QUESTIONNAIRE DESIGN RESEARCH IN A LABORATORY SETTING: RESULTS OF TESTING CANCER RISK FACTOR QUESTIONS

Patricia Royston and Deborah Bercini
National Center for Health Statistics

INTRODUCTION

In October 1985, a questionnaire research facility, called the Questionnaire Design Research Laboratory (QDRL), was established in the National Center for Health Statistics. The QDRL serves as a resource for NCHS and other Federal statistical agencies in the development and testing of survey instruments. The QDRL is the first permanent questionnaire design laboratory to apply research methods commonly used by cognitive psychologists in their studies on memory. (A complete description of the methods employed in the QDRL can be found in an earlier paper [1].)

The first assignment of the QDRL was to develop and test a questionnaire on knowledge, attitudes and practices relating to cancer risk factors (CRF). The CRF questionnaire was to be administered nationwide in 50,000 households as part of the 1987 National Health Interview Survey (NHIS). The purpose of this paper is to present selected findings from the laboratory testing of the CRF questionnaire which illustrate the usefulness of the laboratory approach.

The laboratory tests were designed to complement, rather than to duplicate, the usual NHIS field pretest. Traditionally, the field pretest has been used to identify problems with question flow, skip patterns, transitions between sections, questionnaire length and other overt questionnaire flaws. Laboratory studies, on the other hand, address problems with respondent comprehension of the questions and ability to recall the information. When comprehension difficulties arise, the laboratory attempts to determine the source of the problem, such as an overly complex question structure, unfamiliar terms, vague question concept and so on. Difficulties with recall are investigated by asking respondents about the recall strategies they use. The extent to which respondents employ estimation strategies rather than direct recall and enumeration of individual events, indicates that the question demands information that is not readily accessible in memory or beyond the cognitive processing abilities of respondents in some way. Judgements can then be made about whether the estimation strategies are resulting in responses of an acceptable level of accuracy.

METHODS

The development of the CRF questionnaire was a collaborative effort, with the National Cancer Institute (NCI) as the study sponsor, the Survey Planning Branch of the Division of Health Interview Statistics (DHIS) as the primary survey planner, and the QDRL as advisor. The CRF questionnaire included questions on medical care, cancer screening exams, knowledge and attitudes regarding cancer risk factors, tobacco use, occupational exposure, dietary intake, family history of cancer, reproduction and hormone use, and relationships and social activities.

The development of the CRF questionnaire included five major stages: definition of the data requirements; development of draft questions; laboratory testing of the draft questions in the QDRL; field pretesting of the questionnaire; and final laboratory testing of revisions made after the pretest. These stages were not discrete; definition of the data requirements, for instance, continued throughout the development and testing phases, as sponsors became aware of some discrepancies between what they wanted to learn from the survey and what was possible to collect in a household interview. Consequently, draft questions were continually added, deleted, or revised.

In preparation for the first stage of exploratory interviews, the questionnaire was divided into eight sets of questions. The divisions were made so that similar topics were grouped together, questions requiring special respondent groups were grouped together, and the subsets of questions were approximately equal in terms of respondent burden. The sets of questions were mutually exclusive, with minor exceptions.

Specifications were then developed for the demographic and/or behavioral characteristics of the respondents needed for each section. Volunteer respondents were recruited by means of notices posted in libraries, clinics, stores, and other public places. Volunteers called a telephone number given on the notice and were scheduled for interview. During this initial call, a few items of information (age, sex, race, income, smoking history) were collected in order to assure a reasonable mix of respondents and to assign the volunteer to the appropriate segment of the questionnaire. In some cases volunteers were not scheduled if several respondents with similar characteristics had already been interviewed.

Interviews were conducted in the QDRL facility in Hyattsville, Maryland. In order to derive the maximum amount of information from each respondent, the DHIS and QDRL survey design staff served as interviewers. Respondents were paid $15 for their time and transportation expenses.

At the outset of the interview, respondents were told that the objective was to detect flaws in the questionnaire, rather than to collect data on the respondent. They were instructed to comment on any questions that seemed unclear or difficult to answer for any reason. The first round of interviews were conducted in a "think-aloud" mode, in which respondents were instructed to verbalize their thoughts as they answered questions, to shed light on their interpretations of the questions and the response strategies used in formulating answers. Because many respondents found it difficult to think aloud as they answered, responses were probed extensively to get at the cognitive aspects of the response process. The following are examples

of draft CRF questions and the probes used to get at respondents' thought processes concerning the questions.

CRF QUESTION: How old were you when you first started smoking cigarettes fairly regularly?

PROBES: How did you remember how old you were? Are you remembering how old you were when you first smoked or how old you were when you first smoked regularly? What does "fairly regularly" mean to you?

CRF QUESTION: Where do you get the most useful information about health care? (Respondent is handed card with response categories listed.)

PROBES: What does "most useful" mean to you? (If needed, add:) Does it mean "believable" or "information that you would act on?"

CRF QUESTION: In your opinion, what are the major causes of cancer in this country?

PROBES: What does "major causes" mean to you? (If needed, add:) Does it mean "causes the most cases of cancer" or "is most likely to give me cancer" or something else?

CRF QUESTION: During the past year, how often did you usually eat/drink (list of foods)?

PROBES: (For seasonal foods, ask:) How did you estimate the annual frequency? (For foods asked in a group, ask:) Did you estimate how often you ate each one and combine the answers, did you think about them all together, or did you get the answer some other way? (If estimated separately and then combined, ask:) How did you combine the foods that you eat often with those that you rarely eat?

At the conclusion of a round of interviews, the interviewers met to discuss the problems they had observed. Question revisions that did not alter question concept or intent were made on the spot. Problems which required major revisions in the data collection approach were noted for discussion with the sponsor. (In some cases, the sponsor chose to accept the potential response error in order to maintain comparability with other studies or to avoid additional respondent burden.) When possible, problem questions were revised for further testing.

RESULTS

The think-aloud and probe approach revealed a variety of comprehension and recall problems with the CRF questions. Some questions were extremely complex and consequently difficult to comprehend, some appeared simple but required complex recall strategies, some contained unfamiliar terms or ambiguous words, and some were so broadly worded that they elicited conditional responses. For this paper, five questions were selected which illustrate some of the these cognitive problems. The original questions, cognitive problems, and solutions are shown below.

QUESTION 1

During the past year, how often did you usually (eat/drink) -

(a list of 56 foods and food groups, including:)

Orange juice or grapefruit juice?

Other fruit juices or fortified fruit drinks?

Beans, such as baked, pinto, kidney beans, or in chili? Do not include green beans.

Carrots, or mixed vegetables containing carrots?

Tomatoes, including in salad?

Spinach?

Coleslaw, cabbage or sauerkraut?

French fries or fried potatoes?

Potatoes, baked, boiled or mashed?

Chicken or turkey, baked, stewed, or broiled?

Spaghetti, lasagna, or pasta with tomato sauce?

Vegetable soup, vegetable beef, minestrone, or tomato soup? Do not include other kinds of soup.

Ham or lunch meats?

Peanuts or peanut butter?

Cognitive Problems

This series of questions requires an extremely complex recall and estimation process. Since it is highly unlikely that respondents will actually recall each individual instance of consuming a particular food, they must employ a variety of estimation strategies based on typical consumption patterns and "intuitions". In the household interview setting, respondents may be discouraged from employing the estimation strategies needed to provide accurate answers, because of the time that process would take. The wording of the question also suggests that an exact response is expected. A few single food items may be relatively simple to estimate (fried chicken), but most of the items above require respondents to consider:

- several food items together, e.g., "Coleslaw, cabbage, or sauerkraut"; "Spaghetti, lasagna or pasta with tomato sauce"; "Vegetable soup, vegetable beef, minestrone or tomato soup"
- different modes of preparation, e.g., spinach or potatoes or tomatoes, which can be eaten in salad, casseroles, or alone as a cooked vegetable
- or both grouped foods and mode of preparation, e.g., "Chicken or turkey, baked, stewed, or broiled"; "Peanuts or peanut butter"; "Ham or lunch meats".

Solution

A simple solution is to "decompose" the question, that is, to ask separately about each of the

foods. The resulting increase in accuracy, however, must be weighed against the resulting increase in respondent burden. In this case, it was not feasible to decompose the items in the food list, because the respondent burden was already excessive. Instead, a memory aid in the form of a flash card was used to simplify the respondent's task by showing him/her what would be asked in a particular food category (e.g., fruits, dairy, meats, etc.) so that he/she would know where to report certain foods. The stem of the question was changed to "During the past year or so" to suggest to the respondent that typical patterns were being asked for rather than numerically exact responses.

QUESTION 2

If a person stops smoking completely, do you think his/her chances of developing the following conditions will ever be as low as that of a person who has never smoked?

- Heart disease
- Cirrhosis
- Cancer of the lung
- Cancer of the bladder
- Emphysema

  . . .

Cognitive Problems
This question is difficult to comprehend because of the multiple qualifiers that the respondent must keep in mind when formulating an answer. The question contains at least six separate pieces of information that precede the list of conditions:

If a person stops smoking completely,
do you think
his/her chances of developing
the following conditions
will ever
be as low as that of
a person who has never smoked?

Even before the interviewer reads the first condition, the respondent may be overwhelmed by the task of absorbing and understanding the intent of the question. Some respondents will ask for the question to be repeated; others will assume, correctly or not, that they understand the general intent of the question and provide an answer.

Solution
The concept being measured was simplified. The redesigned question was, "Do you believe that if a person stops smoking completely, his chances of getting (condition) are reduced?" The respondent is asked only if he/she believes that smoking cessation reduces risk. If it had been important to retain the comparative element, a second question could have been asked of those respondents who answered "yes" to the redesigned question: "Would this person's chances of getting (condition) be as low as that of someone who has never smoked?"

QUESTION 3

In your opinion, which of the items on this card can cause cancer?

- Exposure to the sun
- Cigarette smoking
- Air pollution
- Water pollution
- Some cloth dyes
- Eating the wrong kinds of foods
- Exposure to people with cancer
- Viruses

  . . . (Other environmental factors)

Cognitive Problems
1. Respondents had difficulty answering the question because it ignores the hereditary component; a common response was "some people will get it from this (risk factor) and some won't - it depends partly on heredity." Because the question seemed to imply that all people would be affected equally by the listed risk factors, respondents were distracted from the intent of the question.

2. Respondents often felt that the causal relationship between these items and cancer did not always hold, although they believed that exposure to these factors would increase one's risk of getting cancer.

Solution
The question was reworded, "Which of these things do you think increases a person's chances of getting cancer?" The category, "inherited make-up or heredity" was added at the beginning of the risk factor list.

QUESTION 4

I'm going to read a list of foods. Please tell me if you think any of these foods, when eaten regularly, increase or decrease a person's chances of getting certain cancers.

- Bran cereals and muffins
- Doughnuts and pastries
- Whole grain bread
- Lowfat and skimmed milk
- Whole milk and cheese
- Popcorn
- Bacon and eggs
- Fresh fruit
- Fried foods
- Orange juice and citrus fruits

Cognitive Problem
The hidden assumptions underlying this question are that (1) the respondent believes there is a relationship between diet and cancer, and (2) the respondent has knowledge of the nutritional components of foods. Laboratory interviews

showed that neither of these assumptions were true. Some respondents stated that they personally did not believe that diet and cancer were causally related, or if there was a link, it was relatively unimportant. Other laboratory respondents answered the question willingly, often providing the "correct" responses. However, probing revealed that many respondents were guessing on the basis of a general notion of what foods were "good for you" and what foods were not. They were not selecting foods because they knew that a high fiber/low fat diet was associated with lowering cancer risk.

Solution

It was clear that knowledge of the relationship between eating certain foods and cancer risk could not be adequately assessed in a single question. The following series of questions was developed to measure the different levels of respondents' beliefs and knowledge on this topic.

- A question to ascertain agreement or disagreement with the concept that diet can reduce the risk of developing major diseases.
- Respondents who agreed with the concept were asked to name the major diseases that they thought were related to diet (unaided recall).
- Respondents who agreed with the concept but did not mention cancer specifically in the unaided recall question, were asked directly if they believed cancer was related to diet.
- Respondents who mentioned cancer in either the unaided recall question or direct probe were asked two unaided recall questions on what foods people should more of and less of to help prevent cancer.
- All respondents were asked to identify foods high in fiber and foods high in fat from a list of foods.

QUESTION 5

a. Have you ever had a blood stool exam?

(For most recent exam:)

b. Were you told the results of the blood stool exam? (If yes) How were you told the results of the exam?

Cognitive Problems

a. This is one of several questions which ask about cancer screening tests, including breast physical exams, digital rectal exams, mammograms, and proctoscopic exams. Many respondents were unfamiliar with the terms, and used a response strategy based on "If I don't know what it is, I probably haven't had it." Because these exams are often part of routine physicals, this strategy could result in substantial underreporting.

b. Respondents were unsure whether to take the term "told" literally. Often, they were notified of the results by postcard, or told at the time of the exam that if they were not contacted, the results were normal. Thus, a "no" response could mean that they were not told verbally, or they were not told at all. This question was only intended as a screener for the question "How were you told the results of the exam?" which would not be asked if the response to the screener was negative.

Solutions

a. A short description of each exam was read before the question was asked.

b. The question was reworded to include examples of methods of notification, as follows: "How were you told the results of the test - in person, over the telephone, through the mail, or some other way?" "Never told; meaning results normal" and "Never told; DK if problem" were included as precoded response categories, and the screener question was deleted.

CONCLUSIONS

Although the Questionnaire Design Research Laboratory can be used to conduct more basic experimental research on the cognitive aspects of response error, the development of the CRF questionnaire illustrates the contribution of the laboratory approach when applied to specific questionnaire design problems. The intense focus on the response process in a relatively small number of interviews identified potential sources of error that could easily have gone undetected in a traditional field test.

By examining the ways in which respondents formulate their answers, that is, their assumptions about question intent and meaning, level of knowledge about the question topic, ability to recall the requested information, ability to retain complex question wording in short term memory, use of "guessing", and use of various estimation strategies, questionnaire designers can gain insights into the causes of response error and then develop or revise questions that will minimize the error. The laboratory setting permits repeated testing and revision of problem questions, giving researchers multiple opportunities to identify and correct flawed questions. When the source of error cannot be entirely eliminated because it is inherent in the nature of the data to be collected, researchers can use the information from laboratory interviews to make informed decisions about whether the likely level of response error is low enough to justify retaining that approach to obtaining the information.

Laboratory testing of the CRF questionnaire resulted in both outcomes, with major consequences for the final questionnaire. Many questions were tested iteratively until a version that appeared to avoid the identified pitfalls was achieved. Other questions were left

relatively unaltered but the laboratory provided valuable information on the potential limitations of the data. In the case of the food frequency list, for example, the CRF sponsors were willing to accept responses based on partial recall and estimation. This information was "good enough" for their analytical purposes. The questions on the relationship between diet and cancer, on the other hand, underwent many testing and revision phases until a version that adequately measured the various components of this concept was developed. The data to be collected on this topic were going to be used to design a national public education campaign on diet and cancer. Thus it was critical that the questions produce responses that would accurately define the level of awareness and knowledge in the general population.

REFERENCE

[1] Royston, Patricia, Deborah Bercini, Monroe Sirken, and David Mingay. "Questionnaire Design Research Laboratory." 1986 Proceedings of the Section on Survey Research Methods of the American Statistical Association, August 1986. Pages 703-707.