid causing alarm about the confidentiality of the study. Assignment of code numbers is a much better method of identification. Because filling out a questionnaire is, at the very least, an imposition on a respondent's time, both it and the cover letter should be as brief as possible.

Follow-Ups

After a period of about 2 weeks to a month has elapsed, a researcher should correspond with recipients who have not yet returned their questionnaires (that is, nonrespondents). This second mailing can consist simply of another letter soliciting cooperation. It should also include another questionnaire and another stamped, return-addressed envelope in case the respondent cannot find the original ones.

Ordinarily, about one-third to two-thirds of the questionnaires sent out will be returned during the month after the initial mailing. Beyond this period, about 10 to 25 percent can be stimulated to respond by additional urging. If the second mailing (the first follow-up letter) fails to stimulate a response, some researchers send a third mailing. This second follow-up typically takes the form of a postcard and follows the second mailing by about 2 to 3 weeks. Most researchers are unwilling to accept a return of less than 75 to 90 percent (and rightly so). Additional mailings, telephone calls, and a large sampling of nonrespondents (as discussed later) often help to elevate the return. Telegrams or telephone calls may be helpful in generating responses. If a study is worth doing, it is worth striving for the greatest return possible. An example of a follow-up letter is shown in Figure 10.9.

SAMPLE COVER LETTER

FIGURE 10.8

The Florida State University
Tallahassee, Florida 32306-3030
College of Education
Department of Educational Research, B-197

February 18, 1992

Dear Supervisor:

I am currently a doctoral candidate in the College of Education at Florida State University and am working with Dr. Bruce W. Tuckman on a project dealing with teacher feedback. The purpose of my study is to *assist* student teachers in developing their reaching skills through feedback from students.

I am asking you to have each of the student teachers you supervise administer the enclosed test instrument two, times to one of the classes he or she teaches during the eight weeks of student teaching. The first time will be after two weeks of teaching, and the second time after the eighth week. I will provide each student teacher with all the necessary materials needed (pencils, questionnaires).

It should not take students more than five minutes to complete this form, which can be done after the student teacher has completed teaching the lesson. It is a form that is designed to describe the student teacher's style as seen by students.

The student teacher's name, the names of the students who respond, and the name of your school district will not be identified in my study. Instead, I will use a numerical coding process to label and identify my data. To assure the privacy of all involved, under no circumstances will I reveal the identity of the participants to either the school administration or the public,

I deeply appreciate your cooperation and support. Without you and the cooperation of your student teachers, I would not be able to conduct this research project, which hopefully will **shed** light on the improvement of teaching skills in preservice teachers. When the study is completed, I will provide you with a description of the results.

If you have any further questions, feel free to call me at 644-4592.

Sincerely,

Jane R. Richardson

P.S. Enclosed please find a copy of the feedback form

FIGURE 10.9 SAMPLE FOLLOW-UP LETTER

From Forsyth, 1976. Reprinted by permission of the author.

GRADUATE SCHOOL OF EDUCATION • 10 SEMINARY PLACE • NEW BRUNSWICK • NEW JERSEY 08903

5/11/

School of Education
Old Ivy University
Hometown, U.S.A.

Dear Professor _____

The year's end approaches and soon many faculty will scatter for the summer in search of rest and/or individual pursuits. I, however, may sit "dataless" with perplexed and furrowed brow—contemplating for my dissertation Hamlet's immortal question. I need your assistance. The questionnaire I sent you a while back may be buried somewhere so here is another copy for your convenience.

If you were hesitant about completing a questionnaire asking you to list names of individuals, even though anonymously, let us address that issue briefly. The names as such are not significant to this researcher. However, they allow me to reconstruct interaction patterns in this and other academic organizations. The names, once received, are translated into symbols and the original data are destroyed.

A high response rate is essential to the success of this study. Your participation can make the difference. With it many long hours of work will all have been well spent. Again, let me assure you of the confidentiality of your response. The data will not be reported in any way which would allow the most astute student of interaction patterns at the School of Education to identify any individual or group. When you complete the instrument, send separately the signed postcard that will halt follow-up while protecting your anonymity.

If for some reason you still do not wish to participate, it would be greatly appreciated if you would indicate that decision on the return postcard and return it as addressed. In any case, thank you for your time and consideration. One last option: if you are reluctant to indicate names, I can conduct an interview with you which would allow you to do all the translating into numerical data. I can be reached at 201-932-7531.

Sincerely,

Patrick B. Forsyth

Sampling Nonrespondents

If fewer than about 80 percent of people who receive the questionnaire complete and return it, the researcher must try to reach a portion of the nonrespondents and obtain some data from them. Additional returns of all or critical portions of the questionnaire by 5 or 10 percent of the original nonrespondents is required for this purpose.

This additional procedure is necessary to establish that those who have not responded are not systematically different from those who have. Failure to check for potential bias based on nonresponse may introduce both external and internal invalidity based on experimental mortality (selective, nonrandom loss of subjects from a random sample) as well as a potential increase in sampling error.

Obtaining data from nonrespondents is not easy, since they have already ignored two or three attempts to include them in the study. The first step is to select at random 5 to 10 percent of these people from your list of nonrespondents, using the table of random numbers (Appendix C). Using their code numbers, go through the table of random numbers and pick those whose numbers appear first, then write or call them. About a 75-to-80-percent return from the nonrespondents' sample may be all that can be reasonably expected, but every effort should be made to achieve this goal."

CONDUCTING AN INTERVIEW STUDY

Procedures for conducting an interview may differ from those involved in obtaining data by questionnaire, but the aim is the same: to obtain the desired data with maximum efficiency and minimum bias.

Selecting and Training Interviewers

Researchers would obviously prefer to select previously trained and experienced interviewers, but this is an elusive goal. Consequently, many studies employ graduate and undergraduate students. The necessary level of skill will depend on the nature of information you are trying to elicit: personal, sensitive material will require skilled interviewers.

The task of an interviewer is a responsible one, both in the manner of conducting an interview and in willingness to follow instructions. In training, a potential interviewer should observe interviews proceeding in the prescribed manner and then should have the opportunity to conduct practice interviews under observation. Some practice interviews should involve "live" respondents, that is, potential subjects from the study's sample. Practice sessions should also include

[d] One suggested alternative approach calls for comparing the responses of early returners and late returners to check for a response rate bias.

interviews that the researcher has arranged to present certain typical situations that the interviewer will encounter. These "rigged" interviews present trainees with a range of possible situations.

Training should also familiarize prospective interviewers with the forms that will be used in recording responses and keeping records of interviews. To control the sampling, the trainees must learn to determine whom they should interview. They also must know how to set up interview appointments, how to introduce themselves, how to begin interviews in a manner that will put the interviewees at ease, how to use response-keyed questions and other nonlinear approaches, how to record responses, and (if the job includes this task) how to code them.

All interviewers should receive similar training experiences where possible, for differences in interviewer style and approach represent a source of internal invalidity due to instrumentation bias. Interviewers are instruments for collecting data, and, as instruments, their own characteristics should affect the data as little as possible: Interviewers should reflect their respondents and not themselves. Of course, it is impossible to make perfect mirrors out of human interviewers, but if they are chosen from the same population and receive the same training and instructions, they should tend to become standardized against one another as a function of their training. Research may also benefit if the trainers divulge no more about the study to the interviewers than is absolutely necessary; training should not subtly make them confederates who may unconsciously bias the outcomes in the expected directions.

Conducting an Interview

The first task of an interviewer may be to select respondents, although some studies give interviewers lists of people to contact. Unless the interviewers are both highly trained and experienced, study directors should give them the names, addresses, and phone numbers of the people to be interviewed, along with a deadline for completion. The interviewer may then choose the interviewing order, or the researcher may recommend an order.

Typically, an interviewer proceeds by telephoning a potential respondent and, essentially, presenting a verbal cover letter. However, a phone conversation gives the interviewer the advantage of opportunities to alter or expand upon instructions and background information in reaction to specific concerns raised by potential respondents. During this first conversation, an interview appointment should also be made.

At the scheduled meeting, the interviewer should once again brief the respondent about the nature or purpose of the interview (being as candid as possible without biasing responses) and attempt to make the respondent feel at ease. This session should begin with an explanation of the manner of recording responses; if the interviewer will tape record the session, the respondent's assent should be obtained. At all times, interviewers must remember that they are data-collection instruments who must try to prevent their own biases, opinions, or

curiosity from affecting their behavior. Interviewers must not deviate from their formats and interview schedules, although many schedules will permit some flexibility in choice of questions. The respondents should be kept from rambling, but not at the sacrifice of courtesy. (See Chapter 14 for additional information.)

CODING AND SCORING

Objectively Scored Items

Many questions, such as those presented in the form of rating scales or checklists, are precoded; that is, each response can be immediately and directly converted into an objective score. The researcher simply has to assign a score to each point on the list or scale. However, data obtained from interviews and questionnaires (often called *protocols)* may not contribute to the research in the exact form in which they are collected. Often further processing must convert them to different forms for analysis. This initial processing of information is called scoring or *coding.*

Consider Item 13 from the Career Awareness Scale, the sample questionnaire that appears in Figure 10.2:

13. I work at different kinds of part-time jobs. 1. A 2. O 3. S 4. N

You might assign *never* (N) a score of 1, ***seldom (S) a*** score of 2, *often* (0) a score of 3, and always (A) a score of 4. You could then add the scores on all the items to obtain a total score on the scale.[1]

Sometimes items are written in both positive and negative directions to avoid response bias. Consider the following two items on a questionnaire measuring attitudes toward school:

- I enjoy myself most of the time in school.
  strongly agree agree disagree strongly disagree
- When I am in school I usually feel unhappy.
  strongly agree agree disagree strongly disagree

If you were to score strongly agree for the first item as 4, then you would have to score the strongly agree response for the second item as 1, because strong agreement with the first item indicates that a respondent likes school whereas strong agreement with the second item indicates a dislike for school. To produce scores on these two items that you can sum to get a measure of how much a student likes school, you have to score them in opposite directions.

[1] Note that the numbers listed on the scale for the response choices have been reversed for scoring so that a higher score reflects more career awareness behavior.

Often a questionnaire or overall scale contains a number of subscales, each of which measures a different aspect of what the total scale measures. In analyzing subscale scores, a scoring key provides extremely helpful guidance. Typically, such a scoring key is a cardboard sheet or overlay with holes punched so that when it is placed over an answer sheet, it reveals only the responses to the items on a single subscale. One scoring key would be required for each subscale. Using answer sheets that can be read by optical scanners and scored by computers makes this process much easier.

Thus, in scoring objective items, such as rating scales and checklists, the first step is identification of the direction of items-separating reversed and nonreversed ones. The second step is assigning a numerical score to each point on the scale or list. Finally, subscale items should be grouped and scored.

By their very nature, ranking items carry associated scores, that is, the ranks for each item in the list. To determine the average across respondents for any particular item in the list, you can sum the ranks and divide by the number of respondents. All ranking items can be scored in this way. This set of averages can then be compared to that obtained from another group of respondents using the Spearman rank-order correlation procedure (described in the next chapter).

Some scales, such as those using the true-false and yes-no formats, lend themselves primarily to counting as a scoring procedure. Simply count the number of "true" or "yes" responses. However, you must still pay attention to reversed items. A "false" answer on a reversed item must be counted along with a "true" response on a nonreversed item. On a positively phrased item, for example, a "yes" would get a score of 1, and a "no" would get a score of 0. In contrast, on a negatively phrased item, a "yes" would get a score of 0, and a "no" would get a score of 1.

In another scoring procedure, a researcher can count people who fit into a particular category. For instance, if a questionnaire asks respondents to identify their gender, a scorer counts the number who indicate "male" and the number who indicate "female."

Generally speaking then, four scoring procedures apply to objective items:

1. Scale *scoring*. Where the item represents a scale, each point on the scale is assigned a score. After adjusting for reversal in phrasing, you can add a respondent's scores on the items within a total scale (or subscale) to get his or her overall score.
2. *Rank scoring.* A respondent assigns a rank to each item in a list. Here, typically, average ranks across all respondents are calculated for each item in the list.
3. *Response counting.* Where categorical or nominal responses are obtained on a scale (such as true-false), a scorer simply counts the number of agreeing responses by a respondent. This count becomes the total score on the scale for that respondent. Response counting works for a scale made up of more than one item, all presumably measuring the same thing.

4. *Respondent counting.* Where a questionnaire elicits categorical or nominal responses on single items, scoring can count the number of respondents who give a particular response to that item. By properly setting up the answer sheet in advance, mechanical procedures can complete respondent counts. Respondent counting enables a researcher to generate a contingency table (a four-cell table that displays the number of respondents simultaneously marking each of the two possible choices on two items) and to employ chi-square analysis (described in the next chapter). A contingency table is illustrated in Figure 10.10.

Fill-In and Free-Response Items

Although a scorer can apply any one of the four techniques described above to process fill-in and free-response items, the most common is respondent counting. However, before counting respondents, he or she must code their responses. Coding is a procedure for reducing data to a form that allows tabulation of response similarities and differences.

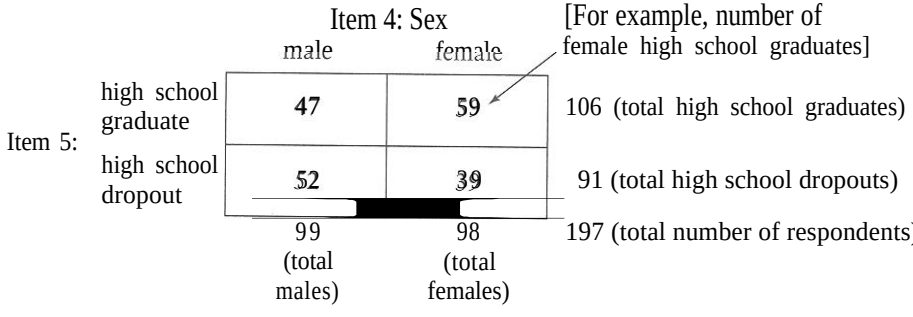
Suppose, for example, that an interviewer asks: Why did you leave school? Suppose, also, that the following potential responses to this question have been identified by the researcher:

___ Couldn't stand it (or some other indication of strong dislike)
_ Wasn't doing well
_ Waste of time
_ Better opportunities elsewhere
_ Other: _____

To maintain efficiency, researchers often establish such precoded response categories for fill-in and free-response items. Although respondents never see these

AN EXAMPLE OF A CONTINGENCY TABLE

FIGURE 10.10



responses (if they did, the item would be a checklist), they appear on the interviewer's answer form; while the respondent is talking, she or he judges which one gives the best fit. Thus, these precoded response categories become a nominal checklist enabling the interviewer to code immediately the unstructured response into checklist form. As an alternative, the interviewer might indicate which of those reasons a respondent gave and rank their apparent importance to the respondent. Coding, therefore, represents a superimposition of a response format onto a free or unstructured response.

Often coding occurs before data collection by supplying interviewers with precoded interview schedules. While they ask open-ended questions and respondents give free responses, the interviewers attempt to catalog the responses into one or more category sets. Here are two examples to illustrate this point:

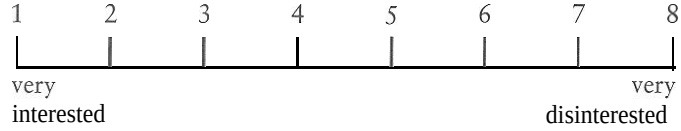
- *Question:* Whom do you consult when you have a problem in school?
  Answer: Mainly I go to my friends, especially my best buddy. Sometimes I talk to my rabbi.
  Coding: _ Parents X Friends
  _ Teacher(s) X Others: clergyman
  _ Counselor
- *Question:* What about school do you like least?
  Answer: I would say the work. I don't find my subjects interesting. They don't have anything to do with what I'm interested in.
  Coding: _ Teachers
  _ Organization
  X Schoolwork
  _ boring
  X irrelevant
  _ too easy
  _ too hard
  _ Other: _____

Of course, the coding scheme you employ in converting a response into analyzable data will be a function of the problem and the hypotheses with which you are working. Consider a hypothesis that youngsters in the upper third of the high school IQ distribution will be more likely to find their school work irrelevant than will youngsters in the middle or lower thirds. To test this hypothesis, you must find out how youngsters view their school work and then code their answers in terms of perceived relevance. The second example above represents an attempt to gather such information.

The extent to which precoding is possible is an indication of the extent to which the question is likely to yield relevant information. Precoding has the additional advantages of eliminating coding as a separate step in data reduction and providing the interviewer with an easy format for data collection.

Any attempt to design response-scoring codes must focus on the basic consideration of the information that you want to find out from the question. If you

are testing to see whether tenured teachers are more or less interested in teaching effectiveness than nontenured teachers, you might ask: How interested are you in the objective determination of your teaching effectiveness? The interviewer could be provided with a rating scale such as:



After listening to the teacher's free response to this question, the interviewer could summarize his or her opinion by placing a check on the rating scale. This method is an example of a *scale-scoring* approach to coding and scoring an open-ended response. An examination of the ratings indicated by the responses of the two groups of teachers would provide data to determine whether tenured or nontenured teachers are more interested in teaching effectiveness. Alternatively, the response could be precoded as simply: _____ seems interested, _____ seems disinterested. This application represents the respondent-counting approach: Simply count the number of teachers in each group (tenured and nontenured) who were seen as interested as well as those seen as disinterested, and place the findings into a contingency table:

	interested in teaching effectiveness	disinterested in teaching effectiveness
tenured teachers		
nontenured teachers		

The discussion of coding so far has focused on applications of precoded categories. The same kinds of coding procedures can work in coding after data collections. Precoding has the advantage of greater efficiency than postcoding, which requires interviewers to record free responses verbatim (usually by tape recorder) or summarize them as respondents speak. These recordings are then transcribed by a typist and finally coded. However, coding after data collection has the advantage of careful preservation of coder reliability.

The reliability of coding judgments becomes an important issue here, just as the previous chapter considered the reliability of rating and coding techniques

that describe behavior. If interviewers code every response, data analysts have no way to check the reliability of those coding decisions, because they lack any record of the responses. When interviewers do all the coding during the interviews rather than making verbatim records of responses, a researcher should be concerned about coding unreliability as a threat to instrumentation validity. To ensure this important priority, at least 20 percent of the responses should be recorded verbatim and then coded by at least two judges or interviewers, thus providing a sample of responses to assess intercoder reliability.

In postinterview coding, the response transcripts allow a second coder to code a sufficient number of protocols to establish reliability with the first coder, or for two coders to code all protocols to increase the reliability of the data.[8] Both first and second coders should be trained in the use of the coding system and complete practice trials under the scrutiny of the researcher. In such instances, reliabilities in the 0.70-to-0.90 range would be sufficient to prevent instrumentation bias in coding.

SUMMARY

1 Questionnaires and interviews provide self-reported data from respondents. Such data reflect what is inside a respondent's head, but they may be influenced by both self-awareness and the desire to create a favorable impression.

2 Questionnaire items represent five formats: (1) direct or obvious questions versus indirect or subtle questions; (2) specific or highly targeted questions versus nonspecific or relatively general questions; (3) fact versus opinion questions; (4) questions versus statements designed to stimulate agreement or disagreement; (5) predetermined questions versus response-keyed questions (those that depend on answers to previous questions).

3. Researchers employ seven response modes: (1) unstructured or open-ended responses; (2) fill-in responses; (3) tabular (table fill-in) responses; (4) scaled responses, in which respondents place themselves along a 5 (or more) point rating scale; (5) ranking responses, in which they rank order certain elements; (6) checklist responses, in which they check one or more selections that apply; (7) categorical responses, in which they check the one of two options that applies.

4. To construct a questionnaire or interview scale, a researcher completes the following five steps: (a) specifying the variables to be measured, or what she or he wants to find out; (b) choosing the question format(s) after considering the relative merits of each; (c) choosing the response modes depending on

[8] Where two coders code all protocols, their judgments can be averaged to obtain final scores. In such cases, the obtained reliability coefficients can be corrected by the Spearman-Brown formula (see Chapter 9), because the average is more reliable than either individual data set.

the type of data desired; (d) preparing either interview or questionnaire items; (e) pilot testing the instrument and evaluating the results using item analysis.

5. Sampling procedures begins with a definition of a study's population (setting its boundary characteristics). From this group, the researcher then draws a sample through simple random or stratified random sampling techniques, the latter requiring the establishment of sampling specifications. This careful process helps researchers to avoid subject selection bias that can affect external validity or generality.
6. Administration of a questionnaire requires (a) an initial mailing to a sample of respondents, accompanied by a cover letter to describe the study's purpose, protective measures for respondents, endorsements, legitimacy, debriefing, needed cooperation, and special instructions; (b) one or more follow-ups to those who do not respond; (c) a systematic attempt to get responses from 5 to 10 percent of the remaining nonrespondents (to evaluate the degree of potential mortality bias).
7. Conducting an interview study requires (a) selection and training of interviewers and (b) interviewing a sample of respondents.
8. Interview and questionnaire responses become usable data only after scoring or coding. For objectively scored items, scorers carry out four procedures: (1) scale scoring-totaling up scale points; (2) rank scoring-averaging ranks across respondents; (3) response counting-adding up the number of agreeing responses; (4) respondent counting—counting up the number of respondents who agree. Scoring for fill-in and free-response items requires a response-coding system that converts each response to quantitative data. The coded results may then be scored by any of the four procedures, most typically by respondent counting.

COMPETENCY TEST EXERCISES

1. Which of the following statements does not describe a purpose for which researchers use interviews and questionnaires!
   a. Finding out what a person thinks and believes
   b. Finding out what a person likes and dislikes
   c. Finding out how a person behaves
   d. Finding out what experiences a person has had

2. Which of the following limitations is not a shortcoming of a questionnaire or interview?
   a. The respondent may not know anything about the interviewer.
   b. The respondent may not know the information requested.
   c. The respondent may try to show himself or herself in a good light.
   d. The respondent may try to help by telling you what you expect to hear.

3. Match up the question types with the descriptions.

a. Indirect question	1. Declarative sentence form
b. Specific question	2. Requests reaction to a single object
c. Question of opinion	3. Next question depends on the response to this one
d. Statement	4. Requests information for inferences
e. Response-keyed question	5. Asks how the respondent feels about something

4. Match up the response types with the examples.

a. Scaled response	1. My favorite subject is (check one): English Chemistry Calculus
b. Fill-in response	2. My favorite subject is calculus. (yes, no).
c. Ranking response	3. How do you feel about chemistry?
d. Tabular response	4. English is a subject I (like a lot, like a little, dislike a little, dislike a lot).
e. Checklist response	5. My favorite subject is
f. Unstructured response	6.
g. Categorical response	7. My order of preference of subjects is:

6.

	English	Chem	Calc
Like			
Dislike			

7. My order of preference of subjects is:
English _____ (1,2,3)
Chemistry _____ (1,2,3)
Calculus _____ (1,2,3)

5. In the list of considerations, write I next to those suited to an interview and Q next to those suited to a questionnaire.
   a. I want to collect data from at least 90 percent of my sample._
   b. I want to keep my problems of data reduction to a minimum._
   c. I do not have very much money to conduct this project._
   d. I want to collect highly reliable data in this study._
   e. I'm not sure what questions respondents will likely answer._
   f. I have to ask some intensive questions, which may lead into sensitive areas._

6. In the list of considerations, write *F next to* those that support or describe the use of the fill-in response mode, S for those that support scaled response, *R* for ranking responses, and C for checklist or categorical responses.
   a. I do not have to anticipate potential responses._
   b. I want to gather ordinal data._
   c. This response mode does not provide for degrees of agreement, so it allows too few options._

d. I'll have a big scoring job. (Prescoring will be a difficult task.)_
e. I may get response bias away from the extremes._

7. You are interested in finding out about the attitudes of teachers toward their school administration, particularly with regard to procedures for ordering classroom supplies. Construct three sequential interview questions.

8. You are interested in finding out about the attitudes of administrators toward teachers, particularly as regards their application of procedures for ordering classroom supplies. Construct three questionnaire items (using three different structured-response modes other than fill-in) to accomplish this goal.

9 You are planning to draw a stratified random sample of 200 from a high school population that contains 60 percent males and 40 percent females. Among the males, 40 percent are college prep majors, 10 percent business majors, 20 percent vocational majors, and 30 percent general majors. Among the females, 50 percent are college prep majors, 25 percent business majors, 5 percent vocational majors, and 20 percent general majors. How many respondents would you need in each of the eight categories?

10. You are going to interview 60 teachers in a school system of 200 teachers. The total includes 100 elementary school teachers-20 men and 80 women; 50 junior high school teachers-20 men and 30 women; and 50 high school teachers-30 men and 20 women. How many teachers in each of the six categories would you include in your sample of 60?

11. Which of the following subjects is not ordinarily discussed in a cover letter?
    a. Protection afforded the respondent
    b. Anticipated outcome of the study
    c. Legitimacy of the researcher
    d. Purpose of the study

12. You are planning to do a study of the relationship between a teacher's length of teaching experience and his or her attitudes toward discipline of students. You are sending out a questionnaire including an attitude scale and a biographical information sheet. Construct a sample cover letter to accompany this mailing.

RECOMMENDED REFERENCES

Berdie, D. R., Anderson, J. F., & Niebuhr, M. A. (1986). Questionnaires: *Design and use* (2nd ed.). Metuchen, NJ: Scarecrow Press.

Fowler, F. J. (1993). Survey research methods (2nd ed.). Beverly Hills, CA: Sage.

Lavrakas, l? J. (1987). *Telephone survey methods: Sampling, selection, and supervision.* Newbury Park, CA: Sage.