

June 1991

Using Structured Interviewing Techniques

Preface

GAO assists congressional decisionmakers in their deliberative process by furnishing analytical information on issues and options under consideration. Many diverse methodologies are needed to develop sound and timely answers to the questions that are posed by the Congress. To provide GAO evaluators with basic information about the more commonly used methodologies, GAO's policy guidance includes documents such as methodology transfer papers and technical guidelines.

This methodology transfer paper on using structured interviewing techniques discusses how GAO evaluators should incorporate structured interview techniques when appropriate to performing our work. It explains when these techniques should be used and what steps should be followed. Overall, it describes techniques for designing a structured interview, for pretesting, for training interviewers, and for conducting the interviews. The original report was authored by Erwin W. Bedarf in July 1985. This reissued version, prepared by Kenneth Litkowski, supersedes the version published in 1985.

Using Structured Interviewing Techniques is one of a series of papers issued by the Program Evaluation and Methodology Division (PEMD). The purpose of the series is to provide GAO evaluators with guides to various aspects of audit and evaluation methodology, to illustrate applications, and to indicate where more detailed information is available.

We look forward to receiving comments from the readers of this paper. They should be addressed to Eleanor Chelimsky at 202-275-1854.



Werner Grosshans
Assistant Comptroller General
Office of Policy



Eleanor Chelimsky
Assistant Comptroller General
for Program Evaluation and
Methodology

Contents

Preface		1
Chapter 1		8
The Role of Structured Interviews in GAO Evaluations		
Chapter 2		11
What Is a Structured Interview and When Should It Be Used?		
Chapter 3		18
Designing a Structured Interview	Identifying Variables and Developing Questions	18
	Composing Appropriate Questions	22
	Selecting a Question Format	23
	Organizing Questions	32
	Layout Considerations	34
Chapter 4		39
More on Interview Design: Avoiding Problems	Appropriateness of the Language	39
	Level of the Language	40
	Use of Qualifying Language	41
	Clarity of Language	42
	Bias Within Questions	44
	Considerations for Telephone Interviewing Instruments	48

Contents

Chapter 5		49
Pretesting and Expert Review		49
	Purpose of Pretest	49
	Pretest Procedures	50
	Purpose of Expert Review	55
	Instrument Redesign	56
<hr/>		
Chapter 6		57
Training Interviewers		57
	Training Methods	57
	Interviewer Qualifications	62
<hr/>		
Chapter 7		64
Selecting and Contacting Interviewees		64
	Selection of Interviewees	64
	Contacting Potential Interviewees	64
	Interview Arrangements	70
	Protecting the Interviewee	72
<hr/>		
Chapter 8		74
Conducting Interviews		74
	Developing Rapport and Showing Interest	74
	Giving the Interviewee a Reason to Participate	76
	Helping the Interviewee to Be Responsive	76
	Asking Questions in a Prescribed Order and Manner	77
	Ensuring Understanding	78
	Ensuring Nonbias	79
	Obtaining Sufficient Answers	79
	Showing Sensitivity to Interviewee Burden	80
<hr/>		
Chapter 9		81
Analyzing the Data		81
	Nonrespondent Problem	82
	Data Analysis	85
	Analysis of Open-Ended Questions	88

Chapter 10 The Role of Evaluators and Specialists on Each Task		90
Bibliography		94
Glossary		96
Papers in This Series		99
Tables		
	Table 1.1: Evaluation Questions and Strategies	8
	Table 2.1: Comparison of Data-Collection Techniques	12
	Table 3.1: Identifying, Developing, and Selecting Questions	21
	Table 10.1: Functions and Responsibilities of Evaluators and Specialists	91
Figures		
	Figure 3.1: Question 1	25
	Figure 3.2: Questions 2, 3, and 4	26
	Figure 3.3: Question 5	26
	Figure 3.4: Question 6	27
	Figure 3.5: Question 7	27
	Figure 3.6: Question 8	28
	Figure 3.7: Question 9	28
	Figure 3.8: Question 10	32
	Figure 3.9: Question 11	32
	Figure 3.10: Structured Interview Text	35
	Figure 4.1: Question 1	44
	Figure 4.2: Question 2	46
	Figure 4.3: Question 3	46
	Figure 4.4: Question 4	47
	Figure 4.5: Question 5	47

Contents

Figure 7.1: Interviewee Contact Procedures Example	65
Figure 7.2: Example of Telephone Contact With Potential Interviewee	67
Figure 7.3: Interviewee Contact Log	68
Figure 7.4: Interviewee Contact Log Filled In	69

Abbreviations

CATI	Computer-assisted telephone interviewing
DCI	Data-collection instrument
GAO	General Accounting Office
PEMD	Program Evaluation and Methodology Division
QPL	Questionnaire Programming Language
QUEST	A GAO computer program for formatting interviews and questionnaires

The Role of Structured Interviews in GAO Evaluations

A major responsibility of the General Accounting Office (GAO) is to audit and evaluate the programs, activities, and financial operations of federal departments and agencies and to make recommendations toward more efficient and effective operations.

The broad questions that dictate the objectives of a GAO evaluation and that suggest the evaluation strategy can be categorized as descriptive, normative, or impact (cause-and-effect).¹ A descriptive evaluation, as the name implies, provides descriptive information about specific conditions of a program or activity, while a normative evaluation compares an observed outcome to an expected level of performance. An impact (cause-and-effect) evaluation aims to determine whether observed conditions, events, or outcomes can be attributed to the operation of the program or activity. According to the type of evaluation questions to be answered, different evaluation strategies are used, as shown in table 1.1.

Table 1.1: Evaluation Questions and Strategies

Type of question	Strategy
Descriptive	Sample survey
	Case study
	Available data
Normative	Sample survey
	Case study
	Available data
Impact (cause-and-effect)	Field experiment
	Available data

¹We use the term "evaluation" throughout this paper; however, many of the interviewing concepts and procedures apply equally to GAO audits. The categories of questions are discussed fully in the methodology transfer paper entitled Designing Evaluations. See the bibliography at the end of this paper.

Chapter 1
The Role of Structured
Interviews in
GAO Evaluations

In a sample survey, data are collected from a sample of a population to determine the incidence, distribution, and interrelationship of events and conditions. The case study is an analytic description of an event, process, institution, or program based on either a single case or multiple cases. The field experiment compares outcomes associated with program operations with estimates of what the outcomes would have been in the absence of the program. Available data refers to previous studies or data bases previously established and currently available.

The design of a GAO evaluation encompasses seven elements:

- the kind of information to be acquired,
- sources of information (for example, types of respondents),
- methods to be used for sampling sources (for example, random sampling),
- methods of collecting information (for example, structured interviews and self-administered questionnaires),
- the timing and frequency of information collection,
- the basis for comparing outcomes with and without a program (for cause-and-effect questions), and
- the analysis plan.

This paper focuses on the fourth design element—specifically, structured interviews. Like self-administered questionnaires, structured interviews are often used when the evaluation strategy calls for a sample survey. Structured interviews can also be used, however, in field experiments where information must be obtained from program participants or members of a comparison group. Similarly, when essentially the same information must be obtained from numerous people for a multiple case-study evaluation or a single case-study evaluation, it may be beneficial to use structured interviews.

Chapter 1
The Role of Structured
Interviews in
GAO Evaluations

Structured interviews (and other forms of structured data collection, such as the self-administered questionnaire) are often used in conjunction with a design that employs statistical sampling. This combination provides data that can be used to make projections about the entire population from which the sample was drawn. We discuss sampling methodology and generalization in depth in the methodology transfer paper entitled Using Statistical Sampling.

It should be noted, however, that the steps in the evaluation design process—defining the questions that dictate the objectives of the study, selecting the method of collecting the information, and preparing an analysis plan for using the collected information to answer the questions—are interrelated and iterative. If, for example, a structured interview is used to collect information to answer an evaluation question, the question will determine the contents or subject matter of the interview form. Any constraints in identifying and selecting a sample (for example, the lack of a universe listing of the target population) may make it necessary to refine the original evaluation question. Many more examples could be given to demonstrate the iterative nature of this process. The point to remember is that the use of structured interviewing to collect information is not an isolated process and cannot be thought of as a sequential task unrelated to or independent of other tasks in the process of answering an evaluation question.

What Is a Structured Interview and When Should It Be Used?

For years, GAO evaluators have collected data through various techniques such as reviewing records and interviewing government and contractor officials, employees, and program participants. Increasingly since 1972, we have used what have come to be called data-collection instruments (DCIs) on assignments that require the same or uniform information on numerous cases. A DCI is a document containing questions presented in a systematic, highly precise fashion; its purpose is to enable the evaluator to obtain uniform data that can be compared, summed, and, if it is quantitative, subjected to additional statistical analysis. The form of a DCI varies according to whether it is to be used in a structured interview, as a self-administered questionnaire (either mailed to individuals or organizations or completed by individuals in a group setting), or as a pro forma schedule to obtain information from records.

An interview that uses a DCI to gather data, either by telephone or face to face, is a structured interview, one in which evaluators ask the same questions of numerous individuals or individuals representing numerous organizations in a precise manner, offering each interviewee the same set of possible responses. In contrast, an unstructured interview contains many open-ended questions, which are not asked in a structured, precise manner. Different evaluators interpret questions and often offer different explanations when respondents ask for clarification.

Given the need to collect uniform data from numerous persons or organizations, when should the evaluator use a structured interview rather than a mail questionnaire or a questionnaire administered in a group setting? There is no hard-and-fast answer. We discuss some of the advantages and disadvantages of interviews and questionnaires in the

Chapter 2
What Is a Structured
Interview and When Should It
Be Used?

Table 2.1: Comparison of Data-Collection Techniques*

Characteristic or advantage	Extent of advantage				
	Structured interview		Questionnaire		
	By telephone	Face to face	By mail	Group	Audit of records
Methodology					
Allows use of probes	3	5	1	2	na
Controls bias of collector	3	2	5	4	5
Can overcome unexpected events in data collections	4	5	2	3	4
Facilitates feedback about instrument or collection procedures	4	5	2	5	2
Allows oral and visual inquiry	1	5	2	5	na
Allows oral and visual response	1	5	2	2	2
Evaluator can control collection procedures	3	5	1	4	5
Facilitates interchange with source	4	5	2	5	na
What contents allow					
Inclusion of most relevant variables	3	5	4	4	3
Complex subject matter to be presented or derived	3	5	3	4	4
Collection of real-time data	5	5	4	5	3
Acquisition of historical data	4	4	4	4	5
Universe or sample					
Relevant universe to be sampled can be identified	4	5	4	5	4
Facilitates contacting and getting sample	3	2	4	4	5
Allows use with large sample	4	3	5	4	5
Allows identity of source to be known	4	5	3	5	3

(continued)

Chapter 2
What Is a Structured
Interview and When Should It
Be Used?

Characteristic or advantage	Extent of advantage					Audit of records	
	Structured interview		Questionnaire		By mail		Group
	By telephone	Face to face					
Reduces problems from respondent's illiteracy	4	5	1	3		na	
What time, cost, and resources minimize							
Instrument-development time	2	3	1	1		5	
Instrument-development cost	3	1	1	1		5	
Number of field staff	5	?	5	?		?	
Travel by staff	5	?	5	?		?	
Staff training	2	1	5	3		5	
Time required to carry out activities	?	?	3	?		?	
Overall cost	3	1	5	4		1	
Results, response, and quality of data							
Maximize rate of return of data after source is contacted	4	5	3	5		na	
Minimize multiple contacts of sources	2	2	3	4		na	
Minimize follow-up after initial response	5	5	3	4		5	
Increase chance source will be accurate	4	4	4	4		3	
Allow reliability to be checked	5	5	3	4		4	
Allow validity to be checked	4	4	2	4		5	
Facilitate recall of data by source	4	5	3	4		na	

^aKey:

- 1 Little or no extent
- 2 Some extent
- 3 Moderate extent
- 4 Great extent
- 5 Very great extent
- ? Depends greatly upon study specification
- na Not applicable

Chapter 2
What Is a Structured
Interview and When Should It
Be Used?

following paragraphs. In addition, the characteristics of various data-collection techniques are systematically compared in table 2.1.

In the job design phase of an evaluation or in a one-of-a-kind interview during the data collection and analysis phase of an evaluation, the less-structured, less-guided type of interview may be more useful.

Face-to-face interviews and telephone interviews are generally more successful with respondents whose reading levels are low in comparison with the complexity of the questions. In this radio and television age, some respondent groups understand spoken words and sentences better than written ones.

The telephone interview and, even more, the face-to-face interview enable the interviewer to establish rapport with the respondents. Individuals who would ignore mail questionnaires entirely or who would not answer certain questions on them can be persuaded to provide truthful answers in a telephone or face-to-face interview. Also, a well-trained interviewer can recognize when a respondent is having a problem understanding or interpreting a question and can employ the proper techniques to assist the interviewee without jeopardizing the integrity of the interview.

In comparison to the telephone interview, the face-to-face interview gives the interviewer the opportunity to observe as well as listen. For example, if it is required or desired that the interviewee's living arrangements be noted, the face-to-face interview would be the choice. Also, more complex questions can be asked in a face-to-face interview than in a telephone interview. Respondents can be shown cards with the complete set of possible responses, making it easier for them to remember and consider all the choices. In addition, more questions can be asked. Twenty to 30 minutes is the usual limit for

Chapter 2
What Is a Structured
Interview and When Should It
Be Used?

telephone interviews, while face-to-face interviews can last up to an hour.

Computer-assisted telephone interviewing (CATI) is one form of telephone interviewing. In CATI, the questionnaire or DCI is stored in a computer, questions are displayed on the computer screen during the interview, and the interviewer directly enters the responses into the computer. Telephone interview costs generally fall somewhere between the lower mail survey costs and the higher personal interviewing costs. Also, depending on the size of the sample, the number of interviewers available, the number of questions, and question complexity, telephone surveys can be completed quickly.

In comparison with mail questionnaires, face-to-face and telephone interviews are much faster methods of gathering data. The need to train interviewers and their time spent traveling and contacting and interviewing respondents, however, make the face-to-face interview much more expensive than telephone interviews or mail or group questionnaires. Both forms of questionnaire can be longer and can include more complex questions (if the respondent group is one that reads well) than is possible with the telephone interview.

To administer a questionnaire in a group setting requires that it be practical to assemble the respondents. Thus, it is normally used in situations in which the sample is an entire group or a large portion of it, such as an Army company or battalion or all or many agency employees in one location. Group questionnaires are faster than mail questionnaires and permit some clarification of questions (but not to the same extent as interviews). As with mail queries, however, the language complexity used in group questionnaires must be commensurate with the reading level of the respondents.

In the past, GAO has used structured, face-to-face interviews to study such topics as

Chapter 2
What Is a Structured
Interview and When Should It
Be Used?

- the self-reported experience of Work Incentive participants while in and after leaving the program,
- the experience of participants trained by Opportunities Industrialization Centers, and
- the opinions of Drug Enforcement Agency agents concerning various operating procedures for two time periods.

We used face-to-face interviews in the first two cases because the respondent groups were not ones that tend to respond in large numbers to mail questionnaires, the subject matter was complex in relationship to their reading levels, and the interviews were too long to be done by telephone. In the Drug Enforcement Agency evaluation, the face-to-face interview was used because time did not permit a mail survey, the interview was too long for a telephone survey, and the agents could not be assembled in a group.

GAO used structured telephone interviews to study such topics as the satisfaction of

- small businesses with management assistance provided by Small Business Administration contractors and
- individuals having weatherization work done on their houses under a federal program.

In both cases, telephone interviews were used because the number of questions to be asked was small and time precluded a mail questionnaire.

Questionnaires were administered in a group setting as part of GAO studies of

- cadets and midshipmen at the four service academies in regard to attrition and
- employees of the Federal Communications Commission in regard to a proposed move from the Washington, D.C., area to Pennsylvania.

Chapter 2
What Is a Structured
Interview and When Should It
Be Used?

In both instances, it was practical to assemble the respondents. Also, in the latter case, time did not permit a mail questionnaire.

In general, GAO uses mail questionnaires much more frequently than group questionnaires, telephone interviews, or face-to-face interviews combined.¹ However, an understanding of structured interviewing techniques is essential for situations in which a mail questionnaire cannot be used. Additional discussion of structured interviews, questionnaires, and other DCIs, with examples of GAO applications, appears in chapter 10.1 of the GAO Project Manual.

¹Questionnaires are discussed in the methodology transfer paper entitled Developing and Using Questionnaires.

Designing a Structured Interview

Designing a structured interview requires more than just writing down a set of questions to be asked. In this chapter, we first examine the process by which the interview questions are identified, developed, and selected; then we describe standard procedures for composing and formatting the questions. These procedures aim to ensure that the data collected are reliable and valid and to facilitate trouble-free editing and analysis of data, while keeping the burden on the interviewee to a minimum.

Reading or even studying this transfer paper will not make anyone an expert in writing questions for structured interviews. We suggest, therefore, that you work with measurement specialists from the design, methodology, and technical assistance group in the division programming the assignment when you are planning to use a structured interview.

The DCI for structured interviews should be reviewed by a design, methodology, and technical assistance group if it involves 10 or more private citizens, private firms, or local governments; 5 or more state governments; or 25 or more federal agency officials or employees.¹ In certain executive agencies, GAO has designated representatives and established procedures that must be followed when using structured interviews and questionnaires.²

Identifying Variables and Developing Questions

The first step is to formulate the broad, overall questions to be answered by the evaluation or audit. Why is the study being done? What do we hope to be able to say or prove? Are we primarily describing what has taken place in a program? Do we want to compare what has happened with some

¹See GAO's *Project Manual* chapter 10.1 on methodology.

²For example, see GAO's *Operations Manual*, order 0175.5 (A-91), "Coordination of General Accounting Work at the Department of Defense, Defense Agencies and the Military Departments/Bureaus."

established or implied standard, a normative-type question? Or do we want to determine if a program has made a difference, a cause-and-effect type question? Examples of such questions are

- Descriptive: “How do graduates of the XYZ program for the unemployed seek out and find jobs in the community?”
- Normative: “How well does the program meet its goals for placing graduates in jobs?”
- Cause-and-effect: “Why do some graduates find jobs and others do not find jobs?”

The type of question asked will dictate the evaluation strategy. Also, certain strategies are more appropriate to answering certain questions.³ However, structured interviews, being simply a method of data collection, can be used with several evaluation strategies and, thus, in a variety of GAO assignments.

After the broad overall questions are developed, they must be translated into measurable elements in the form of hypotheses or questions. For the example mentioned above, to evaluate how participants found jobs would require developing such measures as the sources through which participants learned of available jobs, the number of employers contacted, and the number of job interviews arranged. Next, the target population must be identified. The target population is the source level (individuals, groups, or organizations) at which the information is to be gathered. Thus, in the study of how program participants found jobs after leaving

³Formulating overall evaluation questions and selecting evaluation strategies that provide answers is discussed in the methodology transfer paper entitled Designing Evaluations.

the program, the target population is the individual participants of the program who were trained.⁴

Next, develop a pool of questions that attempt to measure the variables under consideration. The questions may include various ways of measuring the same variable. For example, for age, you might ask, "How old were you on your last birthday?" or "On what day, month, and year were you born?" Both questions help you determine the individual's age, but the second elicits much more information. Decide which to use. From the pool of questions, then, the most useful or appropriate are chosen.

The identification, development, and selection of questions for our example, a study of how program participants found jobs after leaving a job-training program, are illustrated in table 3.1.

⁴Later in the evaluations, data analyses may actually be done at a higher (more aggregated) level. In the example above, the XYZ program may be conducted at several locations in a city, in many cities in a state, and in many states. Thus, several levels of analysis would be possible. The objectives of the evaluation and the sampling plan devised to meet those objectives, however, dictate the level or levels of data analysis.

Chapter 3
Designing a
Structured Interview

Table 3.1: Identifying, Developing, and Selecting Questions

Task	Example
Formulate overall questions	How do program participants find jobs after leaving the XYZ program?
Determine the kind of information needed	<ol style="list-style-type: none"> 1. Sources through which participant learned of available jobs 2. Number of employers contacted 3. Number of job interviews arranged 4. Number of interviews attended 5. Number of jobs offered 6. Time (in days) it took to secure a job 7. Time (in days) since participant left program to date of data collection 8. Relationship of job obtained to skill. . .
Identify target population	Program participants who have left the program (random sample)
Create a question pool	<ol style="list-style-type: none"> 1.1 How did you look for jobs? <ol style="list-style-type: none"> 1. Look in the newspaper? 2. Ask friends? 3. Go to a state employment office? 4. Go to a private employment office? 5. Look in the telephone book? 6. Drop in on companies? 7. Get information from radio or TV? 1.2 About how many jobs that you were interested in did you find out about from <ol style="list-style-type: none"> 1. The newspaper? 2. A friend? 3. The state employment service? 4. Private employment services? 2.1 How many employers did you contact about a job since you left the program? 2.2 Since you left the program, about how many employers did you contact about a job that you heard about from <ol style="list-style-type: none"> 1. The newspaper? 2. A friend? 3. The state employment service? 3.1 How many . . .
Select questions	<ol style="list-style-type: none"> 1.1 . . . 2.1 . . . 3.1 . . .

Composing Appropriate Questions

When composing interview questions, be sure they are appropriate—that is, relevant to the study, directed to the proper persons, and easily answered.

- Relevance: Questions should be relevant to the study being conducted and should have a good probability of yielding data needed for the final report. Although this would seem obvious, evaluators sometimes go on “fishing expeditions” and want to include all sorts of variables that can create an unnecessary burden on the interviewee and distract attention from the central purpose of the interview.
- Selection of respondents: Give preliminary consideration to which people can be expected to answer given questions. A question may be relevant to a given study, but the choice of persons to answer it may be inappropriate.
- Ease of response: Interviews are meant to obtain data that may otherwise not be documented or, if documented, may need some interpretation. This includes opinions and feelings about the study topic. You should attempt to construct questions that are relatively easy to answer and do not cause undue burden to the interviewee.

Avoid questions that require the interviewee to perform “audit work” to answer—that is, to consult records or other information sources. If used at all, such questions should be reserved for mail questionnaires. For telephone interviews, the questions should be even less complex, because there is less of an opportunity to help the interviewee understand. It is possible to send the questionnaire beforehand to the person who will be interviewed, requesting that he or she gather the necessary information in preparation for the interview.

Other questions (or the manner in which they are presented) that cause the interviewee discomfort should be avoided or used with extreme care. The

same is true of questions that would tend to incriminate or show the interviewee in a bad light, particularly since the interview might terminate if they were asked. Likewise, avoid personal questions about private matters that do not belong in a GAO study, as well as questions whose sole purpose is to embarrass the interviewee (such as testing or questioning the intelligence of the interviewee or seeking information about private habits).

If needed, ask sensitive questions in a mail questionnaire, where confidentiality or anonymity can be granted.⁵ Also avoid questions that could cause unnecessary confrontation, causing the interviewer and interviewee to take sides and do battle. This detracts from the interview task, may cause bias, and can seriously affect the validity of the answers given.

Also avoid questions that have no answers and avoid questions that, if you attempt to ask them, produce unusable results. These are not to be confused, of course, with questions for which the legitimate answer might be “no basis to judge” or “no opinion” (presumably, some interviewees will not have a basis to make a judgment or give an opinion).

Selecting a Question Format

Considerations in deciding on the format or type of question to use include how the question is delivered or presented, what the interviewee is asked, and available response alternatives. Among the types of questions we use are open-ended, fill-in-the-blank, binary-choice, and scaled-response, as discussed below.

⁵See our discussion on confidentiality and anonymity in chapter 7.

Open-Ended
Questions

The open-ended question provides no structure for the answer, allowing the interviewee to discuss what he or she wishes, not necessarily what the interviewer wants to know. By sharpening the question, you can focus it. For example:

- Broad question: “What happened to you while you were unemployed?”
- Focused question: “How did you manage to pay your bills while you were unemployed?”

Open-ended questions are easy to write. For initial research, they can be used successfully to elicit answers that contribute to the formulation of more specific questions and response alternatives. For a small number of respondents and where analysis may be qualitative, rather than quantitative, open-ended questions may also be useful. If possible, avoid using open-ended questions with larger numbers of respondents, whose answers need to be tabulated. Under such circumstances, content analysis should be done before attempting to tabulate.⁶

In CATI questionnaires, the questions should be designed in accordance with the guidelines established for structured telephone surveys.⁷ In addition, other practices apply. For example, open-ended questions should be avoided as much as possible, primarily because of the time it takes to type the answer. If the topics addressed in the questionnaire are at the exploratory stage, a CATI is not recommended. A CATI requires some degree of maturity in the understanding of the issues under investigation. To the extent that open-ended questions are included in a CATI, they should be designed for easy typing. Such questions take up considerable space in the computer data files. To

⁶Discussed in the methodology transfer paper entitled Content Analysis: A Methodology for Structuring and Analyzing Written Material. Also see chapter 9.

⁷Discussed in the methodology transfer paper entitled Developing and Using Questionnaires.

the extent possible, they should be moved to the end of the questionnaire and the interviewer should attempt to record the answers "off-line." These questions have the potential for interrupting the flow of the CATI and deflating the interview.

A question that actually is closed can be presented in such a way that to the interviewee it appears to be open-ended. Do this by preparing a list of potential answers and checking these off during the interview, as the interviewee mentions the various alternatives. Do not, however, read the choices to the interviewee. Such questions are more focused and specific than simple, open-ended questions and allow the range of possible answers to be narrowed. Question 1 in figure 3.1 illustrates the technique.

Figure 3.1: Question 1

1. Why weren't you satisfied with the plan? (DO NOT READ CHOICES) (*Check all that apply.*)

1. Didn't get training
2. Didn't get kind of job I wanted
3. Didn't get needed education
4. Didn't get further counseling after plan was formulated
5. Other (*specify*) _____

Fill-In-The-Blank Questions

This type of question has a simple answer, usually in the form of a name, frequency, or amount. Again, you may prepare a list of alternative answers to check off during the interview. Questions 2, 3, and 4 in figure 3.2 illustrate this type of question.

Figure 3.2: Questions 2, 3, and 4

2. Who completed your last performance appraisal? _____

3. How many hours did you work last week? _____ hours

4. What was your pay before deductions for last month? \$ _____

**Binary-Choice
Questions**

This is the typical yes-no, true-false type of question, a good format for obtaining factual information but generally not opinions or feelings. Since the interviewee is asked to make a commitment to one extreme or another, binary choice is considered a forced choice. Figure 3.3 shows an example.

Figure 3.3: Question 5

5. Have you ever served in the U.S. military?
(Check one.)

1. Yes

2. No

**Scaled-Response
Questions**

In the scaled-response question, you read or show to the interviewee a scale—a list of alternative responses that increase or decrease in intensity in an ordered fashion. There are three types: balanced, unbalanced, and rating and ranking scales.

Balanced Scales

The end points of the balanced scale are usually adjectives or phrases with opposite meanings—for example, very satisfied and very dissatisfied. As its name implies, the balanced scale contains an equal number of responses on each side of a reference

point or neutral response, as shown in question 6 in figure 3.4.

Figure 3.4: Question 6

6. How satisfied or dissatisfied are you with the typing ability of the secretaries in your division?
(Check one.)
1. Very satisfied
 2. Generally satisfied
 3. Neither satisfied nor dissatisfied
 4. Generally dissatisfied
 5. Very dissatisfied

This scale expands the binary-choice answer discussed above, permitting a range of answers that better reflect the way people hold opinions.

Unbalanced Scales

The unbalanced scale is used when no negative response is possible. It has a reference point (usually a “zero” point or “none”) and the value of the attribute increases for successive points on the scale. Intensity ranges from none to very great.

Figure 3.5: Question 7

7. On your last assignment, how much opportunity, if any, were you given to help develop staff working for you?
(Check one.)
1. Very great opportunity
 2. Great opportunity
 3. Moderate opportunity
 4. Some opportunity
 5. Little or no opportunity

Rating and Ranking Scales

In a rating question, the interviewee is asked to assign a rating to persons, places, or things according to specified criteria. The points on the scale can be either numeric or verbal. An example of a verbal scale is shown in figure 3.6.

Figure 3.6: Question 8

8. Using the rating scale, rate each of the individuals listed below on their ability to do Band I Evaluator work. (Check one for each person.)

Staff	Exceptional	Superior	Fully successful	Borderline	Unacceptable
1. Brown					
2. Green					
3. Johnson					
4. Martin					
5. Smith					

Whether verbal or numerical, a rating scale implies that the distance from one point to the next is the same on all parts of the scale.

In a ranking question, the interviewee is asked to place items in order according to a specified criterion, as shown in question 9 in figure 3.7.

Figure 3.7: Question 9

9. Rank the following individuals on their overall ability to do Band II evaluator work. Use 1 for the best, 2 for the second best, 3 for third best, 4 for the fourth, and 5 for the last. (Enter number for each.)

- _____ Brown
- _____ Green
- _____ Johnson
- _____ Martin
- _____ Smith

Ranking questions may have several types of instructions. You can ask the interviewee to rank all, as in the example, or to select the first (best) and the last (worst), the top three, or some other combination.

In contrast to rating, ranking does not imply that the distance between points is the same on all parts of the scale. For example, if Johnson, Green, and Smith were ranked 1, 2, and 3, respectively, the interviewee may not necessarily think that the gap between Johnson's and Green's performance is the same as the gap between Green's and Smith's.

When it is necessary to obtain the interviewee's opinion as to the distance between items (for example, how much better or worse one evaluator is than others), use a rating question. While a rating question may also produce an ordering, a respondent may well give two or more items the same rating. If you want the interviewee to choose between seven or fewer items but you do not care how much better he or she believes one item is than the others, a ranking question is likely to give you what you want. When a larger number of items must be ordered, however, it will probably be easier for the interviewees to rate them than to rank them. It is difficult to judge the order of a large number of items and avoid ties between items, especially in interviews. A final order can be produced by averaging the ratings over all respondents.

Number of Cues

The number of cues (alternative responses) for scaled-response questions depends on the type of interviewee and type of analysis expected. There is a physical limit, generally, to the number of cues to which an interviewee can react, probably around seven. GAO usually uses five-point scales. Respondents with a keen interest in the study can be expected to handle a greater number of cues. The more points on the scale, the better will be the eventual analysis of the data, since more cues provide a more sensitive measure and allow the analyst greater flexibility in selecting ways to analyze the data.

An even number of cues used in a balanced scale generally eliminates a middle or neutral point on

the scale and forces the interviewee to commit to a positive or negative feeling. The use of an odd-numbered scale permits a neutral answer and more closely approximates the range of opinions or feeling that people can have.

When the possible responses do not include “no basis to judge,” “can’t recall,” or “no opinion,” the interviewee may feel forced to select an answer that is inaccurate. The point is that some people honestly may be unable to answer. If you have good reason to believe this is so for members of the respondent group, include in the list of cues read or shown to the interviewees the most applicable of the alternatives—“no basis to judge,” “can’t recall,” or “no opinion.” If you do not do this, the interviewee may guess, make up an answer, or ignore the question.

Order of Cues

The order in which the cues are presented can be used to help offset possible arguments that the interviewees are biased toward answering the question in a particular way. Consider a situation in which GAO had preliminary evidence that participants in a training program were not getting job counseling. The following question could be asked:

“Job counseling involves someone talking to you about how to apply for a job, how to behave in an interview, etc. To what extent did you receive job counseling while you were in this program?”

The choices presented to the interviewee would be

- “very great extent,”
- “great extent,”
- “moderate extent,”
- “some extent,”
- “little or no extent.”

In this example, the order of presentation biases the choice slightly in favor of the program. Some interviewees who did not take a strong interest in the question might select the first choice, indicating that they received job counseling to a very great extent. This would tend to give us an overall answer that was slightly biased toward receiving job counseling.

When the cues form a scale, only at great expense could we totally eliminate the bias inherent in the order in which the alternative responses are presented.⁸ To repeat, the bias is slight. But when it does exist, we should use the logic of biasing the question against the hypothesis we are examining.

Wording of Cues

As indicated in the previous example, the scale used in the cues was the "extent" to which some action was performed. When an action or process is being assessed in a question, it is preferable to present the question and the cues in terms of the action. The previous question would generally be rephrased as "How much job counseling did you receive?" The cues could be rephrased as "A very great amount of counseling," "A great amount of counseling," "A moderate amount of counseling," and so on.

Unscaled-Response Questions

In an unscaled-response question, a list of cues is read or shown to the interviewee, who is asked to choose one from the list or to select all that apply. The list should consist of mutually exclusive categories. An "other" category is usually included as a last alternative, either to provide for many possible

⁸To totally eliminate this type of bias requires that half the sample be presented the cues in one order and the other half be presented the cues in the opposite order. In our example, half the sample would be presented a card on which "very great extent" was the first (or top) cue and "little or no extent" was the last (or bottom) cue. The other half of the sample would be presented a card on which "little or no extent" was the first cue and "very great extent" was the last cue.

(but thought to be rare) answers or if it is thought that some interviewees will come up with unique answers. Question 10 in figure 3.8 is an example of a question in which only one response is to be given; question 11 in figure 3.9 is a question in which the interviewee may check several responses.

Figure 3.8: Question 10

10. Educationally, what is the highest level that you have achieved?

(Check one.)

1. High school graduate
2. Some college
3. BS or BA degree
4. MS or MA degree
5. PhD
6. Other (specify) _____

Figure 3.9: Question 11

11. Please check the following possible eligibility requirements that you will be using to determine when you will offer remedial education to youths in your program.

(Check all that apply.)

1. Scoring below a specific performance level on a READING test.
2. Scoring below a specific performance level on a MATH test.
3. Teacher recommendation
4. Youth or parent request
5. Youth must be a dropout
6. Age limits
7. Other (specify) _____

Organizing Questions

In any DCI, the order in which the questions are presented is important. Early questions, which set the tone for the collection procedure and can influence responses to later questions, also help you get

to know the interviewee and to establish the rapport essential to a successful interview.⁹ For example, in an interview with participants in the XYZ program, the first few questions could review for accuracy data obtained from agency files such as family composition, age, and education.

The next questions should also be ones that can be answered with some ease, as you are still developing rapport with the interviewee. Should these early questions be too difficult or too sensitive for the level of relationship developed, the interviewee might end the interview. Remember also that the questions should hold the interviewee's attention; thus, you must begin to introduce some "interesting" questions and the sensitive areas covering the attitudes of the interviewee.

Present the questions in a logical manner, keeping the flow of questions in chronological or reverse order, as appropriate. Avoid haphazardly jumping from one topic to another.

Also, avoid introducing bias in the ordering of questions. For example, to determine what the interviewee thinks a program's advantages and disadvantages are, do not mention the possible advantages or disadvantages earlier in the interview.

Generally, the set of questions asked varies from interviewee to interviewee. Many questions are asked only if there is a specific response to a particular question. As a result, several questions may be skipped. These interrelationships among the questions constitute the skip pattern of the DCI. For face-to-face interviews and telephone interviews that do not use a CATI system, the complexity of the DCI's skip pattern should be kept to a minimum. Otherwise, it becomes very difficult for the interviewer to find the next question to be asked.

⁹Establishing rapport is covered in more detail in chapter 8.

One of the important advantages of a CATI questionnaire is that it allows for considerable complexity in the skip pattern, since the branching is handled entirely by the computer. Any number of paths can be followed through the questionnaire. Usually, the computer displays the next question in sequence. Alternatively, conditional skips can be programmed to go from one specific question to another somewhat later in the questionnaire. These skips can be based on how the interviewee answers a single question or on the responses to several questions.

One drawback to a CATI questionnaire is that multiple-choice questions permitting several answers are not easily handled. It is difficult for an interviewee to remember all the options when several can be chosen. As a result, multiple-choice questions allowing the interviewee to check all that apply (as illustrated in figure 3.9) should be broken down into separate questions, each of which is an alternative response that is "checked" or "not checked."

Layout Considerations

The layout or form of a printed DCI (for non-CATI applications) (see figure 3.10 for an example) is important; it is what you carry into the interview and use as a guide to conducting it. It gives you on-the-spot instructions for each question and allows you to record the answer. Later, the form is used to facilitate editing, keypunching, and the subsequent computerized analysis.

**Chapter 3
Designing a
Structured Interview**

Figure 3.10: Structured Interview Text

<p>Now I'd like to find out what you are doing.</p> <p>23. Are you now receiving any AFDC?</p> <p><i>If, "Yes" - Is this a full grant or reduced grant? (Check one.)</i> (20)</p> <p>01. <input type="checkbox"/> Yes - Full grant <i>(Go to Question 24.)</i></p> <p>02. <input type="checkbox"/> Yes - Reduced grant <i>(Go to Question 24.)</i></p> <p>03. <input type="checkbox"/> No <i>(Go to Question 29.)</i></p> <p>24. What is your status with WIN? Are you registered in training or what?</p> <p><i>(Listen. Insert comments and try to determine what code to assign. If necessary check records or check with WIN staff afterwards. Check one.)</i> (21-22)</p> <p>Comments:</p> <hr style="border: 0; border-top: 1px solid black; margin: 5px 0;"/> <hr style="border: 0; border-top: 1px solid black; margin: 5px 0;"/> <p>01. <input type="checkbox"/> 10. Working registrant status</p> <p>02. <input type="checkbox"/> 11. Part-time employment</p> <p>03. <input type="checkbox"/> 15. Working nonregistrant</p> <p>04. <input type="checkbox"/> 20. Institutional training</p> <p>05. <input type="checkbox"/> 30. Work experience</p> <p>06. <input type="checkbox"/> 31. WIN/OJT</p> <p>07. <input type="checkbox"/> 32. WIN/PSE</p> <p>08. <input type="checkbox"/> 33. Suspense to training</p> <p>09. <input type="checkbox"/> 34. Suspense to employment</p> <p>10. <input type="checkbox"/> 40. Intensive employability services</p> <p>11. <input type="checkbox"/> 41. IES/Group job seeking activities</p> <p>12. <input type="checkbox"/> 50. Other WIN noncomponent activity</p> <p>13. <input type="checkbox"/> 60. Unassigned recipient</p>	<p>25. Are you looking for work (different work)? <i>(Check one.)</i> (23)</p> <p>1. <input type="checkbox"/> Yes <i>(Go to Question 26.)</i></p> <p>2. <input type="checkbox"/> No <i>(Go to Question 30.)</i></p> <p>26. How are you going about looking for work? <i>(Do not read choices; indicate 1 = mentioned, 2 = not mentioned.)</i> (24-26)</p> <p>1. <input type="checkbox"/> On my own</p> <p>2. <input type="checkbox"/> Through WIN</p> <p>3. <input type="checkbox"/> Through CETA</p> <p>4. <input type="checkbox"/> Through Employment Services (ES, SES)</p> <p>5. <input type="checkbox"/> Through private employment agency</p> <p>6. <input type="checkbox"/> Other <i>(Specify.)</i> _____</p> <hr style="border: 0; border-top: 1px solid black; margin: 5px 0;"/> <p>27. To what extent are you having difficulty finding a job? <i>(Read choices; check one.)</i> (30)</p> <p>1. <input type="checkbox"/> Very great extent <i>(Go to Question 28.)</i></p> <p>2. <input type="checkbox"/> Great extent <i>(Go to Question 28.)</i></p> <p>3. <input type="checkbox"/> Moderate extent <i>(Go to Question 28.)</i></p> <p>4. <input type="checkbox"/> Some extent <i>(Go to Question 28.)</i></p> <p>5. <input type="checkbox"/> Little or no extent <i>(Go to Question 31.)</i></p>
--	--

Here are some considerations when designing the DCI.

Typeface. Generally the text to be read to the interviewee is set off in a different typeface from the instructions that you do not read to the interviewee. In the example presented in figure 3.10, the text to be read to the interviewee is presented in upper- and lowercase, the instructions in upper- and lowercase italics.

Continuation of questions. Generally, do not continue a question in the next column or on the next page, as you risk not having the entire question or all the response alternatives presented to the interviewee.

Boxes and lines. Provide open-top boxes for the interviewer to record answers to questions that require written responses. Place the box or line in a standard place beside each question to aid the interviewer and to facilitate editing, data entry, and subsequent analysis of completed questionnaires.

Keypunch numbers. These should be placed in a standard place beside each question to facilitate keypunching, when data are to be entered into computer files.

Skipping questions. If a certain response to a question means that interviewers are to skip the next question, specify this by placing a "GO TO" instruction beside the response.

Two computer programs are available to assist in the design and layout of structured interviews and CATI questionnaires. They are QUEST (developed by GAO's Office of Information Management and Communications) and QPL (Questionnaire Programming Language, developed by the design, methodology, and technical assistance group in the Human Resources Division).

QUEST makes it possible for an evaluator to create questionnaires using laser printers, incorporating typographic and graphic elements (such as check boxes, arrows, and italic type).¹⁰ QUEST automatically handles many of the layout considerations mentioned above. Draft and pretest versions of typeset questionnaires, incorporating current desktop publishing concepts, can be generated quickly. The designer develops a WordPerfect file employing codes that identify questionnaire elements and control typographical layout on the page. These codes make it possible to correct questionnaires easily, automatically renumbering pages, questions, and choices and altering keypunch instructions.

QPL is designed to automate many of the activities involved in gathering and preparing survey data for analysis.¹¹ It was developed primarily to implement CATI questionnaires within GAO; it can also be used as a data entry program for other DCIs. In this system, the questionnaire is first written in QPL, using a word processing program, and then compiled. The compiled version displays the questions on the computer screen, one at a time, and then waits for the interviewer to type a response. The interviewer can page forward and backward through the questionnaire to make corrections or review answers. The record of the interview is then added to a data file. The compiled version can also be converted into SAS and SPSS statistical analysis

¹⁰The question examples in this paper were prepared with QUEST. A manual is available from the Office of Information Management and Communications, Publishing and Communications Center. As new technology becomes available, QUEST will be improved to make it more powerful and even easier to use. More advanced desktop publishing software may also facilitate questionnaire development.

¹¹See the QPL reference manuals: QPL Reference Manual, Version 2.0 (HRD Technical Reference Manual 1, March 1990); QPL Data Collection Program (HRD Technical Reference Manual 2, March 1990); and QPL Data Editing Program (HRD Technical Reference Manual 3, March 1990).

programs that can process QPL data files. One of the programs in the QPL system reformats the computer questionnaire into a written questionnaire, numbering all the questions, drawing open-top boxes for the answers, specifying card and column locations for each answer in the data file, and writing skip instructions. Unlike QUEST, the questionnaire does not incorporate typographic and graphic elements, but QPL makes it easy to review and revise a questionnaire.

More on Interview Design: Avoiding Problems

In this chapter, we suggest further ways to compose good interview questions and to forestall problems with comprehension or bias. As an evaluator writing such questions, you need to consider the appropriateness and level of language used in the interview, the effects of qualifying language, and the importance of clarity. We also discuss the various kinds of bias that can creep into the wording of interview questions and their effect on the validity of the evaluation results.

Appropriateness of the Language

Whether interviewing language is appropriate or inappropriate may relate to what is said, how it is said, or when it is said, as discussed below.

What is said in the interview is basically dictated by the written, structured data-collection instrument. The DCI is prepared in advance and pretested and the interviewers are trained to use it; thus, to some extent, the appropriateness of the language has been tested. It is the task of the interviewer to transmit faithfully to the interviewee the meaning of the questions. In addition to wording the questions precisely, you may include supplemental language in the DCI, to be used if the interviewee does not understand the original wording of a question. If, in the course of the interview, the interviewee still does not understand and different language must be improvised, such improvisations should be noted and considered before the data are analyzed.

How it is said concerns the speech and mannerisms of the interviewer who controls the “presentation” and whose delivery of questions may alter their intended meaning. More detailed information on this topic appears in chapter 8.

When it is said refers to the context of the interview in which each question is placed. Although, in designing the DCI, you should be precise about the order in which questions are asked, you may introduce some variation during the actual interview to

clarify the questions, review information, or postpone potentially sensitive questions. Or, if the interviewee expresses concern or sensitivity to a given question, changing the language of a subsequent question might defuse the concern.

Level of the Language

When composing interview questions, consider the level of the language used. Seek to communicate at the level the interviewee understands and to create a verbal setting that is conducive to serious data-gathering yet one in which the interviewee is comfortable. In chapter 3, we touched on some of the writing approaches to use; here we deal with how the questions sound and the atmosphere the language creates. One problem often encountered is maintaining a level of language that is neither above nor below the interviewee's level of understanding.

Speaking over the interviewee's head includes the use of complex, rare, and foreign words and expressions, words of many syllables, abbreviations, acronyms, and certain jargon. Such language, while it may seem appropriate to the interviewer or evaluation team, may not be understood by the interviewee.

For example, when interviewing participants in a training program, the terms "OJT" and "PSE" in a question may be nothing but alphabet soup to the interviewees; even the words they represent, "on-the-job training" and "public service employment," may be over their heads. In conducting the actual interview, you would most likely have to give further definitions or examples of what was meant. When interviewing training program directors, however, the use of "OJT" or "PSE" would be appropriate if the interviewees use the terms daily.

Thus, to speak over the interviewee's head hinders communication. Interviewees who are embarrassed at their lack of understanding may either not

answer or guess at the meaning, which can lead to incorrect answers. Or the interviewee may get the impression that you really do not care about the answer and lose interest in the interview.

Speaking down to an interviewee is just as bad. You can oversimplify the language in the DCI to the point where the interviewees feel you regard them as ignorant. This approach is demeaning. You have contacted these individuals because they have important information to impart. To treat a person condescendingly—or to let it appear that you do—negates that importance.

Likewise, take care in using slang, folksy expressions, and certain jargon. While such language may help you develop rapport with the interviewee, the exactness of the communication may be lessened.

To avoid error in either direction, pretest both the final wording of the DCI and the interview approach.¹

Use of Qualifying Language

After composing an interview question, you may find it requires an adjective or qualifying phrase added or a time specified to make the item complete or to give the interviewee sufficient or complete information. For example, “How many employees do you have?” might become “How many full-time-equivalent employees do you have?” and “How many times have you gone to a physician?” might become “How many times have you gone to a physician in the past 6 months?”

If feedback is possible in the actual interview, the interviewee can ask for further qualification, where needed. If you have not included the necessary qualifiers in the DCI, however, another interviewer may qualify in a different way. This could make the resulting data difficult to summarize and analyze.

¹More detailed information on pretesting appears in chapter 5.

Also, interviewees, not realizing that qualifying language is absent, may answer the question as they interpret it. Thus, different interviewees would be responding to different questions, based on their own interpretations.

Clarity of Language

The style in which a question is couched can affect its clarity of communication. We discuss below such matters as question length, complexity, and clutter; double-barreled questions; double negatives; extreme language; and defining terms.

Length, Complexity, and Clutter

A question that contains too many ideas or concepts may be too complex for the interviewee to understand, especially if it is presented orally, which makes it difficult for the interviewee to review parts of the question. While the interviewee might be responding to one part of the question, the interviewer may be interpreting the response as a response to the entire question. You should set up more than one thought in separate sentences and give the interviewee the proper framework. For example, “How satisfied or dissatisfied were you with the amount of time devoted to helping you get a job while you were in the XYZ program?” becomes “Think about the training experiences you had while in the XYZ program. How satisfied or dissatisfied were you with the amount of time devoted to helping you get a job?”

Likewise, a sentence may contain clutter—words that do not clarify the message. Word questions concisely. Here are a few tricks to reduce sentence clutter:

- Delete “that” wherever possible—for example, “Others suggest [that] training can be improved.”
- Use plain language. For example, for “aforementioned,” use “previous” or “previously mentioned.”

- Avoid the passive voice. Substitute pronouns (“I,” “we,” or “they”) and active verbs; instead of “It is necessary to obtain,” use “We need.”
-

Double-Barreled Questions

A double-barreled question is a classical example of an unclear question. Consider the following: “Did you get skill training while in the program and a job after completing the program?” This question attempts to determine if there is a relationship between skill training and getting a job. But if the interviewee answers “yes,” this could mean “yes” to both parts, “yes” to the training part only, or “yes” to the job part only. Other interviewees, finding the question confusing, might not respond. You are presenting two questions but the opportunity to record only one answer. Both interviewee and interviewer may see the need for only one answer. State the questions separately.

Double Negatives

In phrasing a question, avoid the double negative, which is difficult to answer. For example, “Indicate which of the organizational goals listed below are not considered unattainable within the 2-year period” should be reworded to read “Indicate which of the organizational goals listed below are considered attainable within the 2-year period.”

Extreme Words

Avoid such words as “all,” “none,” “everything,” “never,” and others that represent extreme values. Rarely is a statement using such a word true, and the use of extreme words causes interviewees to avoid the end points of a scale. There are cases when the use of “all” or “none” is appropriate, but they are few. Where “yes” or “no” answers are expected, the results can be misleading. For example, if one employee is not covered in a question like “Are all of your employees covered by medical insurance?” a “yes” answer is impossible. A better question would be “About what percent of your employees are covered by medical insurance?”

Alternatively, choices can be provided, as in question 1 in figure 4.1.

Figure 4.1: Question 1

<p>1. What portion of your employees are covered by medical insurance? (READ THE CHOICES) (<i>Check one.</i>)</p> <p>1. <input type="checkbox"/> All or almost all</p> <p>2. <input type="checkbox"/> More than half but not all</p> <p>3. <input type="checkbox"/> About half</p> <p>4. <input type="checkbox"/> Some but less than half</p> <p>5. <input type="checkbox"/> None or hardly any</p>

Defining Terms

Where possible, define key words and concepts used in questions. For example, when speaking of "employees," define and clarify the term. Are we talking about part-time, full-time, permanent, temporary, volunteer, white-collar, blue-collar? An example of how this might be done is

"Consider people who work for your company, are paid directly by your company, work at least 35 hours per week, and are viewed as permanent employees. What percent of these employees . . . ?"

Of course, not all questions need be preceded by such a definition. As earlier questions are developed, definitions evolve. You may wish to list definitions in a separate section or on a card to hand to interviewees for reference.

Bias Within Questions

A question is biased when it causes interviewees to answer in a way that does not reflect their true positions on an issue. An interviewee may or may not be aware of the bias. Problems result when the interviewees are

- unaware of the bias and influenced to respond in the way that is directed by the wording or
- aware of the bias and either deliberately answer in a way that does not reflect their opinions or

- refuse to answer because the question is biased.

Bias can appear in the stem (or statement) portion of the question or in the response-alternative portion. Bias may also result when a question carries an implied answer, choices of answer are unequal, “loaded” words are used, or a scaled question is unbalanced. These are discussed below.

Implied-Answer Bias

A question’s wording can indicate the socially acceptable answer. An example is the question “Most GAO employees have subscribed to the U.S. Savings Bond program. Have you subscribed?” Interviewees who are concerned about being different from the norm may answer “yes,” even if they have not subscribed. The question could be restated as “Have you subscribed to the U.S. Savings Bond program?”

Questions can be worded so as to impel some people to answer in one direction and others in another. Yet both types of interviewee could be unaware of any bias in the wording. Such bias usually occurs when additional qualifying or identifying information is added to the question. There is bias in the question “Which plan is more acceptable to you: the one designed by Pat Brown, our chief economist, or the one designed by Chris Green, the consultant we hired?” The interviewee who is not familiar with either plan may answer on the basis of whether the plan was generated internally or externally to the organization, although this may have little or nothing to do with the quality of the plan. A better presentation would be “Whose plan is more acceptable to you: Pat Brown’s or Chris Green’s?”

Bias Resulting From Unequal Choices

When response alternatives are created, it is important that they appear to be equal. If undue emphasis is given to one, it may be easier for the interviewee to select that one. Question 2 in figure 4.2 illustrates a question with unequal emphasis,

and question 3 in figure 4.3 corrects the unbalance. Alternative 3 in question 2 is isolated from the two others because of the words “high-paid,” which sets those individuals apart from the others, and by the fact that alternative 3 is longer than the others.

Figure 4.2: Question 2

2. Who do you feel is most responsible for the poor quality of the training program?
(Check all that apply.)
1. Instructors
 2. Counselors
 3. High-paid managers who run the centers

Figure 4.3: Question 3

3. Who do you feel is most responsible for the poor quality of the training program? *(Check all that apply.)*
1. Instructors who teach the courses
 2. Counselors who advise which courses to take
 3. Managers who run the centers

Bias From Specific Words

When used in almost any context, certain words can be considered “loaded,” because they evoke strong emotional feelings. “American,” “freedom,” “equality,” and “justice” generally evoke positive feelings, while “communist,” “socialist,” “bureaucrat,” and “nuclear holocaust” may evoke negative feelings. Since it is difficult to control the emotional connotations of such words, it is usually best to avoid them.

Bias From Lack of Balance

When using a scaled question, avoid bias in the stem as well as in the response alternatives. A question that seeks to measure satisfaction with something should mention both ends of the scale in a balanced fashion. For example, question 4 in figure 4.4 shows unbalance in both the stem and the alternatives, while question 5 in figure 4.5 shows how this bias is eliminated.²

Figure 4.4: Question 4

4. How satisfied were you with the answers you received? (*Check one.*)
1. Extremely satisfied
 2. Very satisfied
 3. More satisfied than not
 4. Neither satisfied nor dissatisfied
 5. Not satisfied

Figure 4.5: Question 5

5. How satisfied or dissatisfied were you with the answers you received? (*Check one.*)
1. Very satisfied
 2. More satisfied than not
 3. Neither satisfied nor dissatisfied
 4. More dissatisfied than not
 5. Very dissatisfied

²Proper use of an unbalanced scale was discussed in chapter 3.

Considerations for Telephone Interviewing Instruments

In general, the same principles described above apply to the development of questions and answers for telephone interviews. However, some additional considerations come into play. The primary additional factor is that the cues available in face-to-face interviews are absent. It is not possible to observe the interviewee's reactions (including confusion, uncertainty, or hostility) and make allowable adjustments in conducting the interview.³ Making questions shorter, breaking multiple-choice questions into binary questions, and conducting some pretests face-to-face will overcome some of these difficulties.

Another loss in telephone interviewing arises from the impersonal nature of the telephone. An interviewer has a tendency to become flatter in presentation. The interviewer must counter this tendency by being continually aware of the enunciation of questions. In the QPL system, some words are capitalized, underlined, or put into bold type to help the interviewer maintain appropriate pitch and emphasis.

In summary, designing a structured interview form is not simple. It involves many considerations and choices: the specific questions to be asked, their format, language order, and layout. In this chapter and chapter 3, we have covered briefly the basic principles that should be followed in making these choices.⁴

³See chapter 8 for details.

⁴For more information, consult Bradburn and Sudman (1981) or Sudman and Bradburn (1982), as listed in the bibliography.

Pretesting and Expert Review

Pretesting and expert review constitute perhaps the least appreciated phase in the development of a structured interview.¹ In the desire to meet deadlines for getting the job done, staff may ask “Why not eliminate the pretest?” or “Do we need outside opinions on the interview form?”

But these are perhaps the most important steps in the development of the interview, an iterative process that uses continuing input from evaluators and technical specialists to derive the final product. As Cannell et al. (1989) indicate, when the evaluator has little experience with a topic or when the interviewee has difficulty with a question, substantial work may be necessary to develop questions that will obtain the desired results. Research has shown that question formulation may alter results by as much as 50 percent. The pretest and expert review processes give the evaluators feedback as to whether its efforts stand a chance of doing what they are designed to do.

Following pretesting and expert review, the DCI is redesigned as needed—an iterative process that occurs after each pretest or group of pretests.

Purpose of Pretest

In pretesting, we test the DCI with respondents drawn from the universe of people who will eventually be considered for the study interviews to predict how well the DCI will work during actual data collection. The pretest seeks to determine whether

- the right questions are being asked to obtain the needed information,

¹The term “pretest” is not interchangeable with “pilot.” “Pretest” is usually used in connection with the testing of a structured interview or questionnaire, while “pilot” implies a test of all or most of the complete study design at one field location before proceeding to implement the design at all selected locations.

- the contents or subject matter of each question is relevant to the respondent and the respondent has the knowledge to answer the question, and
- the wording and procedures used in conducting the interviews are adequate to ensure that valid and reliable results are obtained.

Research (Cannell et al., 1989) has shown the following to be among the types of problems that arise with survey questions:

- difficulties in asking the question because of complex sentence structure, tongue twisters, or words that are difficult to pronounce;
- difficulties in comprehending the question because of difficult vocabulary, complex sentence structure, or lack of specificity about information or the form of information that is needed;
- lack of common understanding of terms or concepts in the question because they are interpreted differently by different respondents or they are interpreted differently from what the interview designer intends; and
- difficulties in answering the question because the information is inaccessible or unavailable to the respondent, to the extent that the respondent does not want to make the effort to obtain the answer.

Pretest Procedures

The number of pretests typically varies depending on the size of the survey and the range of conditions that may affect the survey results. For structured interviewing of thousands of respondents, 25 to 75 pretests might be conducted. Sometimes, when the sample is less than 500, a dozen or fewer pretest cases are sufficient, provided they bracket the range of data collection conditions. Discuss the exact number with the measurement specialist who designed the DCI. To a great degree, the pretest procedures for the structured interview simulate what would be done during actual data collection. It is important to test as many of the procedures involved in conducting a structured interview as

possible, including the selection of and contact with the interviewees. In part, pretests should be conducted in the same mode to be used in the actual interviews—that is, the face-to-face interview pretested in person and telephone interviews over the telephone. However, telephone and mail surveys should also be tested in part in face-to-face interviews. For CATIs, which generally have fewer than 300 interviews, a dozen pretests might be sufficient. These pretests should be conducted both in person and over the telephone.

Who Conducts the Pretest

Two types of staff should represent GAO at the pretest:

- the evaluator working on the job, because he or she can best address questions on the content of the DCI and the background of the evaluation, and
- the measurement specialist who designed the DCI, because he or she needs to evaluate the interview process, including how the DCI works, and suggest improvements.

The measurement specialist acts as the interviewer—that is, asks the questions on the first and perhaps the second pretest—while the evaluator observes. On subsequent pretests, the evaluator asks the questions and the measurement specialist attends as observer.

Selecting and Contacting Pretest Interviewees

Pretest interviewees are drawn (not necessarily randomly) from the universe being considered for the final study. If the universe is relatively homogeneous—for example, welfare recipients—the pretest subjects need not be exactly balanced as to various attributes. With a heterogeneous group, such as taxpayers or U.S. citizens, however, try to obtain pretest interviews with high- and low-income people, old and young, the highly educated and less-educated, and women and men. Ideally, the DCI is

pretested with several of each of the different kinds or types of individuals in a heterogeneous group.

Contact pretest interviewees by telephone or in person to arrange a pretest session. If possible, follow procedures similar to those proposed for actual data collection. Identify yourself, describe what kind of agency GAO is and what it does, explain the nature of the study, and indicate the importance of their participation. If this is a face-to-face pretest, ask the interviewee to participate by arranging to meet in a place that is convenient to the interviewee and free of distractions. If this is a pretest of a telephone interview, arrange a time that is convenient for the interviewee. (For a more detailed explanation and copies of text to be followed, see chapter 7.)

Conducting the Pretest

The initial steps of a pretest are the same as for actual data collection. Give the interviewee any appropriate background information, even if you have covered this while setting up the interview appointment. Since an interview is interactive, the interviewee will probably provide a great deal of feedback in addition to answering the questions. Problems with the DCI or procedures often become evident immediately and may be dealt with then, so that the interview can proceed. Often, if an instruction, word, or concept is not understood, the interview cannot continue.

Ideally, however, it is desirable to run through the entire interview without getting sidetracked. This way, you can examine the flow of the interview and estimate the total time needed to complete it.

During the pretest, then, your tasks as interviewer are to

- carry on the normal introduction and questioning of an interview without too much interruption in the flow,

- provide explanations or try alternative wordings when the interviewee cannot or will not answer a question and note the changes introduced,
- record the answers on the DCI so the recording procedure and coding technique can be tested,
- make notes on situations that occur during the interview that indicate problems with the instrument or procedures, and
- conduct a debriefing at the end of the interview to learn what the interviewee thought of the interview but did not express.

With respect to the second item, providing explanations or alternative wording must be done carefully, since interviewer bias can occur. The interview is written as bias-free as possible. In deviating from the prescribed text, you may not have time to rephrase the question adequately and can make a slip in wording that favors or is slanted toward your approach to the situation.

For telephone interviews, it may be easier to conduct the pretests and they may be more informative. The interviewee should be informed that a measurement specialist will be listening for purposes of refining the instrument. It may be possible to use a speaker phone to allow more members of the team to listen, take notes, and record answers without intruding. With the interviewee's permission, the interview may be taped to allow for more detailed examination of problems. With these possibilities, pretesting telephone interviews may be a lot smoother than pretesting face-to-face. However, as mentioned above, remember to include some face-to-face interviews.

Identifying Problems

After a pretest, the evaluator and the measurement specialist review the interview process and attempt to identify any problems that the interviewer has in asking the questions or the interviewees appear to have in answering questions. If the pretests disclose problems such as ambiguous interpretation, or

other difficulties (discussed below), you must revise the interview and continue the tests until the problems are resolved, even if this requires unplanned extra time. Premature termination of pretests can result in questionable data. Major indicators of problems include the following:

- slight changes by the interviewer in reading the question;
- major changes by the interviewer in reading the question or not finishing reading the question;
- interrupting the question-reading with an answer;
- difficulty in answering, as evidenced by the interviewee's nervousness, inconsistency in answers from question to question, inconsistency by reflecting back and changing answers, or taking too long to answer;
- clarification, explanation, or repetition of the question requested by the interviewee;
- qualified answer provided by the interviewee;
- inadequate answer provided by the interviewee, including giving a different answer than one listed on the instrument or known to the interviewer;
- "Don't know" answer given by the interviewee; and
- refusal to answer by the interviewee.

The problems fall into two basic categories—those related to instrument design or administration and those concerning the interviewee's lack of knowledge or reluctance to answer. The first type can be controlled by the staff designing the instrument and are covered in chapters 3 and 4, while the second is merely recorded as observed behavior.

Research has found (Cannell et al., 1989) that pretest interviewers are not consistent in identifying problems with the questions or providing guidance for their revision. Responses can vary by as much as 50 percent when there are no adequate controls over the quality of the questions and procedures. Two techniques (categorization of respondent behavior and use of probe questions) that have been developed are useful particularly when the

number of interviewers is large. The first method simply involves tabulating for each question how often each one of the problems mentioned above occurred across all interviews. A small percentage of interviews is expected to have some problem for each question. If, however, for a given question, a high percentage of interviews has a specific problem, this suggests that a question needs revision.

The second method, use of probe questions, can be used by itself or to clarify the nature of the problems identified from the first method. Special probe questions may be included in the interview or may be used at the end to ask interviewees to elaborate an answer, explain how they interpreted the questions or answers, or describe any difficulties. There are three types of probes:

- general probes ask for more information about particular questions or the entire interview,
- information retrieval probes ask whether interviewees had difficulty in recalling the information necessary to respond to the question, and
- comprehension probes ask interviewees how they interpreted particular questions or phrases.²

Purpose of Expert Review

Because no instrument is perfect, it is generally useful to seek outside commentary on our approach. We seek expert review on assignments using structured interviews to help us determine whether

²For more information on conducting a pretest debriefing, see the methodology transfer paper entitled Developing and Using Questionnaires.

- the questions being asked and the manner in which they are asked are adequate to answer the overall question posed in the evaluation,
- the intended interviewee group will have the knowledge to answer the questions, and
- the instrument is as well constructed as possible within state-of-the-art confines.

In many instances, officials from the agency whose program is under review serve in this capacity. By obtaining agency input at this stage, we avoid potential problems after data collection, when time and money have already been spent. In other cases, staff in other design, methodology, and technical assistance groups, PEMD staff, or individuals with subject-area or evaluation expertise can provide expert review. In particular, subject-matter experts in membership associations who provide us with lists of the respondent universe or sample can provide expert review.

Persons providing expert review are not acting as interviewees. They do not answer the questions but instead provide a critique.

Instrument Redesign

The evaluator and the measurement specialist consider the results of the pretest and expert review and make appropriate changes to the DCI. If changes are minor, the instrument can be used without further pretests; if extensive, another series of pretests may be necessary.

If pretesting can be spread over a longer period of time, more versions of the instrument can be tested and a smaller number of interviewees can be used with each version. Changes that are obviously needed can be made and the revised version can be used in the next pretest. This allows us to use a relatively more perfect version on each round of pretests.

Training Interviewers

In most cases, our own evaluators conduct structured interviews for GAO studies, but occasionally we use employees of other agencies or contractors. Regardless, the interviewers must be trained in the purpose of the evaluation and the procedures for conducting the interview.

Training Methods

GAO uses various ways of training its interviewers and helping them maintain their skills throughout the data-collection period: a job kickoff conference, an interview booklet, role-playing and field practice, and supervisory field visits and telephone contacts. In addition to the items discussed below, interviewer training should emphasize the skills described in chapter 8 for conducting the interviews, with particular attention to structured interview tips, probing techniques, and reinforcements. These are also discussed below.

Kickoff Conference

For most projects of any size, a GAO division holds a kickoff conference to tell the staff from the regions and other divisions the purpose of the evaluation, to make assignments, and to answer questions. When a project is to include structured interviewing in the data-collection phase, the conference is usually extended so the interviewers can be given detailed instructions on the use of the DCI. Preferably, all potential interviewers should attend.

If a region sends only one representative to the kickoff conference, for example, it should be an individual who will be conducting interviews for the study. Not all aspects of the training can be written into the interview booklet (discussed in the next section); thus, practice sessions must involve, along with the measurement specialist, those who will actually conduct interviews and possibly will train others in the region to do so.

The training begins with the evaluator in charge and the measurement specialist reviewing the purpose of the study and how the interview data will fit into its overall objectives. Then, the data-collection procedures are covered in detail, using the interview booklet. The trainers discuss the interview form, question by question, including the need for the data, possible rephrasing to be used if a question is not understood by the interviewee, how to record the answers, and other matters they think could arise. The trainees can ask questions, clarify items, catch typographical errors in the DCI, and suggest possible changes from their experience. Even at such a late date as the kickoff conference, changes can be made in the DCI to preclude problems being carried into the actual interviews.

Among the potential problems that the trainers usually make special efforts to address is making sure that the interviewers

- know what an adequate answer to each question is. Without this knowledge, they may accept an inadequate answer. A structured interview is set up to show the interviewer, by means of the response choices, what is adequate and what is inadequate. For this to be learned, the interviewer must understand the DCI.
- ask the questions correctly. The words are there on paper; the interviewers need to be persuaded to use them in the way they are presented to ensure standardization of meaning and delivery and elimination of bias. Even though the instrument is pretested, some interviewees will still have trouble understanding the language. The interviewer must know enough about the question that if rewording it for clarity is the only recourse the interviewer has, the rewording will not violate the intent of the question.
- do not omit questions they think are answered by other questions. Answers are expected to all questions, unless instructions call for an item to be skipped or the interviewee refuses to answer.

(Refusal can be considered an answer.) If the interviewee gives the answer to a question before it is asked, the interviewer should either ask the question anyway or give it as a statement for the interviewee to affirm.

- do not introduce bias in the way they ask the questions (see the discussion of this in chapter 4).

Interview Booklet

Where the interview questions are limited in number and not very complex or difficult and the staff members who will conduct the interviews helped develop the DCI, we use the kickoff conference alone to inform the interviewers in detail how each question should be handled.

If, however, a large-scale interview effort is undertaken, GAO project staff may prepare a booklet that discusses in detail each question in the DCI. (The booklet is similar to that issued by the Bureau of the Census to its enumerators.) Typically, GAO's booklets cover not only the interview questions but also other matters such as sampling procedures, contacts with interviewees, and coding procedures. These are discussed below.

Sampling Procedures

Where statistical sampling procedures are to be used to select interviewees, the booklet shows the interviewer how to identify the universe and select the sample. The booklet may include a random-number table, when necessary, and describe both simple random samples and more complex two-stage procedures.

Interviewee-Contact Procedures

Rules are provided for contacting the potential interviewee and deciding what to do if the person refuses or cannot be located. An example is given of a phone conversation to set up the interview. Also covered is the log interviewers must keep of all interview contacts to ensure that proper sampling is maintained. The log makes it possible to adjust the

universe later and to examine possible effects of nonresponse. For CATIs, many of the contact and logging procedures are handled automatically by the computer. How this is to be accomplished should be described to the interviewers during training.

Coding Procedures

The booklet shows interviewers how to code the various types of question to facilitate editing and keypunching the answers and reviews different types of questions. This is handled automatically for CATIs.

Role-Playing
Practice

This is nothing more than two staff members taking turns "playing" interviewer and interviewee, a training method that should start at the kickoff conference as a group session with the measurement specialist observing and critiquing. The role-playing can continue when the staff members return to their regions, particularly if regional staff members who did not attend the conference will also be conducting interviews.

Such role-playing gives staff members the chance to become familiar with the instrument from both sides of the interview. The person playing the interviewee should challenge the interviewer by giving him a "hard time," perhaps refusing to answer questions or pretending not to understand them. Sometimes this serves to show the weaknesses of questions that are unclear or lack sufficient response alternatives. If so, the evaluator in charge or measurement specialist should be notified so the items can be changed or clarification can be given to all interviewers.

For CATIs, the interviewers must also be trained in the software requirements. This should be done after the training in the details of a paper version of the DCI. The computer training first focuses on the mechanics of using the computer for a CATI,

showing the interviewers how to start the CATI, move the cursor and step through each screen, move back and forth between questions, and identify particular situations that may arise.

After the essentials of the DCI and the computer have been covered, the interviewers can proceed to role-playing, this time using the computer set up for office-to-office mock interviews. The evaluator in charge or measurement specialist should observe these sessions to identify not only weaknesses in the DCI but also any difficulties in using the computer. This role-playing should be practiced for a half to a full day.

Field Practice

Once evaluators are in the field at the first site, they should oversample the number of interviewees needed for that site and use some for field-practice interviews. These interviews are planned as throw-away cases, identified as such in advance of the interview. The data derived from an interview are not used in the final analysis, regardless of whether the interview went well or poorly. Interviewing real interviewees who do not count gives interviewers a chance to get rid of any anxiety and test out their approach. The interviewees, however, should not be told that this is a practice session. To them, this is the real thing; they will, therefore, exhibit all the cautions and concerns of any interviewee.

Obviously, field practice takes some time and should be built into the project schedule. After practice, the interviewers should discuss any problems they had and decide where they need to change their approach or learn more. Any lasting concerns should be relayed to the evaluator in charge or the measurement specialist.

Supervisory Field Visits

Normally, the evaluator in charge makes field visits during the course of an evaluation. A visit early in the data-collection phase, when interviewing has

just begun, is valuable, allowing the evaluator in charge to review the procedures being used to conduct the interviews and observe some interviews firsthand. This quality-assurance checking enables the evaluator in charge to ascertain that interviewers are carrying out the standard practices designed into the structured-interview procedures. If possible, the measurement specialist should participate in some of the visits.

For CATIs, it may be more difficult to maintain supervisory controls. To the extent possible, each interview should be recorded on paper as entries are made into the computer, so that the accuracy of the computer input can be verified. Large organizations that conduct CATIs frequently provide the capability of a supervisor to monitor calls by interviewers, usually at random or when the interviewer experiences problems. This is not usually possible for GAO CATIs. In some instances, it may be useful to tape initial interviews, with the interviewee's permission, in order to remove any final problems associated with the interview administration.

**Supervisory
Telephone Contacts**

The evaluator in charge and measurement specialist form a team that keeps interviewers informed of changes in procedure and receives comments from the field on progress and problems encountered. These telephone contacts serve as the final step in training interviewers.

**Interviewer
Qualifications**

Many GAO interviews are highly sensitive, and the data to be obtained can be influenced by subtle elements that are in the control of the interviewer. When GAO uses outside sources to supply interviewers, it usually retains the right to examine the work of interviewers and, if there is cause, suggest that some be replaced. The same applies to GAO evaluators whom the region or division assigns to the project. Staff members who are reluctant to conduct the necessary interviews or exhibit some bias

may not be right for the job and could jeopardize the data-collection effort.

For CATIs, the skill level and content knowledge of interviewers can be lower than for face-to-face interviews because the questions are generally simpler and fewer probes need to be used. As a result, GAO contracts for CATIs or the use of short-term, part-time staff have been quite successful and provide alternatives to the use of evaluators.

The qualifications that interviewers exhibit during the various training opportunities should be evaluated by supervisors. If there are any problems that cannot be corrected through retraining, these interviewers should be replaced.

Selecting and Contacting Interviewees

This chapter touches briefly on the selection of interviewees and then discusses in some detail contacting the prospective interviewees, arranging the interview, and protecting the interviewee (through informed consent and guarantees of confidentiality or anonymity).

Selection of Interviewees

For some structured interviews, because there is only one person who fits the category of interviewee (for example, state officials responsible for welfare programs), no selection process is needed. More-complex selection procedures that are required—for example, when the sampling plan calls for a random sample of program participants or other respondent groups—are covered in some depth in the methodology transfer paper entitled *Using Statistical Sampling*. When complex sampling techniques are used and a list of interviewees is generated by computer, control over the selection and contact of interviewees can be automated, as described in more detail below.

Contacting Potential Interviewees

Once the potential interviewees have been selected, you must contact them, explain what GAO is doing and why you need their assistance, and arrange an appointment. The interview booklet sets out rules to be followed in contacting the interviewees.

Frequently, when structured interviews are used, interviewees are program participants or beneficiaries of federal programs. The universe list is developed for a given point in time and a sample is drawn. By the time the sample is contacted for interviews, months may have passed. This means some of the people selected for initial telephone contact will have moved away, died, or otherwise become inaccessible to GAO interviewers. Thus, we oversample and set up rules for replacing individuals who cannot be located. Such provisions are illustrated in figure 7.1, which contains rules that

Figure 7.1: Interviewee Contact Procedures Example

In order to provide a comprehensive assessment of the XYZ program nationwide, a complicated sampling plan has been devised to select participants to interview.

The sampling plan will allow us to interview as few as 16 participants at each selected location, thereby alleviating limitations on our staff and time. However, since only 16 participants will be representing all participants at a site, the sampling and interviewing rules for selecting the participants must be strictly adhered to. Failure to follow the rules will seriously jeopardize the validity of our review.

The rules for randomly selecting the participants for possible interview should be closely followed to yield the 16 planned interviews. Log sheets will be provided for you to record your attempts to contact potential interviewees. The rules for random selection require that the people interviewed must be the first ones selected. Only if you absolutely cannot reach one of the first ones can you move down the list to try the next participant for possible interview. We have set up some rules to follow which allow you to drop a participant from the list.

You may drop a potential interviewee

1. if the participant has no telephone and you cannot contact him by phone through his job or through the XYZ office,
2. if you contact the person and he absolutely refuses to be interviewed
3. if you reach someone other than the participant at his number and that person indicates that the participant is out of town and will be back after you have left that site, or
4. if you have called the participant four times and received no answer and the four calls were made morning, mid-day, and evening of one day and once the next day.

Other rules and suggestions will be discussed at the kick-off conference.

Example log sheets follow this section.

GAO used to review a nationwide program requiring the interviewing of program participants.

When contacting the interviewee by phone, use a *standardized approach*. This ensures that you do not omit any important information. An example of such an approach is presented in figure 7.2. For CATIs, the introductory script can be put onto the computer. Naturally, if unexpected events occur, you may have to deviate from this guide. The interview booklet may contain some samples of unexpected events and provide some guidance on how to deal with them.

Maintain a log of all attempted contacts, with a record of each interviewee's name and address, telephone number, date and time of the attempted contact, and the result. This information will be of use later in *determining the possible effects of nonrespondents on the results*. Also, it gives the analyst a means of adjusting the universe and plays a role when response-weighting is used. An example of such a log appears in figure 7.3 and how it looks partially completed is in figure 7.4.

Chapter 7
Selecting and
Contacting Interviewees

Figure 7.2: Example of Telephone Contact With Potential Interviewee

WHEN YOU GET THE INTERVIEWEE ON THE PHONE, YOU SHOULD SAY SOMETHING LIKE:

Hello, (name of interviewee), my name is (give your name). I work for the U.S. General Accounting Office. We work for the U.S. Congress. Currently, we are doing a study of services provided under the XYZ program. That is the program that provides (briefly describe the program). When can we set up an appointment for you to spend 30 minutes or so with me to answer some questions about the program and your experiences with it?

IF THE INTERVIEWEE AGREES, SET UP THE APPOINTMENT.

IF HE OR SHE REFUSES, EXPLAIN THE IMPORTANCE OF THE INTERVIEW BOTH TO THOSE WHO PARTICIPATE IN THE PROGRAM AND TO THE GOVERNMENT. YOU CAN SAY:

We are trying to determine if the program is helping those like yourself who are participating in it. Congress has asked us to find out what is good about the program and what should be improved. To do this, we must talk to you and others who have been in the program. We will only take about 30 minutes or so of your time. We will try to arrange it when you have time.

IF HE OR SHE STATES THAT IT IS NONE OF THE GOVERNMENT'S BUSINESS, YOU CAN SAY:

Well, the government is providing the money for the program. If it is a good program, they should know that; if it is not doing the job, it should be changed. Our report will help the government decide what should be done. That's why we need to talk to people who have been in the program and really know what is going on.

IN ANY CASE, TRY NOT TO LOSE THE INTERVIEW.

IF ALL EFFORTS FAIL, RECORD THE REASON FOR THE REFUSAL IN THE RESULT COLUMN OF THE LOG SHEET.

Chapter 7
 Selecting and
 Contacting Interviewees

Figure 7.4: Interviewee Contact Log Filled In

INTERVIEWEE CONTACT LOG

Job title Review of WIN Universe size 84 Unit no. 350
 Job code 205021 Sample size 20 Page 1 of 1
 Auditor D.L. Stone

Rank no.	Name address phone number	Case I.D.	Attempts to contact & results			
			Date & time	Result	Date & time	Result
36	James Noyes 123 Line Ave. City, ST 04321 432-7999	-	2/4 3:00p	No Answer		
			2/4 8:00p	No Answer		
			2/5 9:30a	No Answer		
			2/5 12:15p	No Answer: dropped		
49	William Cash 333 Best St. City, ST 04321 432-5566		2/4 3:15p	Moved to another state		
06	Fred Boh 661 Way Way City, ST 04321 432-1131	1	2/5 9:00 a	Out shopping - will call back		
			2/5 1:00 p	Call sat for 2/6 10a.		
			2/6	Interviewed		
80	Grace Store 720 Green St. City, ST 04321 432-6049	2	2/5 1:30p	No Answer		
			2/5 2:30p	Int. sat for 2/6 1:00p.		
			2/6 1:00p	No show		
			2/6 3:00p	Answer to 2/6 4:00p.		
			2/6 4:10p	Interviewed		

Note: Attach completed interview forms to these logs.

For CATIs, particularly those using the QPL system, a data base of respondents can easily be generated and used to provide automated call sheets. Certain

information, such as the time at which an interview was conducted, can be entered automatically by the computer. In addition, if information about the interviewee is already available from other sources, it can be entered directly into the record being generated without the need to ask the interviewee (unless some verification is wanted). Finally, when the interview is completed, selected information can be transferred to the automated call sheets to record the progress in administering the DCI.

A main objective when selecting and contacting interviewees is to avoid bias. By following set procedures, you can minimize wrong selections made by mistake or because of ease in contacting them.

Interview Arrangements

When you interview an individual for a GAO evaluation or audit, the interviewee usually is doing GAO a favor. You should, therefore, make the interview arrangements, including time and site, as convenient as possible for the interviewee.

This may mean conducting the interview at what is for you, the interviewer, an inconvenient hour, such as early morning or late evening. The location might be a GAO office, an audit site, space provided by the agency under review, or some other public place. If this is not convenient for the interviewee, you may have to travel to his or her home or place of employment or to some other such location. For example, if you must interview farmers, you cannot expect them to take time from their work routine to travel to a place to meet you; you would need to go to the farms.

If the interview contains sensitive questions, holding the interview in certain locations might create difficulties. For example, if you are questioning participants of a welfare program about the services and treatment they are receiving, it would be unwise to conduct the interview in the welfare office. Such a setting might cause interviewees to

omit negative comments about the office and its personnel out of fear that this information would be overheard and affect their benefits.

When interviewing people in their homes, you may encounter frequent interruptions from other family members, neighbors, and telephone calls. Television and radio programs can also be distracting. Interruptions and distractions also occur when people are interviewed at work. Nevertheless, there are advantages to interviewing people in their own settings: they generally feel more comfortable, they have not been inconvenienced by having to travel to the interview, and they may have records and other sources of information, including other people, at their disposal. Thus, choose the interview setting carefully. On balance, it is more important to conduct the interview in a setting in which the interviewee feels comfortable than to insist on a setting that offers no distractions.

For CATIs, the same general principles are used to set up an interview. You should assume that your first contact with the interviewee will not be a satisfactory time to conduct the interview and ask the interviewee when would be a good time to set aside a specific amount of time. You should be particularly alert to any impatience on the part of the interviewee in arranging the time. One method is to mail a postcard to the interviewee asking for a convenient time. Sometimes it is useful to highlight the main topics the interview will cover. In addition, if there is specific information that the interviewee will need to gather before the interview, this should be included on the postcard or in a letter.

If the interview is to be taped or if a supervisor is listening, the interviewee should be so informed. If the interviewee objects, alternative arrangements should be available. Acceptable alternatives should be described in the interview guide.

Protecting the Interviewee

You may encounter interview situations that result in the interviewees' speaking of themselves or others negatively. This could come from asking questions on such sensitive issues as personal habits or behavior, attitudes (for example, political or religious views), or reactions to an employer, boss, or employees.

To obtain cooperation from interviewees and improve the quality of the data and the response rate, you may need to grant some kind of assurance to the interviewees that the data collected will not be used in a manner that could harm them.

When you first contact interviewees and again when meeting for the interview, usually give them some idea of what types of question you wish to ask and seek their cooperation. This is called obtaining informed consent: revealing the contents of the interview in advance of the actual questioning, thus giving the interviewee a chance to refuse to comply with the interview request. GAO does not use the more defined procedure in which the interviewee is asked to sign a statement of understanding. Providing advance information is preliminary to GAO's actual guarantee of protection, which takes the form of confidentiality or anonymity.

Confidentiality means that the evaluator could associate the interviewee's name with specific responses but promises not to do so. Evaluators must obtain specific written approval before making pledges of confidentiality. GAO management grants approvals and, in certain cases, the congressional requester must also agree to honor the pledge. Once a pledge of confidentiality has been given, GAO safeguards the information consistent with the pledge. For details of pledges of confidentiality, refer to GAO's General Policy Manual, chapter 7.0, and GAO's Project Manual, chapter 7.1.

Anonymity assures that GAO staff performing the work on the evaluation will be unaware of the

responses of individual interviewees. When data are collected through face-to-face interviews conducted by GAO interviewers, granting anonymity to the interviewees would be impossible.

Conducting Interviews

Each participant in the interview—interviewer and interviewee—has a role to perform and a set of behaviors that assist in the performance. Because the role and behaviors of each one influence the conduct of the interview, they affect the other participant. The interviewer's role and behaviors can be prescribed and acquired through training, while the interviewee's role and behaviors must be observed by the interviewer, who seeks to modify them as necessary to successfully complete the interview.

To oversimplify, the role of the interviewer is to ask the questions, while that of the interviewee is to respond with answers. Actually, the interviewer must perform at least eight major tasks:

- develop rapport with the interviewee and show interest,
- give the interviewee a reason to participate,
- elicit responsiveness from the interviewee,
- ask questions in a prescribed order and manner,
- ensure understanding,
- ensure nonbias,
- obtain sufficient answers, and
- show sensitivity to the interviewee's burden.

These tasks, which are not isolated but must be integrated into the interview procedure, are discussed below from the viewpoint of the interviewer and his or her responsibilities.

Developing Rapport and Showing Interest

Seek to establish a balanced relationship between the interviewee and yourself as an empathetic, friendly individual who is not too different from the interviewee but who is also an independent, unbiased, and honest collector of data. Your appearance, verbal mannerisms, body language, and voice will determine the rapport, starting with the contact that sets up the interview. Since this is usually done by telephone, your voice and verbal mannerisms are

extremely important (as they are later in the interview setting). Of course, for CATIs, voice and verbal mannerisms are key factors in the success of the interview. Remember that this interview is an artificial situation in which nonverbal cues are not available.

Make your verbal and voice cues calm and unflustered. Speak so the interviewee need not strain to hear and understand. Changes in voice inflection, sighs, or other noises give clues to your feelings or moods, as do your facial expressions and body language. Control these so that the interviewee does not pick up impatience, disapproval, or other negative feelings. Ideally, you should not experience such feelings during the interview, since you are supposed to be an impartial, unbiased, and tolerant observer. Likewise, you should control expressions of positive feelings or agreement with what the interviewee is saying.

It is important that the interviewer be aware of characteristic nonlinguistic cues such as change in voice, facial expressions, or gestures, since as much as half of the communication that takes place during the interview is conveyed by these modes of expression. Failure to understand these cues may result in miscommunication. More details on understanding and using these cues are presented in the methodology transfer paper entitled Designing and Using Questionnaires.

Your appearance is still another variable that influences rapport and, therefore, the tone of the interview. Dress to fit both the interview and the interviewee. If the interview is with a state welfare official in his office in the capitol, it is appropriate, perhaps mandatory, to wear office-type clothing (suit and tie for men, and suit or dress for women). This is what you would expect the interviewee would be wearing. Try to live up to the expected standards of the interviewee in this case. Not doing so might get the interview off to a bad start.

If, however, the interview is to take place at a construction site or with young people at a summer youth-recreation site, wear more casual clothing. This makes sense in that it gives interviewees the feeling that you understand the nature of their circumstances. Also, you are not set off as being totally different from the interviewee.

Giving the Interviewee a Reason to Participate

Generally, interviewees do not benefit directly from the information that they give to GAO. Why, then, should they agree to give you their time for an interview? The reasons are various. Some interviewees are obliged to cooperate with GAO because of their positions and to provide information on how federal money is being spent. Such individuals usually understand why they should participate and need only be told something about the evaluation procedures. In other cases, where interviewees are not operating some part of a federal program but are the recipients of funds, such as program beneficiaries and contractors, greater explanation may be required.

Interviewees who are not aware of the importance of the evaluation and how they can help may not give sincere and well-thought-out answers. Your explanations to them, therefore, are important to the validity of the resulting data.

Helping the Interviewee to Be Responsive

Many people you may contact, especially program beneficiaries, have never before been interviewed during an evaluation or audit. They may have had job interviews and interviews prior to receiving benefits, where they have given name, address, age, number of children, work experience, and the like. But generally they have not been asked for their opinions and feelings.

Thus, the individual may need to learn how to act as an interviewee. The interviewer should help in this process, and while this should not include hints

on how questions should be answered, it does involve making the interviewee comfortable and capable as a respondent. For example, you will impart information that helps the interviewee learn to use an answer format that has been programmed into the structured interview. Where responses form a closed set, the interviewee must know how to choose from the alternatives given.

For CATIs, the development of the questions must be sensitive to helping the interviewee. The success of the interview is very sensitive to question wording and ensuring that the interview takes on a conversational tone. For a CATI, there is a critical tradeoff between formality (with longer, more difficult questions) and conversation (with shorter and easier questions).

During the interview, it may help to reinforce the interviewee with such verbal cues as

- “Un-huh, I see.”
- “Let me get that down.”
- “I see” (repeat answer).
- “I want to make sure I have that right.”
- “It’s useful to get your ideas on this.”
- “Thanks, it’s important to get your opinion on that.”
- “I see, that’s helpful to know.”

Asking Questions in a Prescribed Order and Manner

The order in which the questions appear in the structured interview is not accidental. Questions are ordered so as to lead the interviewee through various topics, correctly position sensitive questions, and hold the interviewee’s interest. To the greatest extent possible, you must maintain this order. The words and phrasing used in the questions also have been carefully chosen and tested. For the sake of standardization and understandability, it is important that these be used as planned. For CATIs, this is made simpler, since the order is preprogrammed.

The following tips may help:

- ask the questions exactly as they are worded in the questionnaire,
- ask the questions in the order in which they are presented in the questionnaire,
- ask every question specified in the questionnaire,
- read each question slowly (two words per second),
- repeat questions that are misunderstood or misinterpreted,
- do not let the respondent stray from the questions in the interview, and
- keep nonverbal cues as neutral as possible.

Remember that for telephone interviews, the lack of visual contact decreases the ability to make the interviewee understand.

Ensuring Understanding

At times, an interviewee will not understand a question, as indicated either by telling the interviewer so, by not answering, or by providing an answer that seems inconsistent or wrong. When this happens, you should use an appropriate probing technique such as the following:

- repeat the question;
- give an expectant pause;
- repeat the respondent's reply; and
- make neutral questions or comments, such as "Anything else?" "Any other reason?" "Any others?" "How do you mean?" "Could you tell me more about your thinking on that?" "Would you tell me what you have in mind?" "What do you mean?" "Why do you feel that way?" "Which would be closer to the way you feel?"

To maintain the meaning of the questions and not to bias them, do this "probing" with care. These kinds of probes should be worked out during the pretest. Rephrasing the question or adding new questions should be avoided. If all probes have been tried and

rephrasing or adding questions is the only alternative, notes to that effect should be added next to the responses.

Ensuring Nonbias

In earlier chapters, we covered bias in the way a question is written or in the selection of interviewees. There can be bias also in the way you pose the contents of the query, in the introduction of your own ideas into a probe, or in your adding certain verbal emphasis or using certain body language. All these can destroy the neutrality that should characterize your presentation. When listening to the interviewee's answer, you can filter out portions of the message that alter the true response.

Obtaining Sufficient Answers

You must learn to judge when an answer is sufficient before going to the next question. If the answer is incomplete or vague, you should ensure that the question is understood (as discussed above) or draw more out of the interviewee to complete the answer. At times, the interviewee is allowed to answer questions in an open-ended fashion, while you match each answer to one of a set of responses on the interview form. You must be sure that the interviewee has sufficient information to select one of the answers. Sometimes, you must select two or more responses (not just one) from the set and ask the interviewee which one best matches his or her answer. This should be done, however, only as a last resort and only after giving the respondent ample time to respond.

On other occasions, an interviewee may not have the answer in mind but may need to refer to documents or ask someone else. If this can be done conveniently and within a short time, encourage the interviewee to do so.

You can also check the accuracy of the answers given by asking for supporting information from

the interviewee. Sometimes the design of the instrument has built into it questions to which answers have already been obtained from files or from other people in advance. Use these to check the accuracy with which the interviewee is answering. Underreporting of information is often found. As the length of time since a subject event increases, there is a greater tendency for the interviewee either to forget the event occurred or to recall it only partially.

**Showing
Sensitivity to
Interviewee
Burden**

Before conducting an interview, give the interviewee a general statement of how long it is expected to take. You are then under some obligation to adhere to this time limitation.

Frequently, interviewees prolong their answers by adding examples, critical incidents, or other stories. If neither you nor the interviewee has a time problem, this extension of the interview is acceptable. If time is critical, however, use techniques to speed up the interview so as not to lose valuable answers at the end. Besides the length of time taken, the interview can be burdensome because of the amount of work the interviewee needs to go through to produce the information requested. If a relatively unimportant question requires a significant amount of time or energy by the interviewee, it may not be worth pursuing.

Analyzing the Data

The purpose of all the work that has gone into designing, pretesting, and revising the structured interview (perhaps many times), obtaining expert review, and finally using it is to obtain data that, when properly analyzed, will answer the evaluation or audit questions. If you have followed the procedures outlined in previous chapters, chances are great that you now have those data. All the hard work and time will have been worth the cost.

You now have uniform data that can be used to answer the evaluation or audit questions. The questions in your structured interviews were as clear and precise as the state of the art permits. Your interviewers were carefully trained and instructed as to what explanations were to be given when interviewees did not understand or had trouble with questions. The pretests you conducted and the expert review you obtained ensured that the people you interviewed could give you the data you need.

You need not worry about analyzing narrative responses to a long list of open-ended questions, because your interviews contained few, if any, of them. Through your preliminary research, your interviews with program officials and outside experts, and your pretests, you identified most of the possible replies to your questions. Thus, you were able to convert what started out as open-ended questions to closed-format questions with sets of alternative responses that minimized the use of "other, please specify" responses.

Transferring the data from the completed interview forms to computer files is comparatively simple if you have used keypunch boxes. You can have the data keyed directly to disks or magnetic tape and then entered directly into computer files. (Of course, for CATIs, the data base will have been generated as you went along.) After verifying the accuracy of the keypunching, you are almost ready to begin the analysis. But before you do this, you must determine if you have a nonrespondent problem.

Nonrespondent Problem

When you draw a sample of persons from a universe to conduct interviews, you intend that all the sample be interviewed. Indeed, it is part of the GAO analysis plan that what these selected persons say will stand for what the entire universe would have said if all could have been queried. Rarely can the entire sample be interviewed, however, because of deaths, inability to locate people, refusals to be interviewed, and so on. For telephone interviews, the nonrespondent problem is usually significantly less, since they frequently have 80- to 90-percent response rates.

To combat the nonrespondent problem, normally you will randomly sample a greater number of people than is statistically required. Nonrespondents can be replaced by randomly drawn substitutes. For example, if the sampling plan calls for 50 persons to be interviewed, you might randomly select 75 names. If the 8th, 20th, 31st, and 49th individuals you try to contact have died, you would use cases 51 through 54 as substitutes.

Usually, making a small number of substitutions has no effect on analysis of the final data. When a larger number of substitutions is made, for example 20 percent or more, you may have some concern that the people you were unable to interview represent a unique portion of the universe. For example, if all these persons died, they may represent the older persons in the universe; your data collection therefore would not adequately represent the opinions of older individuals. If queried, this portion might have given dramatically different answers to all or some of the questions and altered the final results of your data collection.

There are several ways to assure yourself that the data would not have changed much had these individuals been contacted: analyzing the reason for the

nonparticipation, interviewing by telephone a subsample on critical questions, comparing demographic information, and assuming a “worst case” answer. These are discussed below.

Analyzing Reasons for Nonparticipation

When you are trying to contact individuals to set up interviews, you have the opportunity to talk to them or someone who has information about them. If the potential interviewee could not be contacted, record in your log the reason given for not being able to set up the interview: for example, death, moved out of the area, whereabouts unknown, or apparently at the location but unable to be contacted.

The potential interviewee, when reached, may decline to be interviewed, giving such reasons as “too busy,” “I don’t give interviews,” “it’s none of your business,” or “I don’t understand why you want to talk to me—I never participated in that.”

If you have no data other than this, you may make some attempt to determine whether the reasons given for nonparticipation are related to critical questions in the interview. For example, if you are relating social services received to the recipients’ state of well-being and have missed many interviews because of potential interviewees’ deaths, this could mean loss to your sample of many interviewees who would have reported poor well-being. You might then have to place some limitation on the final conclusions of the study. There is no statistical test of the excuse data that can be used to make this decision.

Interviewing a Subsample on Critical Questions

A second approach to the nonrespondent problem is to select a subsample of those not available for an interview (or the entire group, if it is small enough) and conduct a short phone survey of them, using some of the critical questions on the instrument. Of course, this does not help if the people could not be

located in the first place or were deceased. If most, however, were found but at first refused an interview because of time considerations, you may be able to collect data on some questions on the phone. The answers are then compared to those collected in the normal interviewing process, using statistical procedures to test for significant differences. Questions on which the two groups differ significantly might then be eliminated from the final analysis. There are some exceptions to this approach, as discussed below under "worst-case assumption."

Comparing Demographic Information

Many times, we have a rich data base on a collection of demographic variables for all potential interviewees. For example, the program file information for welfare recipients may contain information on their sex, age, race, education, marital status, number of children, and work experience. Thus, if you cannot obtain partial interview information from a subsample, as discussed above, you can compare the demographic variables for those interviewed and those not.

Significant differences on a certain proportion of critical demographic variables would cast doubt that the two groups were essentially the same and indicate that the absence of these individuals could alter the overall results.

Assuming the Worst Case

Some of the questions you pose will have binary-choice answers. Your task may be simply to determine whether more people have done something than have not. Suppose you are to interview a random sample of 100 people, but 20 people could not be contacted and were not replaced by the next 20 people on your random list. For the 80 you do interview, your results show that 47 said "yes" and 33 said "no." You are now asked about the 20 you were unable to interview. Could they have changed the outcome?

Taking a conservative view, you could attribute all their responses to one of the categories. If a majority of “yes” votes would allow you to defend a particular finding, then you would want to make the assumption that all 20 would have answered “no.” This would make the final outcome 47 “yes” and 53 “no.” Under these circumstances, the finding would not have support. In other words, a 47-to-33 split of the data with 20 uninterviewed persons is too close to make a decision. Had the split been something like 55 to 25, your case would have stood a chance.

Some of these methods require the collection of additional information. With each method, assumptions and limitations can influence the eventual interpretation of the data collected during the structured interview. Nothing short of obtaining the interviewees’ answers to the questions will be fully satisfactory.

Data Analysis

The edited data now reside in computer files and you have dealt with the nonresponse problem, if any. You can now begin the data-analysis phase. This will probably be the most enjoyable part of the job, as you will begin to see results and imagine how the report will read. Yet this phase will not be easy and will require your full attention.

At this point, it is useful to address the question, What do you mean by data analysis? In GAO, data analysis carries with it various meanings, ranging from such simple tasks as learning how many members of a surveyed population are 25 years of age or older and how many are under 25 (first-level analysis) to investigating causal relationships between the different achievement levels (if any) of only children, children with only older siblings, children with only younger siblings, and children with both older and younger siblings (third-level analysis).

The analysis to be done will be determined to a great degree by the project objectives that have been established for the structured interview. You might consider three levels of analysis.

First-Level Analysis

In first-level analysis, you concentrate on a description of the data—for example, how many responded to each response alternative, both in absolute numbers and on a percentage basis. For example, a question may have asked, “Did you complete high school?” A description of the data would show how many and what percentage responded “yes” and how many “no.”

In the language of analysis, this type of description of the data is commonly referred to as frequency tabulations or frequency tables. Although not the only analytic activity under this first-level analysis, it is normally the most significant activity.

You may often make a computer run to obtain frequency tabulations during the data-verification phase, because it will show all values keypunched for every question. A review of the run will disclose possible errors in the data base. In the example above, “yes” answers may be coded as “1” and “no” answers as “2.” Any other number showing up for this question would stem from an error on the part of the interviewer or the keypuncher.

Second-Level Analysis

Second-level analysis begins where the description of the data stopped. In this next level of analysis, perhaps the most useful to most GAO efforts, you first analyze the data, one question at a time. Certain statistics, such as the mean and median, can be obtained with the description of the data for questions where such statistics would be useful or appropriate. Remember that if a sample other than a simple random sample has been used, the number and percentages shown on the frequency tabulations’ run must be weighted

before you make projections. Therefore, it would be wise to consult a sampling statistician before using the numbers in a briefing or report.

Having completed the single-question analysis, you then move to testing the associations between pairs of questions in response to hypotheses established during the design phase. For example, is there an association between a person's sex and whether or not the person completed college? If the data show that a larger percentage of women complete college than do men, is the difference statistically significant, or could it stem from the fact that we studied a sample and not the entire population? Such statistical measures as chi-square analysis and correlation analysis are often used to determine how certain we can be that apparent associations between responses to two questions do not stem from chance. On many GAO assignments, second-level analysis is as far as the analysis of questionnaire or interview data goes.

Third-Level Analyses

Third-level analyses are more complex than the other levels of analysis. They normally take into account many variables at one time and address more complex questions. Third-level analyses often address differences between subgroups of surveyed cases—what factors differentiate students who repay federal loans in a timely manner from those who do not—or investigate the influence that a set of factors may have on a single variable—what factors influence the amount of loans made by the Small Business Administration.

Two of the many analytic tools available to investigate these more complex analytic questions are multiple regression analysis and discriminant function analysis.

It is not our intent here to provide a detailed account of the analytic tools available in survey research. We do, however, want you to understand

that the nature and complexity of the analysis phase of a project can vary dramatically, depending primarily upon the objective established for the project. The analysis that addresses cause-and-effect questions will be much more difficult than the analysis for descriptive or normative questions. Regardless of the type of question being addressed, a large number of statistical tools are available for the analysis phase of the evaluation. Selecting the most appropriate is not easy. Once again, we strongly advise that evaluators and auditors obtain the assistance of a specialist from the design, methodology, and technical assistance group, the regional technical assistance group, or PEMD for this phase of the assignment.

Analysis of Open-Ended Questions

Answers to open-ended questions may range from a few words to several sentences. Interviewees typically give the interviewer some salient ideas that come quickly to mind but leave out some important factors. Open-ended questions do not help interviewees consider an identical range of factors. After conducting several interviews, interviewers may supplement the question by asking the interviewee about factors not mentioned, but such supplemental questions will not be standard among interviewers. Thus, the interviewees as a group are not responding to identical questions.

As mentioned briefly in chapter 3, the proper analysis of open-ended questions requires the use of a complicated, time-consuming process called "content analysis." In brief, you must read and reread a substantial number of the written responses, come up with some scheme to categorize the answers (in essence, develop a set of alternative responses), and develop rules for assigning responses to the categories. Even with a set of rules, people can categorize answers differently. Therefore, three or four people must go through each completed interview and categorize the answers. A majority of them must agree if you are to have a reliable data base.

Because content analysis is so time-consuming, the answers to open-ended questions are often left unanalyzed. The evaluator or auditor in reporting may quote from one or a few selected responses, but open-ended interviews generally do not produce uniform data that can be compared, summed, or further analyzed to answer the evaluation or audit questions.

The Role of Evaluators and Specialists on Each Task

This paper has discussed the major tasks that must be performed to collect data by structured interview. Some of these tasks are short and seem relatively uncomplicated, such as identifying the target population and selecting variables from the variable pool. Other tasks appear lengthy and rather complex, such as designing the interview form and conducting the interview.

All these necessary tasks are carried out in a cooperative effort by evaluators and specialists. Generally, three types of technical expertise are involved: measurement, sampling, and data analysis. The expertise required need not, however, come from three separate individuals. Most specialists are capable of handling more than one function, depending on the complexity of the job. Examples of functions each staff member might perform during the evaluation appear in table 10.1. In addition to data-collection tasks 1 to 17, which are discussed in this paper, the table includes for completeness seven tasks that concern data handling, analysis, interpretation, and reporting.

**Chapter 10
The Role of Evaluators and
Specialists on Each Task**

Table 10.1: Functions and Responsibilities of Evaluators and Specialists

Task	Staff function^a			
	Evaluator in charge and staff	Measurement specialist	Sampling specialist	Data analyst
1. Formulate overall question	Initial work	Acquire job understanding and review	Optional	Optional
2. Determine kind of information needed	Initial work	Review	Optional	Optional
3. Identify the target population	Initial work	Review	Optional	Optional
4. Create question pool	Contribute	Contribute	Optional	Discuss analysis options and problems
5. Select questions from the pool	Check for relevance	Primary role	Optional	Review
6. Decide on final data collection method	Check for job-related constraints	Outline advantages and disadvantages; recommend solutions	Design sampling plan	Review
7. Plan data analysis	Participate	Participate	Participate	Participate
8. Design interview form	Assist and review	Primary role	Optional	Optional
9. Obtain subject matter review	Lead	Participate	—	—
10. Conduct pretest	Participate	Lead	—	—
11. Review of expert review and pretest	Participate	Lead review	Optional	Review
12. Revise interview form	Review	Write	Optional	Consider coding

(continued)

Chapter 10
The Role of Evaluators and
Specialists on Each Task

Task	Staff function ^a			
	Evaluator in charge and staff	Measurement specialist	Sampling specialist	Data analyst
13. Kick-off conference	Give information on how instrument fulfills job needs	Give instructions on question meaning and instrument use	Give instructions on sampling	Optional (describe how data will be used)
14. Train interviewers to interview	Assist in training	Train using various techniques	—	—
15. Select interviewees	Staff uses sampling plan	Available for consultation	Available for consultation	—
16. Contact interviewees	Staff contact	Available for consultation	Available for consultation	—
17. Conduct interviews	Staff conduct	Available for consultation	—	—
18. Settle nonrespondent problems	Acquire needed data	Recommend solutions	Recommend solutions	Participate
19. Edit raw data	Staff edit	Consult	—	Direct
20. Key punch	—	—	—	Contract arrangement
21. Edit key punch data	Staff edit	—	—	Direct
22. Analyze data	Consult	Consult	Consult	Analyze
23. Interpret data	Participate	Participate	Participate	Participate
24. Draft report	Write	Write, consult, review	Write, consult	Write, consult

^aA staff member may perform more than one function during a given task in the course of an evaluation.

This table can serve as a guide when reaching an agreement about responsibilities when assembling personnel at various times during the development and use of the DCI. Staff responsible for a given function may not be required during some tasks, yet their presence could prove useful. In other cases, a large number of staff might be distracting, making it better to work with fewer persons for that task.

Bibliography

Babbie, E. H. Survey Research Methods. Belmont, Calif.: Wadsworth, 1973.

Bradburn, N. M., and S. Sudman and Associates. Improving Interview Methods and Questionnaire Design. San Francisco: Jossey-Bass, 1981.

Cannell, C., et al. New Techniques for Pretesting Survey Questions. Ann Arbor, Mich.: Survey Research Center, University of Michigan, 1989.

Dawes, R. M. Fundamentals of Attitude Measurement. New York: John Wiley and Sons, 1972.

Dillman, D. Mail and Telephone Surveys. New York: John Wiley and Sons, 1978.

Fowler, F. J. Survey Research Methods. Newbury Park, Calif.: Sage Publications, 1988.

Fowler, F. J., Jr., and T. W. Mangione. Standardized Survey Interviewing: Minimizing Interviewer-Related Error. Newbury Park, Calif.: Sage Publications, 1989.

Frey, J. H. Survey Research by Telephone, 2nd ed. Newbury Park, Calif.: Sage Publications, 1989.

Groves, R. M. Survey Errors and Survey Costs. New York: John Wiley and Sons, 1989.

Groves, R. M., et al. Telephone Survey Methodology. New York: John Wiley and Sons, 1988.

Heverson, M. E., L. L. Morris, and C. T. Fitz-Gibbon. How to Measure Attitudes. Beverly Hills, Calif.: Sage Publications, 1978.

Lavrakas, P. J. Telephone Survey Methods: Sampling Selection and Supervision. Newbury Park, Calif.: Sage Publications, 1987.

Bibliography

Sudman, S., and N. M. Bradburn. Asking Questions: A Practical Guide to Questionnaire Design. San Francisco: Jossey-Bass, 1982.

U.S. General Accounting Office. Content Analysis: A Methodology for Structuring and Analyzing Written Material, methodology transfer paper 3. Washington, D.C.: 1982.

U.S. General Accounting Office. Designing Evaluations, methodology transfer paper 10.1.4. Washington, D.C.: 1991.

U.S. General Accounting Office. Developing and Using Questionnaires, methodology transfer paper 7. Washington, D.C.: 1986.

U.S. General Accounting Office. Using Statistical Sampling, methodology transfer paper 6. Washington, D.C.: 1986.

U.S. National Center for Health Statistics. Data Evaluation and Methods Research: A Summary of Studies of Interviewing Methodology. Vital and Health Statistics, series 2, number 69. Rockville, Md.: 1977.

Warwick, D. P., and C. A. Lininger. The Sample Survey: Theory and Practice. New York: McGraw-Hill, 1975.

Glossary

Bias	Words, sentence structure, attitudes, and mannerisms that unfairly influence a respondent's answer to a question. Bias in questionnaire data can stem from a variety of other factors, including the sequence of questions. Both interviewer and instrument bias can exist.
Closed Question	A question with more than one possible answer from which one or more answers must be selected.
Content Analysis	A set of procedures for collecting and organizing nonstructured information into a standardized format that allows one to make inferences about the characteristics and meaning of written and otherwise recorded material.
Cues	The alternative responses to questions that increase or decrease in intensity in an ordered fashion. The interviewee is asked to select one answer to the question.
Data-Collection Instrument	A highly structured document, sometimes abbreviated DCI, that requires the user or respondent to collect or provide data in a systematic and highly precise fashion.
Demographic Question	A question used in compiling vital background and social statistics, such as age, marital status, and size of household.
Open-Ended Question	A question that does not have a set of possible answers from which to make a selection but permits the respondent to answer in essay form. On a questionnaire, the respondent would write an essay or short answer or fill in a blank. During an interview, the respondent would give the interviewer an unstructured, narrative answer. The interviewer

would record the response verbatim or select salient features. If a structured interview were used, a question might appear to be open-ended to the interviewee but could be “closed down” by the interviewer, who would have a set of alternative answers to check.

Probe

To examine a subject in an interview in depth, using several questions.

Qualitative Analysis

An analysis that ascertains the nature of the attributes, behavior, or opinions of the entity being measured. In describing a person, a qualitative analysis might conclude that the person is tall, thin, and middle-aged. See also Quantitative Analysis.

Quantitative Analysis

An analysis that ascertains the magnitude, amount, or size, for example, of the attributes, behavior, or opinions of the entity being measured. In describing a person, a quantitative analysis might conclude that the person is 6 feet 4 inches tall, weighs 165 pounds, and is 45 years old. See also Qualitative Analysis.

Reliability

The extent to which a measurement process produces similar results on repeated observations of the same condition or event.

Stem

The statement portion of a question.

Structured Interview

An interview in which questions to be asked, their sequence, and the detailed information to be gathered are all predetermined; used where maximum consistency across interviews and interviewees is needed.

Target Population

The level (item, individual, group, organization, or the like) at which data are collected. Data can be collected at the individual level (for example, program participant) and analyzed and reported at the organizational level (for example, the employment office). Participants might be asked how many hours of counseling they received from the employment office, whereas the analysis of the data might reflect the number of offices that provided given amounts of counseling.

Validity

The extent to which the question being asked measures the concept that the evaluator wants it to measure.

Papers in This Series

This is a flexible series continually being added to and updated. The interested reader should inquire about the possibility of additional papers in the series.

The Evaluation Synthesis. Transfer paper 10.1.2, formerly methods paper 1.

Content Analysis: A Methodology for Structuring and Analyzing Written Material. Transfer paper 10.1.3, formerly methodology transfer paper 3.

Designing Evaluations. Transfer paper 10.1.4, formerly methodology transfer paper 4.

Using Structured Interviewing Techniques. Transfer paper 10.1.5, formerly methodology transfer paper 5.

Using Statistical Sampling. Transfer paper 10.1.6, formerly methodology transfer paper 6.

Developing and Using Questionnaires. Transfer paper 10.1.7, formerly methodology transfer paper 7.

Case Study Evaluations. Transfer paper 10.1.9, formerly methodology transfer paper 9.

Prospective Evaluation Methods: The Prospective Evaluation Synthesis. Transfer paper 10.1.10, formerly methodology transfer paper 10.

Ordering Information

The first copy of each GAO report and testimony is free. Additional copies are \$2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. VISA and MasterCard credit cards are accepted, also. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:

**U.S. General Accounting Office
P.O. Box 6015
Gaithersburg, MD 20884-6015**

or visit:

**Room 1100
700 4th St. NW (corner of 4th & G Sts. NW)
U.S. General Accounting Office
Washington, DC**

**Orders may also be placed by calling
(202) 512-6000 or by using fax number
(301) 258-4066, or TDD (301) 413-0006.**

Each day, GAO issues a list of newly available reports and testimony. To receive facsimile copies of the daily list or any list from the past 30 days, please call (202) 512-6000 using a touchtone phone. A recorded menu will provide information on how to obtain these lists.

For information on how to access GAO reports on the INTERNET, send an e-mail message with "info" in the body to:

info@www.gao.gov

**United States
General Accounting Office
Washington, D.C. 20548-0001**

**Official Business
Penalty for Private Use \$300**

Address Correction Requested

<p>Bulk Rate Postage & Fees Paid GAO Permit No. G100</p>
