'5181a

Using Structured Interviewing Techniques

Methodology Transfer Paper 5

GAO
United States General Accounting Office

Program Evaluation and Methodology Division
July 1985

Request for copies of GAO reports should be sent to:

U.S. General Accounting Office
Document Handling and Information
Services Facility
P.O. Box 6015
Gaithersburg, Md. 20877

Telephone (202) 275-6241

The first five copies of individual reports are free of charge. Additional copies of bound audit reports are $3.25 each. Additional copies of unbound report (i.e., letter reports) and most other publications are $1.00 each. There will be a 25% discount on all orders for 100 or more copies mailed to a single address. Sales orders must be prepaid on a cash, check, or money order basis. Check should be made out to the "Superintendent of Documents".

Using Structured Interviewing Techniques

Methodology Transfer Paper 5

Program Evaluation and Methodology Division

July 1985

PREFACE

Structured interviewing is a technique that can aid General Accounting Office (GAO) evaluators in gathering data needed to conduct a GAO evaluation or audit. The purpose of this paper is to help an evaluator

- recognize when structured interviews may be useful to his/her assignments,
- prepare workable first drafts of structured interviews, and
- understand and practice the basic procedures necessary to successfully complete the interviews.

This paper also provides a basic reference for GAO evaluators whose assignments require them to review structured interviews used by government agencies or their contractors. It is not, however, a comprehensive textbook that would train GAO evaluators to completely plan, write, and execute a structured interview.

Thus, before GAO project staff decide to use, draft, or critique a structured interviewing instrument, we recommend that they request assistance from the Design, Methodology, and Technical Assistance Group (DMTAG) in the GAO division that is programming the assignment or from the measurement assistance staff in GAO's Program Evaluation and Methodology Division (PEMD).

GAO policy on structured interviews is set forth in the General Policy Manual, pages 7-18 and 7-19.

Chapter 1 of this paper discusses how structured interviews are used in GAO evaluations. Chapter 2 defines structured interviews and gives some guidelines as to when they should be used. Subsequent chapters cover

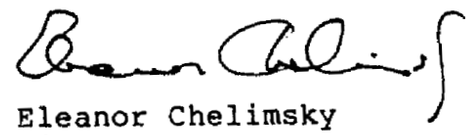
- designing a structured interview (chapter 3),
- some interview flaws and how to correct them (chapter 4),
- pretesting the instrument and obtaining expert review (chapter 5),
- training interviewers (chapter 6),
- selecting and contacting interviewees (chapter 7),
- conducting a structured interview (chapter 8), and
- analyzing the data, including dealing with nonresponse problems (chapter 9).

In the final chapter (10), we list the roles of the GAO evaluator and the measurement specialist (from the DMTAG or PEMD) in each of the tasks necessary to gather information through a structured interview.

Using Structured Interviewing Techniques is one of a series of methodology transfer papers issued by PEMD to give GAO evaluators handy and comprehensive guides to various aspects of evaluation methodology and explain specific applications and procedures. Other papers in the series include:

- Causal Analysis: A Method to Identify and Test Cause and Effect Relationships in Program Evaluations, February 1982,
- Content Analysis: A Methodology for Structuring and Analyzing Written Material, June 1982, and
- Designing Evaluations, July 1984.

The author of Using Structured Interviewing Techniques is Erwin W. Bedarf, formerly of PEMD, now with the Human Resources Division. Readers of this paper are encouraged to send questions or comments on it to me or to Brian Keenan, Principal Survey Methodologist, PEMD.



Eleanor Chelimsky
Director
Program Evaluation and
Methodology Division

Contents

ABBREVIATIONS

DCI	Data-collection instrument
DMTAG	Design Methodology and Technical Assistance Group
EIC	Evaluator in Charge
GAO	General Accounting Office
PEMD	Program Evaluation and Methodology Division, GAO
WIN	Work Incentive Program

CHAPTER 1

THE ROLE OF STRUCTURED INTERVIEWS IN GAO EVALUATIONS

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

The broad questions that dictate the objectives of a GAO evaluation[1] and suggest the evaluation strategy can be categorized as descriptive, normative, or cause-and-effect.[2] 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. A 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 follows:

Evaluation strategy	Type of evaluation question most commonly addressed
Sample survey	Descriptive and normative
Case study	Descriptive and normative
Field experiment	Cause-and-effect
Available data	Descriptive, normative, and cause-and-effect

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 available to GAO.

---

[1]We use the term "evaluation" throughout this paper; however, many of the interviewing concepts and procedures apply equally to GAO audits.

[2]Discussed more fully in Designing Evaluations, PEMD Methodology Transfer Paper, July 1984.

The design of a GAO evaluation encompasses seven elements:

1. Kind of information to be acquired,

2. Sources of information (e.g., types of respondents),

3. Methods to be used for sampling sources (e.g., random sampling),

4. Methods of collecting information (e.g., structured interviews, self-administered questionnaires),

5. Timing and frequency of information collection,

6. Basis for comparing outcomes with and without a program (for cause-and-effect questions), and

7. Analysis plan.

This paper focuses on the fourth design element, specifically structured interviews[3]. 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 or a single case-study evaluation, it may be beneficial to use structured interviews.

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 a PEMD methodology transfer paper expected to be issued later in 1985.

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 content or subject matter of the interview form. Any constraints in identifying and selecting a sample (e.g., lack of a universe listing of the target population) may make it necessary to refine

---

[3]This and other technical terms used in this paper are defined in the Glossary (appendix I).

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.

CHAPTER 2

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 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 respondent the same set of possible responses.

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, but we discuss some of the advantages and disadvantages of interviews and questionnaires in the following paragraphs. In addition, the characteristics of various data-collection techniques are systematically compared in table 1.

In the planning and scoping phase of an evaluation or in a one-of-a-kind interview during the implementation phase of an evaluation, the less structured, less guided type of interview may be more useful.

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

The telephone interview and, even more so, the face-to-face interview enable the interviewer to establish rapport with the respondents. Individuals who would ignore mail questionnaires entirely or not answer certain questions on them can be persuaded to provide truthful answers in a telephone or face-to-face

Table 1

Comparison of Data-Collection Techniques by Characteristics or Advantages

Characteristics or advantages	Extent of advantage: Structured interview: By telephone	Structured interview: Face-to-face	Questionnaires: By mail	Questionnaires: Group	Questionnaires: Audit of records
Methodological					
Allows use of probes	3	5	1	2	n/a
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	n/a
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	n/a
Content					
Allows inclusion of most relevant variables	3	5	4	4	3
Allows complex subject matter to be presented or derived	3	5	3	4	4
Allows collection or real-time data	5	5	4	5	3
Allows 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/getting sample	3	2	4	4	5
Allow use with large sample	4	3	5	4	5
Allows identity of source to be known	4	5	3	5	3
Reduces problems due to respondent's illiteracy	4	5	1	3	n/a

Table 1 (continued)

Characteristics or advantages	Extent of advantage: Structured interview: By telephone	Structured interview: Face-to-face	Questionnaires: By mail	Questionnaires: Group	Questionnaires: Audit of records
Time/cost/resources					
Minimizes instrument-development time	2	3	1	1	5
Minimizes instrument-development cost	3	1	1	1	5
Minimizes number of field staff	5	?	5	?	?
Minimizes travel by staff	5	?	5	?	?
Minimizes staff training	2	1	5	3	5
Minimizes time required to carry out activities	?	?	3	?	?
Overall cost low	3	1	5	4	1
Results/response/quality of data					
Maximizes rate of return of data after source is contacted	4	5	3	5	n/a
Minimizes multiple contacts of sources	2	2	3	4	n/a
Minimizes follow-up after initial response	5	5	3	4	5
Increases chance source will be accurate	4	4	4	4	3
Allows reliability to be checked	5	5	3	4	4
Allows validity to be checked	4	4	2	4	5
Facilitates recall of data by source	4	5	3	4	n/a

Key:

1 Little or no extent
2 Some extent
3 Moderate extent
4 Great extent
5 Very great extent

? Depends greatly upon study specification
n/a Not applicable

interview. Also, a well trained interviewer can recognize when a respondent is having trouble with a question and rephrase it diplomatically and nonintrusively.

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 respondent'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 thirty minutes is the usual limit for telephone interviews, while face-to-face interviews can last up to an hour.

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, 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 questionnaires can be longer and include more complex questions (given that 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 can be used only in situations where 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

- the self-reported experience of Work Incentive (WIN) participants while in and after leaving the program,
- the experience of participants trained by Opportunities Industrialization Centers (OIC), and
- the opinions of Drug Enforcement Agency (DEA) agents concerning various DEA 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 DEA agents' 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.

Structured telephone interviews were used by GAO to study such topics as

- the satisfaction of small businesses with management assistance provided by Small Business Administration contractors and

- the satisfaction of 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 (FCC) in regard to a proposed move from the D.C. area to Pennsylvania.

In both examples, it was practical to assemble the respondents. Also, in the FCC 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; about 80 to 90 percent of all such DCI's are mail questionnaires. We have successfully used them with such diverse groups as executives of Fortune magazine's top 1,000 companies, farmers who received federal loans, Veterans Administration and military doctors, homebuyers, audited taxpayers, federal park rangers, and emergency preparedness teams. Questionnaires are discussed in a PEMD methodology transfer paper to be published in 1985.

Additional discussion of structured interviews, questionnaires, and other DCIs, with examples of GAO applications, appears on pages 11-6 through 11-10 of the GAO Project Manual.

CHAPTER 3

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 insure 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 again suggest, therefore, that you work with measurement specialists from the DMTAG in the Division programming the assignment or from PEMD when you are planning to use a structured interview.

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 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:

- 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 not find jobs?"

The type of question(s) asked will dictate the evaluation stategy. Also, certain strategies are more appropriate to answering certain questions.[1] 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

---

[1]Formulating overall evaluation questions and selecting evaluation strategies that provide answers is discussed in Designing Evaluations, PEMD Methodology Transfer Paper, July 1984.

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. To formulate the questions, the target population must be identified. The target population is the source level (individuals, groups, organization) at which the information is to be gathered. Thus, in the study of how program participants found jobs after leaving the program, the target population is the individual participants of the program who were trained.[2]

Next, develop a pool of questions that attempt to measure the variables under consideration, such as age. The questions may include various ways of measuring the same variable. For example, 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 2.

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.

- Relevancy. Questions should be relevant to the study being conducted and 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, which can create an unnecessary burden on the interviewee and distract attention from the central purpose(s) of the interview.

- Selection of respondent. 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 inappropriate.

- Ease of response. Interviews are meant to obtain data that may otherwise not be documented or, if documented, need some

---

[2]Later in the evaluation, 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.

Table 2

Process of Identifying, Developing, and Selecting Questions: Applied to Example

Task	Example
Formulate overall questions	How do program participants find jobs after leaving the XYZ program?
Determine the kind of information needed	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 ..... n.
Identify target population	Program participants who have left the program (random sample)
Create a question pool	1.1 How did you look for jobs. Did you 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? n. 1.2 About how many jobs that you were interested in did you find out about from 1. The newspaper? 2. Friend? 3. The state employment service? 4. Private employment services? n. 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 each of the following: 1. The newspaper? 2. A friend? 3. The state employment service? n. 3.1 How many. . . . . (n.n)
Select questions	1.1 ... 2.1 ... 3.1 ...

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.

For example, avoid questions that require the interviewee to perform "audit work" to answer, that is, consult records or other information sources. If used at all, such questions should be reserved for mail questionnaires.

Other questions (or the manner in which 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, which do not belong in a GAO study, as well as questions whose sole purpose is to embarrass the interviewee (testing or questioning the intelligence of the interviewee or seeking information about private habits).

If needed, ask such questions in a mail questionnaire, where confidentiality or anonymity can be granted.[3] Also avoid questions that could cause unnecessary confrontation, causing 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 using questions that have no answers and 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 respondents will not have a basis to make a judgement 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.

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:

---

[3]See our discussion on confidentiality and anonymity, pages 42 and 43.

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 formulation of more specific questions. For a small number of respondents and where analysis may be qualitative, rather than quantitative, open-ended questions also may 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[4] should be done before attempting to tabulate.

A question that actually is closed can be structured in such a way that to the interviewee it appears 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. The following question illustrates the technique:

Why weren't you satisfied with the plan? (DO NOT READ CHOICES; INSERT 1 = MENTIONED, 2 = NOT MENTIONED)

/__/ 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. For example:

---

[4]Discussed in Content Analysis: A Methodology for Structuring and Analyzing Written Material, PEMD Methodology Transfer Paper, June 1982.

1. Who completed your last performance appraisal? _____

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

3. 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. For example:

Have you ever served in the U.S. military?

/__/ 1. Yes

/__/ 2. No

Scaled-response questions

In the scaled-response question, you read or show to the respondent 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, as discussed below.

- Balanced scales. The end points of the scale are usually adjectives or phrases with opposite meanings, e.g., very satisfied and very dissatisfied. As its name implies, the balanced scale contains an equal number of responses on each side of a neutral response. For example:

How satisfied or dissatisfied are you with the typing ability of the secretaries in your division?

/__/ 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. Use the unbalanced scale when no negative response is possible. Intensity ranges from none to great, for example:

On your last assignment, to what extent, if at all, were you given the opportunity to help develop staff working for you?

/__/ 1. Very great extent

/__/ 2. Great extent

/__/ 3. Moderate extent

/__/ 4. Some extent

/__/ 5. Little or no extent

● Rating and ranking scales. In a rating question, the respondent 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 numerical scale is:

Using a 7-point scale, where 7 is the highest rating, rate each of the individuals listed below on their ability to do GS-12 Evaluator work:

/__/ Brown

/__/ Green

/__/ Johnson

/__/ Martin

/__/ Smith

For an example of a verbal scale, refer to the adjectives used in the GAO BARS personnel-rating system (exceptional, superior, fully successful, borderline, and unacceptable). 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 respondent is asked to place items in order according to a specified criteria. For example:

Rank the following individuals on their overall ability to do GS-12 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.

/__/ Brown

/__/ Green

/__/ Johnson

/__/ Martin

/__/ Smith

Ranking questions may have several types of instructions. You can ask the respondent 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 respondent 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 respondent's opinion as to the distance between items (e.g., 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 respondent to choose between or among seven or less 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 respondents 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. A final order can be produced by averaging the ratings over all respondents.

* * *

When preparing a list of cues (alternate responses) consider their number and order, as discussed below:

- Number of cues. The number of cues depends on the type of respondent and type of analysis desired. There is a physical limit, generally, to the number of cues to which a respondent 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, however, as 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 generally eliminates a middle or neutral point on the scale and forces the respondent to commit to a positive or negative feeling. 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 respondent 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 the case 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 respondents are biased to answer the question in a particular way. Consider a situation where 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 respondent would be:

1. Very great extent
2. Great extent
3. Moderate extent
4. Some extent
5. Little or no extent

In this example, the order of presentation biases the choice slightly in favor of the program. Some respondents 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.[5] To repeat, the bias is slight. But since it does exist, we use the logic of biasing the question against the hypothesis we are examining.

Unscaled-response questions

In this type of question, a list of cues is read or shown to the interviewee, who is asked to choose one from the list or to

---

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

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 (but thought to be rare) answers or if it is thought that some respondents will come up with unique answers. For example:

> Educationally, what is the highest level that you have achieved? (CHECK ONE FROM THOSE LISTED.)
>
> /__/ High school graduate
>
> /__/ Some college
>
> /__/ BS or BA degree
>
> /__/ MS or MA degree
>
> /__/ PhD
>
> /__/ 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 establish the rapport[6] essential to a successful interview. 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 also should 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 finally the sensitive areas covering the attitudes of the interviewee.

Present the questions in a logical manner, keeping the flow of questions in chronological or a 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.

---

[6]Establishing rapport is covered in more detail in chapter 8, "Conducting Interviews."

LAYOUT CONSIDERATIONS

The layout or form of the printed DCI (see figure 1 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. Here are some considerations when designing the DCI:

- Type face. 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 1, the text to be presented to the interviewee is presented in upper and lower case, the instructions in upper case.

- 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.

- Provide open-top boxes for the interviewer to record numerical answers to all but open-ended questions. Place the box in a standard place beside each question to aid the interviewer and to facilitate editing, keypunching, and subsequent analysis of completed questionnaires.

- Card column numbers are placed under the box.

- 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.

- Numbering and spacing of questions are used to facilitate their reading.

Figure 1

Structured Interview Text (Example)

CURRENT STATUS (CS)

Now I'd like to find out what you are doing.

123. /__/ (38) Are you now receiving any AFDC?
IF YES - Is this a full grant or reduced grant?

1. Yes - Full grant (GO TO Q 124)
2. Yes - Reduced grant (GO TO Q 124)
3. No (GO TO Q 129)

124. /__/__/ (39-40) What is your status with WIN? Are you registered in training or what?
(LISTEN. INSERT COMMENTS AND TRY TO DETERMINE WHAT CODE TO ASSIGN - IF NECESSARY CHECK RECORDS OR CHECK WITH WIN STAFF AFTERWARDS)
Comments: _____

_____

_____

10. Working registrant status
11. Part-time employment
15. Working nonregistrant
20. Institutional training
30. Work experience
31. WIN/OJT
32. WIN/PSE
33. Suspense to training
34. Suspense to employment
40. Intensive employability services
41. IES/Group job seeking activities
50. Other WIN noncomponent activity
60. Unassigned recipient

125. /__/ (41) Are you looking for work (different work)?

1. Yes (GO TO Q 126)
2. No (GO TO Q 130)

126. How are you going about looking for work?
(DO NOT READ CHOICES; INDICATE 1 = MENTIONED, 2 = NOT MENTIONED)

/__/ (42) (1) On my own

/__/ (43) (2) Through WIN

/__/ (44) (3) Through CETA

/__/ (45) (4) Through Employment Services (ES, SES)

/__/ (46) (5) Through private employment agency

/__/ (47) (6) Other (specify) _____

127. /__/ (48) To what extent are you having difficulty finding a job?
(READ)

1. Very great extent (GO TO Q 128)
2. Substantial extent (GO TO Q 128)
3. Moderate extent (GO TO Q 128)
4. Some extent (GO TO Q 128)
5. Little or no extent (GO TO Q 131, P 19)

128. /__/ (49) To what extent is WIN helping you with these difficulties?
(READ)

1. Very great extent (GO TO NEXT PAGE)
2. Substantial extent (GO TO NEXT PAGE)
3. Moderate extent (GO TO NEXT PAGE)
4. Some extent (GO TO NEXT PAGE)
5. Little or no extent (GO TO NEXT PAGE)

CHAPTER 4

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 basically is dictated by the written, structured data-collection instrument. The DCI is prepared in advance and pretested and the interviewers trained to use it; thus, to some extent, the appropriateness of the language has been tested. It is the task of the interviewer to faithfully transmit to the respondent the meaning of the questions. In addition to precisely wording the questions, 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, "Conducting Interviews."

- 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, i.e.:

- 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" or "PSE" in a question may be nothing but alphabet soup to the respondents; 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 respondents use the terms daily.)

Thus, to speak over the interviewee's head hinders communication. The interviewee who is embarrassed at his or her 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 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 interviewee feels you regard him as ignorant. This approach is demeaning. You have contacted this individual because he or she has important information to impart. To treat the person condescendingly--or let it appear that is the case--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.

USE OF QUALIFYING LANGUAGE

After composing an interview question, you may find it requires an adjective or qualifying phrase added or a time-frame specified to make the item complete or to give the interviewee sufficient or complete information. For example:

1. How many employees do you have? might become

   How many full-time-equivalent employees do you have?

2. 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 on the DCI, however, another interviewer may qualify in a different way. This could make the resulting data difficult to summarize and analyze.

Also, the interviewee, not realizing that qualifying language is absent, may answer the question as he or she interprets it. Thus, different interviewees would be responding to different questions, based on their own interpretation.

CLARITY OF LANGUAGE

The style in which the 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 respondent to understand, especially if it is presented orally, which makes it difficult for him or her to review parts of the question. While the interviewee may be responding to one part of the question, the interviewer interprets it 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:

1. Delete "that" wherever possible, e.g., "Others suggest [that] training can be improved."

2. Use plain language, e.g., for "aforementioned," use "previous" or "previously mentioned."

3. Avoid the passive voice. Substitute pronouns, e.g., "I," "we," or "they" and active verbs; instead of "It is necessary to obtain..." use "We need..."

● Double-barreled questions. These are classical examples 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 respondents, 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, e.g.:

Indicate which of the organizational goals listed below are not considered unattainable within the 2-year period.

Reword 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. Where "yes" or "no" answers are expected, the results can be misleading. For example:

Are all of your employers covered by medical insurance?

If one employee is not covered, a "yes" answer is impossible. A better question would be:

About what percent of your employees are covered by medical insurance? or

What portion of your employees are covered by medical insurance? (READ THE CHOICES)

1. All or almost all

2. More than half but not all

3. About half

4. Some but less than half

5. None or hardly any

There are cases when the use of "all" or "none" is appropriate, but they are few.

- Defining terms. Where possible, define key words and concepts used in questions; e.g., when speaking of "employees" define and clarify the term. Are we talking about part-time, full-time, permanent, temporary, volunteer, white-collar, blue-collar, etc.? 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 will evolve. You may wish to list definitions in a separate section or on a card to hand respondents for reference.

BIAS WITHIN QUESTIONS

A question is biased when it causes the interviewee to answer in a way that does not reflect his or her true position on the issue. The interviewee may or may not be aware of the bias. Problems result when the interviewee is

- unaware of the bias and influenced to respond in the way directed by the wording; or
- aware of the bias and either deliberately answers in a way that does not reflect his or her opinion; or
- refuses to answer because it is biased.

Bias can appear in the stem (the statement portion) of the question or in the response-alternative portion.

Also, bias may result when a question carries an implied answer, because 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; for example:

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," although they have not subscribed. The questions could be restated as follows:

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 respondents could be unaware of any bias in the wording. Such bias usually occurs when additional qualifying or identifying information is added to the question, for example:

Which plan is more acceptable to you: the one designed by Ralph Brown, our chief economist, or the one designed by Paul 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: Ralph Brown's or Paul Green's?

- Bias resulting from unequal choices. When response alternatives are created, it is important that they appear equal. If undue emphasis is given to one, it may be easier for the respondent to select that one, for example:

Who do you feel is most responsible for the poor quality of the training program?

1. Instructors

2. Counselors

3. High-paid managers who run the centers

Alternative 3 is isolated from the other 2 because of the words "high-paid," which sets those individuals apart from the others, and the fact that alternative 3 is longer than the others. A better wording would be:

1. Instructors who teach the courses

2. Counselors who advise which courses to take

3. Managers who run the centers

• Bias due to 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 due to lack of balance. When using a scaled question, avoid bias in the stem, as well as in the response alternatives. For example, a question that seeks to measure satisfaction with something should mention both ends of the scale in a balanced fashion. For example:

How satisfied were you with the answers you received?

/__/ Extremely satisfied

/__/ Very satisfied

/__/ More satisfied than not

/__/ Neither satisfied nor dissatisfied

/__/ Not satisfied

Rather, reword this as follows:

How satisfied or dissatisfied were you with the answers you received?

/__/ Very satisfied

/__/ More satisfied than not

/__/ Neither satisfied nor dissatisfied

/__/ More dissatisfied than not

/__/ Very dissatisfied

(Proper use of an unbalanced scale was discussed on page 14.)

* * *

In summary, to design 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. For more information, consult Bradburn, 1981, or Sudman, 1982 (complete bibliographic details appear in appendix II).

CHAPTER 5

PRETESTING AND EXPERT REVIEW

Pretesting[1] and expert review is perhaps the least appreciated phase in the development of a structured interview. In the desire to meet deadlines for getting the job done, staff may question: "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. The pretest and expert review processes are part of that input, giving the study team 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--a serial process that occurs after each pretest or group of pretests.

PURPOSE OF PRETEST

In pretesting, we test the data-collection instrument 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

1. the right questions are being asked to obtain the needed information,

2. the content or subject matter of each question is releevant to the respondent, and the respondent has the knowledge to answer the question, and

3. the wording and procedures used in conducting the interviews are adequate to insure that valid and reliable results are obtained.

PRETEST PROCEDURES

For the typical structured interview, plan between 8 and 12 pretests. 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

---

[1]**Pilot and pretest are not interchangeable terms.** "Pretest" is usually used in connection with the testing of a structured interview or questionnaire, while "pilot" implies a test of most or all of the complete study design at one field location before proceeding to implement the design at all selected locations.

during actual data collection. It is important to test as many of the procedures involved in conducting a structured interview as possible, including selection of and contact with the interviewees. Pretests should be conducted in the same mode to be used in the actual interviews, i.e., the face-to-face interview pretested in person and telephone interviews over the telephone. Questions that work in person may not work by phone.

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, i.e., 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, e.g., 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.

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.

Contact pretest interviewees by phone or in person to arrange a pretest session. Identify yourself, describe what kind of agency GAO is and what it does, and explain the nature of the study. Indicating that this is a pretest, ask the interviewee to participate. Arrange to meet in a place convenient to the interviewee and free of distractions. (For a more detailed explanation and copies of text to be followed, see chapter 7, "Selecting and Contacting Interviewees.")

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

1. carry on the normal questioning of an interview without too much interruption in the flow,

2. provide explanations or try alternative wordings when the interviewee cannot or will not answer a question, and note the changes introduced,

3. record the answers on the DCI so the recording procedure and coding technique can be tested,

4. make notes on situations that occur during the interview that indicate problems with the instrument or procedures, and

5. 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 task 2, providing explanations or alternative wording must be done carefully, as 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 adequately rephrase the question and can make a slip in wording that favors or is slanted toward your approach to the situation (bias).

Identifying problems

Problems can be identified during the pretest session through your observation as interviewer, or by questioning the interviewee afterwards during the debriefing.

Major indicators of problems occur when the interviewee does one of the following:

- refuses to answer,
- exhibits nervousness, fidgets, wrings hands, or lights cigarette,
- exhibits inconsistency in answers from question to question,
- exhibits inconsistency by reflecting back and changing answers,
- takes too long to answer a question,
- gives a different answer than one listed on the instrument or known to the interviewer,
- does not understand a question, or
- asks for questions to be repeated or reworded.

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.

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

- 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 already have been expended. In other cases, PEMD or experts outside GAO can provide expert review.

People providing expert review are not acting as interviewees. They do not answer the questions, but 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 used with each version. Changes that obviously are needed can be made and the revised version used in the next pretest. This allows us to use a relatively more perfect version on each round of pretests.

CHAPTER 6

TRAINING INTERVIEWERS

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

TRAINING METHODS

GAO uses various mechanisms to train its interviewers and help 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. These are 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 data-collection instrument. 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.

First, the evaluator in charge (EIC) and the measurement specialist review 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 feel could arise. The trainees can ask questions, clarify items, catch typographical errors in the DCI, and suggest possible changes based on 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 are 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 presented to assure standardization of meaning and delivery and elimination of bias. Even though the instrument is pretested, some interviewees still will have trouble understanding the language. The interviewer must know enough about the question that rewording it for clarity will not violate its intent.

- 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 their 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 assure

that proper sampling is maintained. The log makes it possible later to adjust the universe and examine possible effects of nonresponse.

- 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.

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 EIC or measurement specialist should be notified, so the items can be changed or clarification given to all interviewers.

Field practice

Once an evaluator is in the field at the first site, he or she 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 are not used in the final analysis, regardless of whether the interview went well or poorly. Interviewing real interviewees who do not count gives the interviewer a chance to get rid of any anxiety and test out his or her 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 EIC or the measurement specialist.

Supervisory field visits

Normally, the EIC makes field visits during the course of an evaluation. A visit early in the data-collection process when

interviewing has just begun is valuable, allowing the EIC to review the procedures being used to conduct the interviews and observe some interviews first-hand. This quality-assurance checking enables the EIC 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.

Supervisory telephone contacts

The EIC 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.

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.

CHAPTER 7

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 the informed consent process 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 (e.g., state officials responsible for welfare programs), you need no selection process. 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 PEMD's forthcoming methodology transfer paper on statistical sampling.

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 2, which contains rules that GAO used to review a nationwide program requiring interviewing of program participants.

When contacting the interviewee by phone, use a standardized approach. This assures that you do not omit any important information. An example of such an approach is presented in figure 3. Naturally, if unexpected events occur, you may have to deviate from this guide.

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 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 4 and how it looks completed in figure 5.

Figure 2

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.

Figure 3

Telephone Contact with Potential Interviewee (Example)

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.

Figure 4

INTERVIEWEE CONTACT LOG

Job title____________________ Universe size________ Unit no.___

Job code____________________ Sample size____ Page ____ of ____

Auditor____________________

Ran-dom no.	Name address phone number	Case no.	Attempts to contact & results			
			Date & time	Result	Date & time	Result

Note: Attach completed interview forms to these logs.

Figure 5

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

Random no.	Name address phone number	Case no.	Attempts to contact & results: Date & time	Result	Date & time	Result
36	James Nolob 123 Line Ave. City, ST 04321 432-9999	—	2/4 3:00p 2/4 8:00p 2/5 9:30a 2/5 12:15p	No Answ. No Answ. No Answ. No Answ.-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 9:00a 2/5 1:00p 2/6	Out shopping-will call back Tel. set for 2/6 10a. Interviewed		
80	Grace Store 770 Green St. City, ST 04321 432-6069	2	2/5 1:30p 2/5 2:30p 2/6 1:00p 2/6 8:00a. 2/6 4:00p	No answ. Int. set for 2/6 1:00p. No show Agreed to 2/6 4:00p. Interviewed		

Note: Attach completed interview forms to these logs.

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, 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 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 farm.

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 the interviewee to omit negative comments about the office and its personnel out of fear this information would be overheard and affect his or her benefits.

When interviewing people in their homes, you may encounter frequent interruptions from other family members, neighbors, and telephone calls. Television and radio programs also can 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.

PROTECTING THE INTERVIEWEE

You may encounter interview situations that result in the interviewees speaking of themselves or others in a negative manner. This could come from asking questions on such sensitive issues as personal habits or behavior, attitudes (e.g., political, 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 content 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, as described below:

- Confidentiality means that the evaluator could associate the interviewee's name with specific responses but promises not to do so. For details of pledges of confidentiality, refer to GAO's General Policy Manual, pages 7-20 through 7-23.

- 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.

CHAPTER 8

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 influence 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, that of the interviewee, to respond with answers. Actually, the interviewer must perform at least eight major tasks:

1. Develop rapport with interviewee and show interest,
2. Give the interviewee a reason to participate,
3. Elicit responsiveness from the interviewee,
4. Ask questions in a prescribed order,
5. Assure understanding,
6. Assure nonbias,
7. Obtain sufficient answers, and
8. Show sensitivity to interviewee burden.

These tasks, which are not isolated but must be integrated into the interview procedure, are discussed more fully 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 emphathetic, 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).

Make these 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.

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 or even work clothes. This makes sense in that it gives the interviewee the feeling that you understand the nature of the circumstances under which he or she works. 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, because of their positions, are obliged to cooperate with GAO and 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 interviewee may need to learn how to act as a respondent. 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.

ASKING QUESTIONS IN A PRESCRIBED ORDER

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.

ASSURING 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 repeat or rephrase the question or add a new question to obtain an adequate response. To maintain the meaning of the questions, do this "probing" with care. These kinds of probes should be worked out during the pretest.

ASSURING NONBIAS

We have covered bias in the way a question is written or in selection of interviewees in earlier chapters. There can be bias also in the way you pose the content 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 assure 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 have 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.

CHAPTER 9

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 that data. All the hard work and time expended will have been worth the cost.

Had you taken the easy way, however, and gone with an unstructured interview, the situation might well be very different. Typically, an unstructured interview will contain many open-ended questions, which are not asked in a structured, precise manner. Different evaluators will interpret questions differently, vary the manner in which they ask questions, and often offer different explanations when respondents ask for clarification.

Answers to open-ended questions may range from a few words to several sentences. Respondents will typically give the interviewer some salient ideas that come quickly to mind, but will leave out some important factors. Open-ended questions do not help respondents consider an identical range of factors. Interviewers, after conducting several interviews, 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 the interviews have not produced uniform data that can be compared, summed, or further analyzed to answer the evaluation or audit questions.

On the other hand, if you, with the help of specialists from a DMTAG or PEMD, have followed the procedures outlined in this

transfer paper, 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 respondents did not understand or had trouble with questions. The pretests you conducted and the expert review you obtained assured 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, as 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 the boxes discussed on page 19. You can have the data keyed directly to disks, magnetic tape, or punch cards, then entered directly into computer files. 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 people from a universe to conduct interviews, you intend that all of the sample be interviewed. Indeed, it is part of the GAO analysis plan that what these selected people 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.

To combat this 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 people 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, if a small number of substitutions are made, this will have 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 people died, they may represent the older people 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

During the time 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; e.g., 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 recipient's 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 have a sample of 100 people that you have interviewed, the sample was random, and 20 people who could not be contacted were replaced by the next 20 people on your random list. Your results show that 57 people said "yes" and 43 said "no." You are now asked about the 20 people 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 voted "no." This would make the final outcome 57 "yes" and 63 "no." Under these circumstances, the finding would not have support. In other words, a 57-to-43 split of the data with 20 uninterviewed people is too close to make a decision. Had the split been something like 65 to 35, your case would have stood a chance.

Some of these methods require collection of additional information; with each method, there are assumptions and limitations that can influence the eventual interpretation of the data collected during the structured interview. Nothing short of obtaining the interviewee's answers to the questions will be fully satisfactory.

DATA ANALYSIS

The edited data now resides 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. Here are some of the levels of analysis that you might consider:

- First-level analysis. Here you concentrate on a description of the data, i.e., 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, "2." Any other number showing up for this question would be due to 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 numbers and percentages shown on the frequency tabulations' run must be weighted before making 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 analyses, 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 shows that a larger percentage of women complete college than do men, is the difference statistically significant, or could it be due to the fact we took 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 are not due to 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 measure 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 research effort. Selecting the most appropriate is not easy. Once again, we strongly advise that evaluators and auditors obtain the assistance of a specialist from the DMTAG, the regional Technical Assistance Group, or PEMD for this phase of the assignment.

CHAPTER 10

ROLES 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. The function each staff member performs during the evaluation appears in table 3. In addition to data-collection tasks 1-17, which are discussed in this paper, the table includes for completeness seven tasks that concern data handling, analysis, interpretation, and reporting.

Use the table as a guide when you must assemble personnel at various times during development and use of the data-collection instrument. 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 exclude some persons from participation in that task.

Table 3

Functions of Evaluators and Specialists
During an Evaluation Using the Structured Interview

Task	Staff functions[a]			
	Evaluator in charge and staff	Measurement specialist	Sampling specialist	Data analyst
1. Formulate overall questions	Initial work	Acquires 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	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

[a]A staff member may perform more than one function during a given task in the course of an evaluation.

Table 3 (Continued)

Task	Staff functions[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 instruction on sampling	[optional]
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	Participat
19. Edit raw data	Staff edit	Consult	-	Direct
20. Keypunching	-	-	-	Contract arrangemen
21. Edit keypunch 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

[a]A staff member may perform more than one function during a given task in the course of an evaluation.

GLOSSARY

Bias
: As used in this paper, words, sentence structure, attitudes, and mannerisms that unfairly influence a respondent's answer to a question. Both interviewer and instrument bias can exist.

Closed question
: A question that has a set of possible answers from which one or more 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
: When used in the context of questions appearing in a structured interview, the alternative responses that increase or decrease in intensity in an ordered fashion. The interviewee is asked to select one as his/her answer to the question.

Data-collection instrument (DCI)
: A highly structured document that requires the user/respondent to collect/provide data in a systematic and highly precise fashion.

Demographic questions
: Questions used to compile vital background and social statistics, such as age, marital status, size of household, etc.

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-type or short answer or fill in the blank. During an interview, the respondent would give the interviewer an unstructured, narrative answer. The interviewer would record the responses verbatim or select salient features. If a structured interview is used, a question may appear open-ended to the interviewee but can be "closed down" by the interviewer who has a set of alternative answers to check.

Probe
: In an interview, to examine a subject in greater depth, using additional 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 quantitative analysis for comparison.)

Quantitative analysis
An analysis that ascertains the magnitude, amount, size, etc., 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 qualitative analysis for comparison.)

Reliability
The extent to which the same answer to a question can be obtained when asked of the same person at another time.

Stem
The statement portion of a question.

Structured interview
An interview conducted using a data-collection instrument designed specifically for that purpose. The structured interview stresses use of closed questions as opposed to open-ended questions, prescribed (standard) wording, and a prescribed sequence of questions; collection of data from a number of people and summarizing of that data; and use of imbedded instructions and definitions by the interviewer.

Target population
The level (item, individual, group, organization, etc.) at which data are collected. Data can be collected at the individual level (e.g., program participant) and analyzed and reported at the organizational level (e.g., the employment office). Participants can be asked how many hours of counseling they received from the employment office, while the analysis of the data reflects 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.

APPENDIX II

BIBLIOGRAPHY

Babbie, E. R. 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.

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

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

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

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.

U.S. General Accounting Office, Program Evaluation and Methodology Division. Content Analysis: A Methodology for Structuring and Analyzing Written Material, Methodology Transfer Paper. Washington, D. C.: 1982.

-----. Designing Evaluations, Methodology Transfer Paper. Washington, D.C.: 1984.

-----. Statistical Sampling, Methodology Transfer Paper. Washington, D. C. (in preparation).