

COGNITIVE TESTING PROPOSED QUESTIONS FOR PRAMS IN THE NCHS QUESTIONNAIRE DESIGN RESEARCH LABORATORY

Barbara Foley Wilson¹, Nedra Whitehead² and Karen Whitaker¹

¹National Center for Health Statistics and

²National Center for Chronic Disease Prevention and Health Promotion

Barbara Foley Wilson, NCHS, 6525 Belcrest Road, Room 915, Hyattsville, MD 20782

Key Words: Cognitive lab, PRAMS

Statement of the Problem

Ongoing surveys typically undergo periodic revision. New questions are proposed for a variety of reasons, from a variety of advisers, and in a variety of formats. Integrating the new questions into the established questionnaire creates many problems, some of which can be solved with studies of the cognitive aspects of survey design. Furthermore, when more questions are proposed than can possibly be added because of time constraints, the cognitive testing results help to inform the decision making process about which questions to include.

In 1999 the Reproductive Health Branch of CDC's National Center for Chronic Disease Prevention and Health Promotion asked the Questionnaire Design Research Laboratory (QDRL) at the National Center for Health Statistics to evaluate 107 questions proposed by maternal and child health groups for addition to the Year 2000 survey of the Perinatal Risk Assessment and Monitoring System (PRAMS).

Background

The Questionnaire Design Research Laboratory (QDRL) was established at the National Center for Health Statistics to conduct cognitive studies of health surveys. The goal of cognitive studies is to reduce response error in surveys by studying the way people process information. The cognitive processes include perception, attention, memory, thought, language and judgment.

Cognitive studies are based on a cognitive model of survey response described by Roger Tourangeau in a 1984 publication which posits four steps in answering a survey question:

1. Comprehension of the question
2. Retrieval from memory of relevant information
3. Judgment about what and whether to answer
4. Response in the requested metric.

There are a variety of techniques that are used in cognitive studies. These include:

- Unstructured interviews adapted from anthropology
- Focus groups
- Expert evaluation
- Vignettes
- Experiments

- Card sorting tasks show how people organize their knowledge.
- Behavior coding of the interview
- Think aloud interviews
- Probes

Think aloud interviews with probing was the method used in this study for testing the PRAMS questions.

PRAMS

PRAMS is a population-based surveillance system of maternal behaviors and experiences before and during a woman's pregnancy and during the early infancy of her child. PRAMS supplements data from vital records for planning and assessing perinatal health programs on a state level. Because PRAMS data are population-based, findings from data analyses can be generalized to an entire state's population of women having live births. Findings from analysis of PRAMS data have been used to enhance understanding of maternal behaviors and experiences and their relationship with adverse pregnancy outcomes.

PRAMS operates through a cooperative agreement between the Centers for Disease Control and Prevention (CDC) and states that have been awarded grants on a competitive basis. Currently PRAMS includes 22 states and covers about 42% of all U.S. births.

Each participating state uses a standardized data collection method developed by CDC. PRAMS staff in each state collect data through statewide mailings and follow-up with nonrespondents by telephone. Every month, a stratified systematic sample of 100-250 new mothers is selected from a frame of eligible birth certificates. Each sampled mother is mailed an explanatory letter that introduces the survey, followed by the 14-page questionnaire at two to six months after delivery. A second questionnaire package, and in most states a third, is mailed to those who do not respond. PRAMS staff telephone those mothers who do not respond to the mailed survey. Therefore the questions must be appropriate for both a Self-Administered Questionnaire and for an interviewer administered telephone survey.

Since its inception, the PRAMS questionnaire has undergone several revisions primarily to capture data on

Questionnaire and for an interviewer administered telephone survey.

Since its inception, the PRAMS questionnaire has undergone several revisions primarily to capture data on new guidelines or emerging issues concerning Maternal and Child Health, such as knowledge of periconceptual folic acid, and sleeping arrangements for infants. Changes have also been made to improve respondents' comprehension of questions.

Protocol for Cognitive Lab Testing

The methodological design of this proposed study is consistent with the design of typical cognitive testing research. The purpose of cognitive interviewing is not to obtain survey data, but rather to obtain information about the processes people use to answer survey questions as well as to identify any potential problems in the questions.

For the study to cognitively test the proposed PRAMS questions in the QDRL, new mothers were recruited through a newspaper advertisement. There were more than 100 responses and the lab manager was able to select mothers of various ages, races, education and socioeconomic levels.

Thirteen mothers were interviewed; seven were under 30 years of age and six were older; six were African American and seven were white; eight were married, two were separated, two had never been married and one would not say; six were employed; and seven had incomes less than \$30,000. They had from one to nine children.

The mothers came to the QDRL to be interviewed. Most brought their children. Interviews were observed by closed circuit TV and videotaped with the mothers' permission. They were given a \$30 incentive.

For the self-administered questionnaire (SAQ) nine mothers read the questions aloud and checked answers in a booklet. For the telephone interviews, four mothers were called from one lab room to another. For both, the interviewer encouraged think aloud comments from mothers using pre-scripted probes as described in Sudman, Bradburn, and Schwarz (1996). Meanwhile, the babies played, cried, ate, shrieked, slept, climbed on the table, banged on chairs, breast fed, and made engine noises. Although distracting, the children's behavior demonstrated to researchers what conditions would be for mothers answering PRAMS at home.

Findings: Problems and Resolutions

Willis and Lessler (1999) described the typical kinds of problems that turn up in cognitive testing of a survey questionnaire:

1. Reading: Interviewers have difficulty reading the question uniformly.

2. Instructions: Conflicting, inaccurate, or complicated instructions.
3. Clarity: Question is lengthy, awkward, ungrammatical, or has complicated syntax; technical terms are undefined; question can be interpreted in multiple way and its unclear what is to be included or excluded; reference periods are missing, not well specified, or in conflict.
4. Assumptions: Inappropriate assumptions are made, the question assumes constant behavior, or it contains more than one implicit question.
5. Knowledge and memory: Respondent may not know the information, may not have an opinion, or may not have formed an attitude; information may not be remembered, or computation of a frequency or rate may be hard and lead to response error.
6. Sensitivity/Bias: Content or wording of the question may refer to behavior that is embarrassing, private, illegal, or not socially sanctioned.
7. Response categories: There may be a mismatch between the question and the response alternatives. There may also be unclear terms, undefined words, vague categories, overlapping categories, missing categories, or an illogical order.

These seven categories of potential problems are kept in mind by QDRL staff when conducting interviews or analyzing results.

In addition to the potential problems described by Willis and Lessler, there may be problems of context where earlier questions effect a respondent's understanding. This can be particularly troublesome when questions are derived from a variety of surveys and juxtaposed in a new context.

General Findings Based on Willis-Lessler schema.

The results of the cognitive lab testing will be presented in the Willis-Lessler schema.

1. Reading and question presentation:

Ideally, the telephone questions for respondents who did not answer the SAQ would be identical to the paper version. In practice however, questions must be adapted to phone mode. Instead of seeing all the response alternatives at once, respondents must listen to them

presented serially over the phone and try to hold everything in memory while forming a response.

This problem cannot be completely resolved. PRAMS staff adapted the final telephone version to improve comprehension. For example, a long list of check-all-that-apply response options, were changed to asking Yes/No for each option. Also, phone interviewers are trained to read slowly and to repeat if respondents appears not to understand question.

A separate problem arose when a question is phrased so that explanatory material followed the real question. For example, “Since your baby left the hospital, has he or she had a well baby check-up by a doctor, nurse, or other health care provider? Don’t count the times you took your baby for care when he or she was sick or visits for WIC.” On paper, a respondent could see that there was more information after the question mark. On the phone, mothers would interrupt with the answer before the additional qualifying material was read.

2. Instructions: Conflicting, inaccurate, or complicated instructions.

With regard to the Instructions, the PRAMS questionnaire was well formatted so that most mothers followed skip patterns with ease. Occasionally the “Go to Question X” gave the wrong direction and some mothers were led astray. In early stages of questionnaire design this can easily happen when questions are moved around and renumbered, or when some mother takes a path that the questionnaire designer did not anticipate. Occasionally, too, mothers overlooked the instruction and answered questions that were not intended for them.

3a. Clarity: Question is lengthy, awkward, ungrammatical, or has complicated syntax.

Several questions were awkward or confusing. For example, “Since you delivered your new baby, **would** you have had someone to help you **if** you were tired and feeling frustrated with your new baby?” Mothers did not easily understand what the question was asking. They repeatedly reread the question. Some finally answered Yes or No, and some just gave up. As originally proposed, the question was one of four with the same stem. That context aided comprehension, but on its own, the question did not work. PRAMS staff solved the problem by putting the question back into the multipart question, where it appeared to work in other surveys. Another solution might have been to take the question out of the hypothetical framework and ask, “Since you delivered your new baby there may have been times when you were tired and felt frustrated with your baby. Did you have someone to help you when that happened?”

Another awkward question was, “How **often is it easy** for you to calm down your new baby when he or she

is crying or fussy.” Mothers would puzzle, “How often is it easy.....? They then saw the response categories from Always to Never and that helped define the question. It would have sounded more natural to ask, “How easy is it for you to calm down your new baby...?” offering the responsive alternatives of Very easy to Very hard.

3b. Inconsistent reference periods

Ideally reference periods should be kept constant or flagged for attention when they are changed. Because the proposed questions came from many different advisory groups, the reference periods in these potential questions varied a lot. Moreover, they even varied in wording when the reference period was the same.

For example, these were used:

- 3 months before you got pregnant.
- 12 months before you delivered
- At any time during your last pregnancy
- Before you had your new baby, did you ever
- During pregnancy
- During the 3 months before your most recent pregnancy
- During the 12 months before you got pregnant
- During the last three months of your most recent pregnancy
- During the past 12 months
- During the three months before you became pregnant
- During your most recent pregnancy
- Last 3 months of pregnancy
- Past 2 years
- Since you delivered your new baby
- Since you delivered your baby
- Since your new baby was born

Recall of past behavior is a memory challenge and is always subject to response error, but shifting time frames certainly add to the cognitive burden. In the lab, mothers were asked to explain how they remembered whether something happened during the 12 months before they became pregnant. One explained a series of significant life events that she used as benchmarks, but after much explanation she realized that her calculations were a year off. Most mothers gave vague descriptions and indicated that they did not take the time frame very seriously when they formulated their answers.

PRAMS staff resolved this problem by changing the questions so that the reference periods are constant, when possible, and that the same terminology is used. They also italicized the time periods so that respondents would be alerted to pay close attention.

4. Implicit unwarranted assumptions

A survey question may make unwarranted assumptions that made it hard for mothers to answer. For

example, the question, “How often is it easy for you to calm down your new baby when he or she is crying or fussy?” This question assumes constant behavior, when in fact, crying and being fussy is highly variable, according to the mothers. It varies from day to day (“He has good days, and bad days.”); as the child ages, (“It’s better now.”); with the mother’s experience (“Depends on whether it’s your first or third child.”); and with the number of helpers available. The mother of nine children said that her older children spend so much time with the baby that “He doesn’t have a chance to get up a good cry.”

Based on these findings, PRAMS staff decided to delete this question.

A second example of implicit unwarranted assumptions was found in the question “During or since your most recent pregnancy, did a doctor, nurse or other health care worker refer you to any of the services listed below? For each thing, please circle Y (yes) if you were referred or N (No) if you were not.

- a. A class or program to help you stop smoking?
N Y
- b. A program to help you stop using alcohol or drugs?
N Y
- c. Counseling services to help you with a family or personal problem? N Y
- d. A women's shelter? N Y”

The question and its response alternatives contain an implicit assumption that the mothers needed such advice and good health care workers would be offering it. This assumption bothered some of the mothers. They could be saying No because they didn’t need the service or No because their provider failed to refer them.

Based on this information, this question was deleted.

5a. Knowledge and memory: Respondent may not know the information.

A question asked “Do you often eat fish caught in local rivers, streams, lakes, or reservoirs?” Most mothers said they don’t know where the fish that they buy are caught. One said Yes. When the Interviewer asked her to explain, she said that Safeway probably buys fish from the Chesapeake Bay, and for her that is local.

5b. Knowledge and Memory: Computation of a frequency or rate may be difficult.

For some questions computation of a frequency or rate lead to response error. Three questions asked about the number of cigarettes smoked in the last two years. Smokers miscalculated the number they smoked when they tried to answer in terms of number of cigarettes per day. They thought in terms of the packs of cigarettes and

they divided incorrectly when they tried to give the number per day.

This problem was resolved by giving respondents the option of answering in terms of cigarettes per day or packs per day.

6. Sensitivity/Bias: Content or wording of the question may refer to behavior that is embarrassing, private, illegal, or not socially sanctioned.

There were five questions about drinking alcohol with varying time frames: in the past two years, during the 3 months before getting pregnant and last 3 months of pregnancy. Most of the mothers reported that they did not drink and skipped out of the questions. The ones who reported drinking, chose very low response categories for the period before pregnancy and no drinking during pregnancy.

Another series of four questions asked about physical abuse--before pregnancy, during pregnancy, by partner, by someone other than partner. All of the mothers we interviewed responded negatively except one who reported, “He pushed me!” Most mothers responded calmly and negatively after briefly searching their memory. However, a few of the mother responded so slowly and tentatively that the interviewer had the impression that they were remembering something that might qualify for an affirmative response. Certainly mothers at home have to be concerned that their answers might be seen by their families, so they might not answer truthfully.

We think that the data yielded by both sets of questions may suffer from under-reporting. However without some system that assures privacy, such as audio-CASI equipment, this problem will be hard for PRAMS staff to resolve.

7. Response categories: Missing alternatives

In the following group of questions, some mothers had problems because they always wore seat belts, before, during and after their pregnancies. When they got to question c, they didn’t know whether to say No or Yes to “I wore my seat belts more often during my pregnancy than I usually do.”

Question: “Listed below are some questions about seat belts. For each question, circle Y (Yes) if it applies to you or circle N (No) if it does not apply to you.

- a. I was worried that wearing my seat belt during pregnancy would hurt my baby N Y
- b. I wore my seat belts less often during my pregnancy than I usually do N Y
- c. I wore my seat belts more often during my pregnancy than I usually do N Y”

To resolve this problem PRAMS staff decided to delete the second and third questions. Two other questions asked for frequency of seat belt use in the 3 months before pregnancy and during the last three months of pregnancy. The other questions were kept and were felt to adequately address this issue. The first question was kept and reworded to a stand-alone question.

8. Confusing context from series.

Question 42 followed up on the information asked in Question 41 and set up a context that confused mothers and resulted in response error :

Question 41. “How often does your new baby sleep in the same bed with you or anyone else?”

- Always
- Almost always
- Sometimes
- Rarely
- Never

42. How many other people sleep in the bed with your new baby?

- One
- Two
- Three or more

Several mothers answered Always, Almost always, or Sometimes to how often baby slept “in the same bed with you or someone else” for question 41. When they got to Question 42 they understood “How many other people” to mean “other than you the mother”. They wanted to have a zero category. One mother even added a zero box and checked it.

Based on these concerns, the second question in this series was deleted.

Discussion

The mothers who participated in cognitive testing of the PRAMS questions clearly were interested by the material and enjoyed working through the booklet of questions. They compared their experiences to the variety of possible answers presented by the response alternatives. They followed along with the implications of the questions and inferred the purpose. When there were alternative versions of questions about a particular topic, such as wantedness of the pregnancy, they paused and compared, trying to be clear about whether the questions were asking about something different, or just asking about the same thing in a different way. And in that case, they would try to decide which way was better at describing their particular experience.

Moreover, the mothers would work very hard to fit their responses into one of the alternatives provided, but that was not always possible. For example:

Question: “**Which of the following statements is truest of your husband or partner during the 3 months before you became pregnant?**

- He wanted me to get pregnant
- He partly wanted me to get pregnant and partly wanted me not to get pregnant
- He didn't care one way or the other whether I got pregnant
- He did not especially want me to get pregnant
- He wanted very much for me not to get pregnant”

One of our mothers had two older children and wanted a third, but getting pregnant required extensive infertility treatment. Her husband initially wanted the third child, but was unhappy about the treatments and felt sorry for his wife. He wanted to give up on the plan. When this mother tried to answer the question there was no response alternative that described her experience. “Partly wanted and partly wanted me not to get pregnant” did not describe her situation well. Unusual experiences like this are hard for a questionnaire designer to anticipate and often the resolution is to provide “Other--please describe.” PRAMS staff report that, in fact, mothers often write extensive notes on their booklets.

Conclusion

The task of adding new material to any ongoing survey is challenging because the question modules typically come from a variety of sources and have different structures, modes, time frames, and vocabularies that create mixed contexts. Laboratory testing with cognitive techniques is a good and relatively inexpensive way to discover problems that may cause response errors in survey data.

References

- Tourangeau, R. (1984). Cognitive sciences and survey methods. In T. Jabine, M. Straf, J. Tanur, & R. Tourangeau (Eds.), *Cognitive aspects of survey methodology: Building a bridge between disciplines* (pp. 73-100). Washington, DC: National Academy Press.
- Sudman, S., Bradburn, N.M., and Schwarz, N. (1996). *Thinking about answers*. San Francisco. Jossey Bass.
- Willis, G. B. and Lessler, J.T. (1999). *Question Appraisal System BRFSS-QAS: A guide for Systematically Evaluating Survey Question Wording*. A report prepared for CDC/NCCDPHP/Division of Adult and Community Health Behavioral Surveillance Branch.