

## QUESTIONNAIRE DESIGN RESEARCH LABORATORY

Patricia Royston, Deborah Bercini, Monroe Sirken, and David Mingay  
National Center for Health Statistics

### INTRODUCTION

The National Center for Health Statistics (NCHS) recently established a Questionnaire Design Research Laboratory (QDRL). The QDRL serves as a resource for developing and testing questionnaires for NCHS surveys and other federally sponsored surveys. While many other statistical agencies and survey organizations conduct questionnaire design research, the QDRL is the first to emphasize research on the cognitive aspects of survey measurement and to incorporate the methods of cognitive science in a laboratory setting.

Field pretesting, the questionnaire testing method most often used by NCHS, provides adequate information on operational and logistical aspects of the survey, such as question flow, skip patterns, respondent burden, and field procedures, but it provides little information on respondents' recall and response processes. For example, some questions encourage guessing, such as questions which impose a complex recall task unsuitable for a fast-paced personal interview, or questions containing unfamiliar terms. The interviewer may routinely get quick, reasonable-sounding responses to such questions, with no indication that the responses are invalid. Other questions may contain vague or ambiguous wording, requiring the respondent to mentally select an interpretation of the question before answering. The QDRL applies cognitive research techniques to the study of these and other recall and response problems.

The concept of the QDRL was first conceived as the result of the Advanced Research Seminar on the Cognitive Aspects of Survey Methodology, known as "CASM", that was conducted by the Committee on National Statistics under a grant from the National Science Foundation [1]. Subsequent to the seminar, the National Science Foundation provided funding to NCHS to conduct a demonstration of how the knowledge and techniques of cognitive science could be applied to improve survey instruments [2]. The conclusion of the demonstration project was that "It is feasible, productive and efficient for a Federal statistical agency to conduct laboratory and statistical research on the cognitive aspects of survey questionnaires." This conclusion, coupled with the 1985 decision by the Office of Management and Budget to limit the number of full-scale field pretests that could be conducted for NCHS surveys, led NCHS to establish the QDRL in the fall of 1985.

### MISSION OF THE QDRL

The QDRL was established as a resource for the NCHS and other Federal statistical agencies in the development and testing of survey instruments. The overall mission of the QDRL is to improve the quality of Federal statistics by improving the validity and reliability of survey instruments. The unusual feature of the QDRL is the emphasis on the use of cognitive research techniques. Several methods frequently used by psychologists in memory and cognition research

are being applied to the investigation of recall and estimation strategies commonly used by survey respondents. The wording of survey questions is thought to affect the types of strategies used by respondents, which in turn affects the accuracy of their responses. The laboratory attempts to identify the most appropriate recall and estimation strategies, and to devise question approaches to promote the use of such strategies by respondents.

When fully staffed, the core QDRL staff will consist of a mathematical statistician (part-time), two survey statisticians, a cognitive psychologist, and a lab manager. When the QDRL conducts studies for NCHS surveys, a fully collaborative arrangement is employed in which the survey staff is involved in developing the research plan, conducting the research, and analyzing the results. From time to time, the staff will be further supplemented by scientists who will consult on applied research problems, and by visiting scientists.

### RESEARCH METHODS

The major responsibility of the QDRL is to assist the NCHS survey staff in the development and testing of their questionnaires. In the past, the typical NCHS questionnaire development procedures consisted of (see Figure 1):

- (1) defining survey objectives and data elements,
- (2) researching topics and drafting questionnaires,
- (3) informally testing questions, (e.g., asking questions of friends and co-workers)
- (4) conducting one or two full field pretests, and
- (5) making final revisions in the questionnaire in preparation for printing.

The QDRL contributes to several stages of the questionnaire development and testing process (see Figure 1). During the earliest planning stages, the QDRL staff is involved in discussions of survey objectives and the definition of data elements so that they can insure that the final questions satisfy the objectives. (These discussions typically continue throughout the questionnaire development process, as research findings indicate which objectives need clarification or revision.) Following these early discussions, the Lab staff arranges interviews with small numbers of respondents to begin to explore the respondents' understanding of concepts and their ability to report data items.

Two methods, focus interviews and free and dimensional sorts, are most appropriate for this work, although other methods can be used.

Focus interviews are unstructured discussions of the subject with individuals or groups. These discussions are led by a trained moderator who encourages the respondents to discuss their reactions to the subject. This approach may be used to gain insights into the

most promising question approaches, or to determine whether it is feasible to ask questions on that subject matter at all. It is helpful to conduct focus interviews with various demographic subgroups to observe the amount of variation in knowledge and attitudes.

Sorting procedures are used to study the natural taxonomies into which people group lists of items such as foods or medical conditions. In the free sort, respondents sort lists of items into groups that "seem to go together." In dimensional sorts, the respondents are asked to rank each item in the list along several scales which are expected to help explain the free sort results. In a recent project in which this method was used to develop categories of chronic conditions, the scales included seriousness of the condition and amount of pain associated with the condition, among others. Sorting procedures are likely to prove particularly helpful in organizing checklists of items in questionnaires.

The results of these exploratory interviews are used to develop draft questions. These questions are reviewed by QDRL staff to identify recall and response issues to be investigated in the lab. Question probes are developed for studying these issues. Interviews are then scheduled with a small number of volunteers to begin to explore the cognitive issues. A variety of methods is available for the exploratory interviews, depending on the cognitive issues to be investigated (see Figure 2). A few of the methods are described below.

Concurrent think-aloud interviews: These exploratory interviews are fairly unstructured; the interviewer uses the draft survey questions and probes as guides in exploring the respondent's thoughts on the topic. The respondent is asked to think aloud as he/she answers the questions. This technique is especially useful for studying recall and estimation techniques used by respondents. The interviews are usually audio taped, so that the interviewer can concentrate on probing the responses and can analyze their content - referred to as protocols - later.

Paraphrasing: The respondent is asked to repeat the question in his/her own words. This can be used effectively to determine whether a question is so complex that the respondent is missing important qualifiers, such as the reference period.

Retrospective think-aloud interviews: These are somewhat more structured than concurrent think-aloud interviews; the interviewer first administers the entire draft questionnaire and then asks respondents about the recall and estimation strategies they used in answering the questions. Respondents are encouraged to think aloud about how and why they responded as they did.

Confidence ratings: After answering selected questions, respondents are asked to rate the degree of confidence they have in the accuracy of their answers. This may indicate to what extent the respondent had difficulty formulating an answer to a question and whether he/she was guessing. Although there are often other clues that the respondent is having difficulty, this is a method for systematically assessing the extent of the problem.

Response latency measurements: Interviews are audio recorded and later reviewed to measure the length of time that elapsed between the question being asked and the respondent providing an answer. This will offer clues as to the complexity of the recall process required by the question, which may be associated with the accuracy of data collected. In the field, complex recall strategies may be replaced with guesses if respondents are anxious to end the interview or if the pace of the interview makes complex recall strategies seem inappropriate to the respondent.

Most of these methods are more time-consuming than the usual interview, so it is often necessary to divide the questionnaire into sections and to study only one section at a time, to avoid excessive respondent burden. After all sections have been studied using one or more of these methods, the questions are revised and combined into one questionnaire. The complete questionnaire is then evaluated in the lab, prior to a full field pretest. Two methods have been used to evaluate the questionnaire: an iterative process and formal experiments.

Iterative process: In some cases, certain questions or groups of questions may continue to cause problems for respondents. The survey designer may prefer to continue to work toward a useable set of questions by testing a version on a small number of people, revising the questions that continue to cause problems, testing the questions again, and so on.

Experiments: In other cases, there may be two or more question versions that appear to work fairly well in the lab, and there may be varying opinions as to which provides the most accurate answers. Hypotheses can be developed and the versions then compared in a controlled laboratory experiment. To analyze these experiments, it is useful to have either a validation source or some independent means for evaluating the accuracy of the responses. For example, in a recent experiment involving questions on fluoridation of home drinking water, we assumed that a positive response indicated that the term "fluoridation" was misunderstood because the area had an unfluoridated water supply.

For both the iterative and experimental procedures, the interviews are usually conducted using retrospective think-aloud interview

techniques, and the respondent is debriefed after all questions are asked. This results in an interview which is as realistic as possible, with information about the recall and estimation processes being collected after the interview has been completed.

Following the evaluation of the questions, the questionnaire is again revised and prepared for field pretesting. If the results of the evaluation interviews do not show one set of clearly superior question wordings, then a split panel approach could be used in the field pretest. If the field pretest indicates that certain questions need more work, then additional laboratory investigations should be initiated. These are likely to be small scale studies designed to address specific problems identified in the field test, and the approach will depend on the type of problem encountered. It may be necessary to begin with more exploratory methods, such as concurrent think-aloud interviews, or it may be sufficient to revise the questions and conduct a small number of evaluation interviews using the iterative procedure.

#### Selection of Laboratory Respondents:

Careful selection of laboratory volunteers is critical to the success of the laboratory studies. The right mix of respondents is needed to assure that all questions can be tested. If some sections are preceded by a screening question which skips out those persons without the required characteristic, then an adequate number of persons with the required characteristic must be recruited to test these questions. In many instances this can be done with a few simple screening questions when volunteers call in to schedule appointments. If persons with the required characteristics are fairly rare, then special recruitment methods must be developed. If the respondents are located only in areas at some distance from the lab, it may be necessary to move the lab temporarily. If respondents can be located through some central record system (e.g., victims of crime, persons with cancer, persons with pacemakers) it may be possible to make an arrangement with the record source to contact persons for the lab experiments.

#### **STRENGTHS AND LIMITATIONS OF THE LABORATORY**

The laboratory is particularly well suited for identifying many kinds of questionnaire problems. The think-aloud methods, for example, provide detailed insight into the respondents' understanding of concepts and ability to report data items. When respondents think aloud as they try to answer questions, problems with complex or vague questions or with terminology immediately become apparent. In a field pretest, however, the respondent may quickly make some assumptions about the meaning of such questions in order to be able to provide an answer, and the interviewer may have no indication of the problem. Laboratory respondents are generally highly motivated and give their full attention to the task; they have agreed to spend the hour or so in the lab, and often have volunteered because they are interested in the topic. Consequently, most respondents are very willing to provide detailed

insights into their thought processes, and the interviews are very productive. It also helps that the interviews are conducted in a controlled, quiet environment, so that the respondents are not distracted from the task by children, pets, the television, or the telephone.

The laboratory has several logistical advantages as well. Special populations can be recruited fairly quickly and efficiently (depending on the ease with which members of the population can be located). In general, laboratory studies are cheaper and faster to organize than field pretests; respondents come into the lab so travel costs are minimal and interviewer time per interview is greatly reduced. Training is less costly because the survey designers and lab staff do nearly all of the interviewing, and they are thoroughly familiar with the questionnaires. Complex experimental designs can be administered more easily than in a field test because of the closely controlled laboratory setting; for example, question variations and interviewers can be assigned using more rigorous methods.

The survey designers' time is also more efficiently used. Given ideal lab facilities with a one way mirror and/or video recording equipment, the survey designers and survey sponsors can observe the interviews inconspicuously to evaluate the questionnaire. As an economical alternative, audio recordings can be made, although nonverbal cues are then missed. Researchers can revise and test a questionnaire repeatedly in a short period of time; as problems with questions are identified, the survey designer can often make revisions within a few hours and test the revision on the next group of subjects. Interviewing can be terminated as soon as enough information is collected on the questionnaire's performance.

While the lab method of testing questionnaires has some advantages over field methods, field pretesting remains a vital component of the questionnaire development process. Occasionally laboratory results may be misleading. Complicated questions or questions which require more than superficial recall efforts may work well in the lab but fail in the fast-paced household interview. Laboratory respondents are volunteers who have agreed to spend a specified period of their day in the lab answering questions for a small fee. They usually give their full attention to the task. In household interviews, respondents are often contacted with no warning other than an advance letter, and are expected to interrupt their activities for an hour or more to answer questions. Consequently, lab respondents may spontaneously employ more thorough recall strategies when answering complicated survey questions, while household respondents may resist interviewers' efforts to encourage the use of those same strategies because of the time and effort required. (If the lab studies indicate that complex recall strategies are needed for some questions, the questions should be simplified before they are tested in the field.) Also, lab respondents tend to volunteer because they are interested in the topic, and thus do not reflect the average level of knowledge in the general population.

### CONCLUSION

The National Center for Health Statistics is revising its traditional methods for developing and testing questionnaires. Heretofore, NCHS has relied completely on field tests to identify questionnaire problems. Laboratory methods which employ the techniques of cognitive research are now being used to test and improve questionnaires prior to field testing. This report describes the current procedures of the NCHS laboratory; it is anticipated that with experience, these procedures will be refined and new procedures will be developed.

### REFERENCES

- [1] Jabine, Thomas J., Miron L. Straf, Judith M. Tanur, and Roger Tourangeau. "Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines." Washington: National Academy Press, 1984.
- [2] Lessler, Judith T. and Monroe G. Sirken. Laboratory-based Research on the Cognitive Aspects of Survey Methodology: "The Goals and Methods of the National Center for Health Statistics Study." Milbank Memorial Fund Quarterly/Health and Society, Vol. 63, No. 3, 1985.

FIGURE 1

## THE QUESTIONNAIRE DEVELOPMENT PROCESS

### TYPICAL NCHS PROCESS

### CONTRIBUTION OF THE LABORATORY

Define survey objectives

Develop detailed list of data items

-----  
Research topics

Study respondents' understanding of concepts and ability to report data items  
(*Methods 1-3*)

Look for previously used questions on same topics

Draft questions

-----  
Informally test/revise questions

Study respondents' interpretation of questions and recall strategies (*Methods 3-7*)

Evaluate questions through experiments and testing (*Methods 5-9*)

-----  
Conduct formal field pretest, revise questions

Evaluate revised questions (*Methods 5-8*)

-----  
Format and print questionnaires

FIGURE 2

## LABORATORY METHODS

- 1. Focus interviews**  
Unstructured discussion of the topic with individuals or groups.
  - 2. Free and dimensional sorts**  
Respondent sorts lists of similar items into groups that go together or ranks the items according to specified scales.
  - 3. Concurrent think-aloud interviews**  
Respondent thinks aloud when answering questions—responses are probed extensively.
  - 4. Paraphrasing**  
Respondent repeats question in his/her own words.
  - 5. Retrospective think-aloud interviews**  
Respondent answers all questions first, then is asked how he/she arrived at the answers.
  - 6. Confidence ratings**  
Respondents rate degree of confidence they have in the accuracy of their answers.
  - 7. Response latency measurements**  
The time between the question being asked and the respondent answering is recorded.
  - 8. Iterative process**  
One version of questions is tested, revised, tested, revised, for several iterations.
  - 9. Experiments**  
Formal hypothesis testing is carried out using multiple questionnaire versions.
- 

FIGURE 3

### ***Strengths of the laboratory:***

Laboratory methods provide more detailed insight into the respondents' understanding of concepts;

Special populations can be quickly recruited and tested;

Researchers can revise and test the questionnaire repeatedly in a short time;

It is cheaper and faster to organize than a traditional field test;

Questionnaire designers and survey sponsors can observe interviews inconspicuously, e.g., through one-way screen or videotape;

Complex experimental designs can be administered more easily than in a field test;

Respondents are generally highly motivated;

Interviewing can be terminated as soon as enough information on the questionnaire's performance is collected;

The respondent is less likely to be distracted or interrupted in the controlled laboratory setting.

### ***Limitations of the laboratory:***

Some questionnaire flaws are missed in the laboratory because:'

Laboratory respondents are more motivated than the average survey respondent and therefore willing to tackle more complex recall tasks;

The laboratory interview is more relaxed than an interview in the field, which may result in different recall strategies being used;

People who volunteer for laboratory studies often do so because they are interested in the topic and thus do not reflect the average level of knowledge in the general population.